Social and communication impairments represent some of the key diagnostic characteristics of autism [3]. Individuals with autism may show delays in language acquisition, with deficits ranging from the complete absence of functional speech, to the existence of adequate linguistic knowledge that is coupled with impairment in the functional use of that knowledge [107,108]. It has been estimated that between 30 and 50% of individuals with autism never develop functional speech [90]. Even when language develops in these individuals, verbal communication is often restricted to the expression of instrumental functions, or simple labeling [106]. If a child remains nonverbal by the age of five or six, the prognosis for social skills and expressive language has traditionally been thought to be poor [22]. However, there is evidence suggesting that the acquisition of language after this “critical period” in autism is also possible [see 89, for a review].

The communication impairment in autism is believed to reflect a lack of understanding of the mind. Theory of mind refers to the ability to understand another person’s mental state, including their beliefs, intents and desires, as separate from one's own thoughts, experiences and behaviors [9,88]. This understanding of the mind typically begins in infancy toward the end of the first year of life, with the emergence of intentional communication such as joint attention (i.e., alerting one another to a stimulus via nonverbal means), simple requesting, and sharing. In particular, joint attention may reflect the child’s motivation to communicate, which is an important prerequisite for social interaction [101]. Theory of mind thus relates to the development of language and social communication because it underlies the fundamental ability to understand actions and intentions of others, and to communicate them effec-
tively. In individuals with autism, theory of mind deficits have been linked to both impairments in executive functioning [54,65,86] and communication difficulties [87,107]. This inability to understand others’ intentions and behaviors may help to explain why language is delayed in children with this disorder, and why a significant proportion of them never acquire language at all [108].

Research has demonstrated a relationship between joint attention and language development in children with autism. In one longitudinal study, children with autism were first assessed between 2 and 6 years of age, and assessed again approximately 8 years later [101]. The results showed that one of the strongest predictors for subsequent language acquisition and expressive language abilities was responsiveness to bids for joint attention at initial assessment. This finding highlights the importance of joint attention in predicting the language and communication deficits that are the hallmarks of autism.

Besides poor joint attention, the communication deficits in autism may be related to imitation difficulties. Imitation involves translating another person’s action into one’s own, and is also considered to be a precursor of language development [39]. Numerous studies have reported imitation deficits in autism [see 117 for a review] and more recent evidence has provided further support for this claim. For example, Vanvuchelen et al. [113] found that individuals with autism showed impaired performance in both gestural imitation and general motor skills, suggesting that their imitation deficits may be part of a broader perceptual-motor problem. Perra et al. [87] found that children with autism performed worse than other children (typically developing children and children with general developmental delay) on imitation and theory of mind tasks. Taken together, these studies indicate that imitation deficits in autism are associated with problems both in comparing self with others and in motor planning. One hypothesis, which we explore below, is the notion that these behavioral deficits may actually be attributable to a common neural mechanism.

1. Mirror neuron dysfunction and communication deficits in autism

Over the past decade, some researchers have proposed that mirror neuron dysfunction might underlie the behavioral manifestations presented in autism [e.g., 39,57,80,81,118]. The mirror neuron system (MNS) was first discovered by recordings in area 44 (Broca’s area) [29,34], an area of the inferior frontal gyrus that is considered to be a precursor of language development [39]. Numerous studies have reported imitation deficits in autism [see 117 for a review] and more recent evidence has provided further support for this claim. For example, Vanvuchelen et al. [113] found that individuals with autism showed impaired performance in both gestural imitation and general motor skills, suggesting that their imitation deficits may be part of a broader perceptual-motor problem. Perra et al. [87] found that children with autism performed worse than other children (typically developing children and children with general developmental delay) on imitation and theory of mind tasks. Taken together, these studies indicate that imitation deficits in autism are associated with problems both in comparing self with others and in motor planning. One hypothesis, which we explore below, is the notion that these behavioral deficits may actually be attributable to a common neural mechanism.

The idea of a mirror-like system in language processing was first posited in the “motor theory of speech perception” [70], well before the discovery of mirror neurons. According to this theory, speech perception relies heavily on observation of the articulatory (motor) gestures of the speaker (e.g., movements of the mouth, lips, and tongue), rather than the acoustic cues of speech sounds. To successfully process spoken language, these motor actions must be represented in the listener’s brain, so that the regions critical to speech production also become activated when the listener sees articulatory gestures. In other words, to attain a shared understanding, there should be a level of synchronicity in motor representation between the speaker and the listener. Another model of speech perception also highlights the importance of facial gestures and manual gesticulations in language comprehension [102].

The discovery of mirror neurons provided support for the involvement of the motor system in auditory speech perception. The shared representations of observed and executed actions in these neurons may serve as a foundation for our capacity to understand the experiences of other people, which is crucial to effective communication and social interaction. Accordingly, it has been hypothesized that an intact mirror neuron system might underlie normal language functions in humans [6,93], and that language comprehension may be achieved through action understanding and mental simulations of sensory-motor structures [10,33,93]. Speech perception is essentially a multi-modal experience in that the development of language typically occurs in the presence of facial articulatory gestures and body movements, rather than by acoustic signals alone [6,102]. Thus, the shared representations of mouth movements and visual and auditory perceptions within the mirror neuron system may help to strengthen the associations between objects and their names [81].

Evidence for the involvement of the motor areas in speech perception comes from neuroimaging studies with normally developing individuals. For example, functional MRI (fMRI) revealed that when participants read sentences containing motor words that were associated with the hand, leg, or head, the regions in the sensorimotor cortex that would normally be involved in the execution of that action became activated [46,47]. Similarly, listening to speech sounds activated speech production motor areas [119], and both seeing and listening to speech activated regions of the putative mirror neuron system (specifically, the superior temporal sulcus and inferior frontal gyrus). Furthermore, transcranial magnetic stimulation (TMS) studies showed an increase in motor-evoked potentials of the tongue when participants listened to speech [28], and that the excitability of motor regions underlying speech production correlated with activity in Broca’s area [115].

Behavioral studies have also provided evidence for the link between motor actions and speech production. For example, when participants observed the grasping of an object while they simultaneously pronounced a syllable, their lip apertures and voice peak amplitudes were greater when the grasped object was large (as opposed to small) [36]. A follow-up study found that the effect of action observation on speech production as a function of object size was more pronounced in children than in adults [38]. This finding indicates that the transfer of action from observed gesture to mouth movement may play a critical role during the period of language acquisition in children.

There is substantial evidence that non-linguistic visual or auditory stimuli can elicit activation in sensorimotor regions. Seeing well-learned actions and/or listening to their sounds produces activity in a network of premotor and parietal brain regions [e.g., 8,24,43–45]. When participants listened to musical passages that were similar to those they had previously learned how to play, the inferior frontal gyrus was activated [89]. Thus, this brain region
appears to be involved in auditory-motor mapping. Regardless of whether a particular action is heard, seen, or performed, a similar network within the mirror neuron system appears to underlie the representation of the same action.

The involvement of this *multisensory* and motor system is particularly evident in experts, such as musicians. Neuroimaging studies using voxel-based morphometry found evidence for structural brain changes such as increased gray matter volume in the inferior frontal gyrus in instrumental musicians compared with non-musicians [35,103]. In addition to parietal regions (supramarginal gyrus and superior parietal cortex) and more dorsal premotor regions, the inferior frontal gyrus constitutes part of a network of multi-modal sensorimotor integration regions. In fMRI studies, it has been demonstrated that practicing an instrument leads to the rapid establishment of an auditory-motor and visual-motor network with increased coherence in this network of brain regions [7,8]. Listening to music, reading musical notation, watching musical performances of pieces that one knows how to play, and actually playing that music, all appear to engage a network of brain regions related to the putative human mirror neuron system [83].

Given that the mirror neuron system is believed to involve both sensorimotor integration and speech representation, it is likely to underlie some of the communication deficits in individuals with autism spectrum disorder (ASD). Recent research has provided preliminary support for this hypothesis. Because individuals with high-functioning autism are over-represented in this literature, the work reviewed here should be viewed with caution when extrapolated to individuals with low-functioning autism, especially those who are nonverbal. Nonetheless, published research studies represent a good starting point for understanding the mechanisms underlying speech-motor impairments in individuals with ASD. For example, Nishitani et al. [78] used magnetoencephalography (MEG) and electromyograms (EMG) to measure brain activity and lip movement when individuals with Asperger's syndrome were asked to imitate orofacial gestures. Compared to controls, the ASD group had EMGs that lasted almost twice as long. Moreover, the ASD group showed weaker activations in Broca's area, which was delayed by 45–60 ms. Similarly, when imitating facial expressions inside the MRI scanner, children with autism showed decreased activity in Broca's area relative to controls [23]. Although behavioral performance between these two groups did not differ, Broca's area activity in the autism group correlated with severity of autism as measured by the social subscales of standardized tests [71,72]. In another study, Theoret et al. [110] used TMS and found that the level of excitability in the primary motor cortex during action observation was lower in individuals with ASD compared to controls. A structural imaging study found that compared to controls, individuals with autism had reduced cortical thickness (decreased gray matter) in regions of the mirror neuron system, including Broca's area, the inferior parietal lobule and the superior temporal sulcus [40]. Moreover, the extent of the reduction in cortical thickness in these regions correlated with the severity of communicative and social symptoms in autism.

Familiarity may also play a role in activating the MNS in individuals with autism. Using EEG, Oberman et al. [82] found that the mu suppression over sensorimotor cortex was greater when children with autism watched videos of actions performed by familiar individuals (their guardian, siblings, or themselves) compared to those performed by strangers. Typically developing children, on the other hand, had similar levels of mu suppression regardless of who performed the actions. Thus, the MNS in individuals with autism (although impaired) does, in fact, respond to observed actions, but only when the person engaged in the action has some personal significance. This finding is consistent with reports of improved social and communication skills when children with autism interact with a familiar, as opposed to an unfamiliar, individual [e.g., 60,66]. It also highlights the importance of family and establishing a level of familiarity in therapy of individuals with autism.

2. Music making as an intervention to engage the mirror neuron system and facilitate expressive language

As reviewed above, there is now growing evidence that links a dysfunctional or broken mirror neuron system (or related network) to the behavioral deficits in autism. The involvement of a sensorimotor system in language processing has received support from neuroimaging data showing motor activity during language tasks [e.g.,46,91], as well as from behavioral data showing modulation of motor performance during language processing [e.g., 13,19,37]. Given the important role that the MNS might play in the understanding of actions and processing of language, and the MNS abnormalities and communication deficits associated with autism, a treatment approach designed to engage the putative human MNS may have significant clinical potential. Development of such an approach is particularly important because at present, there appears to be no evidence-based intervention that consistently produces significant improvements in expressive language in individuals with autism [31].

Recent research has shown that representations of the mirror neurons can be altered by training. Using TMS, Catmur et al. [21] found that a relatively short period of incongruent sensorimotor training (performing index-finger movements while observing little-finger movements, rather than performing movements with the same finger observed) was sufficient to alter the expected pattern of mirror neuron responses during observation of the trained actions. This indicates that experience-dependent plasticity also exists in the mirror neuron system, a finding that is consistent with a large body of literature on brain changes following sensorimotor skills training [e.g., 26,35] and sensory deprivation [e.g., 76,97]. The fact that components of the putative MNS can be manipulated through sensorimotor training highlights the possible benefits of incorporating a motor component in the treatment of a disorder such as autism, that may be related to a dysfunction of the MNS.

Music making is one possible medium through which the putative MNS can be engaged. Music is a unique, multi-modal stimulus that involves the processing of simultaneous visual, auditory, somatosensory, and motoric information; in music making, this information is used to execute and control motor actions [98]. It has been suggested that because music making activities involving imitation and synchronization may engage regions of the brain that overlap with regions that presumably contain mirror neurons, music making activities may be particularly useful for the treatment of developmental disorders such as autism [83]. In their Shared Affective Mirror Experience model, the core idea is that music is perceived not merely as an auditory signal but as an intentional expressive motor act. As described below, our proposed intervention goes beyond music listening alone. Rather, it links and maps the perception of sounds with actions, including both manual and articulatory actions.

It has long been noted that children with autism thoroughly enjoy the process of making and learning music [e.g., 41,112,120]. Listening to music can evoke a great intensity of emotions in these individuals [2,49], who typically have difficulty processing emotions [54]. This positive response to music and music making may help children with autism engage and interact with others, thus allowing them to participate in activities that could facilitate the acquisition of social, language, and motor skills [116]. For example, research has shown that music-based activities facilitate the use of sign language and other nonverbal methods of communica-
tation in children with autism [20]. Verbal instructions that combine melodic and rhythmic patterns with visual cues result in better retention of the words being taught to children with autism [109]. Furthermore, learning through music has been found to improve joint attention behaviors and nonverbal social communication skills in children with autism, with some generalization to settings beyond the music therapy sessions [63]. Clearly, the idea of using music in a therapy context for individuals with autism is not new. However, most existing research involves small-scale case studies, unstructured music listening or music making, and the efficacy of music therapy has rarely been systematically investigated using a controlled design, properly evaluated outcomes or statistical analyses [59]. Furthermore, previous therapy methods have not been well informed by a neuroscientific understanding of the underlying disorder and the forms of music making used as a therapy.

In addition to a strong interest in music, individuals with autism also show enhanced music perception skills. In the first report of autism, Kanner [58] described the exceptional musical skills of several individuals. One notable example was an 18-month-old boy, who was able to discriminate among many symphonies. Subsequent investigations have also reported enhanced music perception abilities in autism. Relative to controls, individuals with autism have been shown to have superior pitch memory [e.g., 4,48] and pitch discrimination skills [e.g., 12,51,52]. Although no study has directly examined the prevalence of absolute pitch in this population, anecdotal reports do suggest unusual absolute pitch abilities [14,50], and it has been suggested that autism and absolute pitch may share biological markers [16].

The value of incorporating principles and practices of music making in the treatment of language disorders (such as autism) is reinforced by neuroimaging research showing overlapping responses to music and language stimuli [e.g., 67,68,85,100]. In particular, fMRI studies have reported activation of Broca's area during music perception tasks [67,111], active music tasks such as singing [84], and imagining playing an instrument [11,74]. Moreover, a common network appears to support the sensorimotor components for both speaking and singing [64,84,91].

From a therapeutic perspective, engaging in musical activities has been shown to improve verbal abilities in language-delayed children [56]. Given the overlap between the language and music systems in the brain, we propose that music making (through singing and/or playing an instrument) may provide an alternative medium for accessing and engaging this system. These activities can potentially enhance social interactions and communication skills in nonverbal children with autism. Evidence relating to this topic is discussed below.

Despite having either a reduced, or complete lack of ability to speak, many children with autism are still able to sing, and accurately reproduce complicated tunes and jingles from television commercials [120]. This dissociation between singing and speaking is strikingly similar to that observed in patients with Broca's aphasia, who are often able to sing the lyrics of a song better than they can speak the same words [e.g., 53]. One successful rehabilitative technique for restoring language function in patients with aphasia is melodic intonation therapy (MIT); more generally, MIT is a form of auditory-motor mapping training (AMMT). MIT emphasizes the prosody of speech through slow, pitched vocalizations [1,104], and engages an auditory-motor mapping network as well as sensorimotor feedback regions through the association of hand tapping and intoned vocal output [79]. MIT, when employed in an intensive course of therapy with patients who have chronic non-fluent aphasia, has been shown to be successful in generating propositional speech [e.g., 1,99,104].

The efficacy of an adapted version of MIT in autism has also been reported in a case study of a 3-year-old nonverbal boy [75]. After 35 sessions, he was able to combine words and could respond to intoned questions or statements. While the results of this study are very encouraging, it is possible that the therapy coincided with the boy's delayed language development. However, another case study of a 6-year-old girl with autism also reported the efficacy of singing in eliciting speech [55]. Thus, it is possible that an adapted version of MIT could provide a useful alternative to traditional speech therapy for children with expressive language impairment linked to autism. Although the nature of language deficits varies greatly among individuals with autism, these two case studies indicate a particular potential of an intonation-based speech production technique in assisting children who are nonverbal.

The application of MIT to the treatment of children with autism requires some modifications to the original procedure designed for the treatment of aphasia. Because the modified technique involves repeated trials of sound-motor mapping, it has been referred to as auditory-motor mapping training, or AMMT [114]. In this intervention, it is important to first establish a comfortable environment for children with autism. Each training session involves a vocalization procedure, during which the child is encouraged to vary the intensity and length of speech sounds. A series of picture stimuli (combined with signs or actual objects) that represent high frequency words, actions, and social phrases are then presented using a procedure adapted from MIT [79]. In addition to intonation, a key component of AMMT is the use of a set of tuned drums (or other tuned percussion instruments) to facilitate sound-motor mapping. The therapist introduces the target words or phrases by simultaneously intoning (singing) the words and tapping the drums on the same two pitches. The child progresses from passive listening, to unison singing, to partially supported singing, to immediate repetition, and finally to producing the target word or phrase on their own. Through intensive repetition in a comfortably structured environment that is often used in the context of therapy for autism, the previously nonverbal child learns to vocalize and possibly associate sound with meaning [114].

How can AMMT help facilitate the acquisition of language skills via regions of the brain that overlap with the putative MNS system? Evidence from prior research indicates that three components may be of particular importance: singing, imitation, and motor activity. Specifically, recent research has shown that compared to speaking, singing engages a bilateral fronto-temporal network more prominently, and that this network contains some components of the MNS [15,84]. Furthermore, an adapted form of MIT [79] can facilitate the production of expressive language in aphasic patients with left-hemisphere lesions by engaging a fronto-temporal network in the right hemisphere [84,99]. Moreover, the motor (hand-tapping) component of MIT may serve to engage a sensorimotor network that controls orofacial and articulatory movements [73]. This overlap may facilitate the auditory-motor mapping that is critical to meaningful vocal communication [69]. Additionally, the communication deficits of children with autism may be due to the oral motor speech deficits observed in language-delayed children with speech apraxia [77], thus highlighting the possible benefits of singing as an intervention [61].

Given the link between MNS and imitation, it has been suggested that imitation could be incorporated as part of an intervention program for individuals with autism. After repeated sessions of imitation, children with autism spent less time in gross motor activity [27] and more time initiating social interactions [27,30]. The benefit of imitation in expressive speech production is already evident in MIT (or any form of AMMT). The intervention involves intensive treatment sessions, which consist of multiple trials where the patient attends to the therapist's orofacial actions, and imitates the intoned phrases produced by the therapist [1,104]. Thus, the positive outcome of imitation observed in children with autism indicates that AMMT, in general, could be a promising intervention.
technique to facilitate acquisition and/or development of social and communication skills.

In addition to activating brain regions that overlap with the putative MNS in humans, AMMT can exert its therapeutic effects in other ways. For example, just as MIT includes the use of hand-tapping to promote engagement of the sensorimotor network in patients with aphasia [79], AMMT has an important hand-motor component which incorporates the use of drums, or any of a variety of other percussion instruments, in children with autism. The use of such instruments would serve the dual functions of promoting motor activity and capturing the child’s interest in the therapy. On one hand, the simultaneous engagement of a number of sensorimotor systems during AMMT has the potential to strengthen the connections between auditory and motor regions. On the other hand, the act of music making itself has the potential to facilitate social communication and interaction in children with autism because it exploits their strong interest in music as well as their positive response to it. The value of incorporating musical instruments within the therapy context has been highlighted by a number of studies. For example, Kern and Aldridge [62] incorporated singing and instrumental playing (e.g., drums, cabasa, and sound tubes) as an intervention for increasing social interaction in children with autism. Through interactive music making, the experimenter modeled turn-taking, appropriate use of body contact, and choice-making behaviors to the child. A peer was also introduced to the child to facilitate play and joint activities. Four children with autism participated in the study, and they all showed increased levels of peer interaction following 20–30 teacher-mediated sessions. A similar finding was also reported by Stephens [105]. In that study, musical segments filled with background music, dancing, instrumental playing (e.g., tambourines and maracas), and action imitation, were interspersed with non-musical segments during which the child would perform action-word imitation with the experimenter. All three children tested showed increased willingness to spontaneously imitate the action and word pairs after the experimenter imitated them during the musical segments. Thus, interactive music making using instruments can provide a useful framework for learning and for the development of social skills in children with autism.

As summarized in Fig. 1, individuals with autism have relative strengths and weaknesses, which may be related to functions and dysfunctions of the putative human mirror neuron system. An intervention or training (i.e., AMMT) that enhances interactions between the auditory and motor systems might represent an effective therapeutic strategy through which individuals with autism can develop their communication skills. Specifically, AMMT engages an action observation, hearing-doing network, which overlaps with components of the MNS. The important function of the MNS in action understanding and vocal production, coupled with the hypothesis that autism is linked with mirror neuron dysfunction, suggests that strategies that engage and stimulate brain regions that are involved in action observation and may be part of the MNS, could potentially ameliorate some of the associated communication deficits. In particular, therapies that incorporate elements of music making may offer a viable approach to help facilitate social skills, interactions with others, and communication, including expressive language in otherwise nonverbal autistic individuals. Interactive music making (using instruments) is useful in facilitating communication and social skills, while singing engages the MNS network that is believed to be deficient in individuals with autism.

Our proposed intervention relies on intentional involvement with the individual, as well as engagement of multiple sensory and motor modalities. These aspects are likely to improve learning independently of any direct involvement of the MNS. However, one important aspect of our intervention is the observation of an action that is coupled with sounds, which is designed to engage the MNS. Although it may be difficult, an ideal research design would test the relative contributions of these features of AMMT. For example, AMMT could be compared with a control treatment that is similarly engaging and multisensory, while lacking an imitation component, or lacking a hand-motor component. Such a comparison may help us to assess the importance of the MNS or the auditory-motor mapping system in inducing changes of speech output versus systems that mediate multisensory integration as well as intention and motivation. If proved to be effective, it is hoped that future research will focus on combining such methods of music making with speech therapy. This effort may ultimately lead to specialized treatments

Fig. 1. Graphical representation of the potential utility of AMMT in facilitating expressive language in individuals with autism via the mirror neuron system.
for autism that maximize the individual's potential for developing or re-learning expressive language function (e.g., through interactive AMMT), and thus, improve the quality of life for people with autism and their families.

Conflict of interest

The authors have no conflict of interest.

Acknowledgement

We sincerely thank the Nancy Lurie Marks Family Foundation for their support.

References


