Teaching receptive labelling to children with autism spectrum disorder : A comparative study using infant-directed song and infant-directed speech

Simpson, Kate, Keen, Deb and Lamb, Janeen

Simpson, K., Keen, D. and Lamb, J. (2015). Teaching receptive labelling to children with autism spectrum disorder : A comparative study using infant-directed song and infant-directed speech. *Journal of Intellectual and Developmental Disability*, 40(2), pp. 126-136. https://doi.org/10.3109/13668250.2015.1014026.

Journal of Intellectual & Developmental Disability, doi: **10.3109/13668250.2015.1014026**

Running header: *Teaching receptive labelling to children with ASD*

Keywords: *autism spectrum disorder, communication, children, intervention, music*

ORIGINAL ARTICLE

Teaching receptive labelling to children with autism spectrum disorder: A comparative study using infant-directed song and infant-directed speech

KATE SIMPSON[1,2], DEB KEEN[2] & JANEEN LAMB[1]

[1]*Faculty of Education and Arts, Australian Catholic University, Brisbane, Australia, and* [2]*School of Education and Professional Studies, Griffith University, Brisbane, Australia*

Correspondence: Kate Simpson, School of Education and Professional Studies, Griffith University, Brisbane, Qld 4122, Australia. E-mail: k.simpson@griffith.edu.au

Abstract

Background There is a growing body of literature investigating the efficacy of music interventions for children with autism spectrum disorder (ASD); however, little empirical research has been conducted into the use of musical elements to facilitate language learning.

Methods This crossover-design study compared the responses of 22 children with ASD (*M* age = 5.88 years) to sung and spoken instructions embedded into a computer-based communication intervention designed to teach receptive labelling.

Results There was no significant difference between the sung and spoken conditions. Following both conditions, there was a significant increase in receptive labelling skills; skills were generalised and were maintained at follow-up. A difference in group performance was found.

Conclusion Further research is required to investigate child characteristics that may impact on children's performance using this approach.

Introduction

Children with autism spectrum disorder (ASD) have impairments in expressive and receptive language that generally manifest early in their development, with abilities ranging from within normal limits to severe delays in development (Charman, Drew, Baird, & Baird, 2003; Luyster, Lopez, & Lord, 2007). Receptive language abilities, in particular, appear to be more impaired in comparison to expressive language abilities in some children with ASD (Hudry et al., 2010; Maljaars, Noens, Scholte, & van Berckelaer-Onnes, 2012). Early intervention to facilitate language learning among young children with ASD is recommended because of the strong associations between language abilities and later functional outcomes (Luyster et al., 2007). Identifying factors that may aid language abilities and ways to prevent or treat language impairments is therefore an important focus for researchers. One factor that appears to promote language acquisition in typically developing children is infant-directed speech. This study was conducted to determine if the addition of musical elements to this type of speech, as found in infant-directed song, may assist receptive language learning for children with ASD.

Infant-directed speech

Infant-directed speech is the ubiquitous speech caregivers use with infants that is characterised by higher pitch, wider pitch range, longer pauses, shorter utterances, and more repetitions than speech directed to adults (Fernald & Simon, 1984; Fernald et al., 1989). This speech is used by caregivers to attract the infant’s attention and communicate meaning, and is preferred by infants over other spoken stimuli (Fernald, 1985; Papoušek, Papoušek, & Symmes, 1991; Stern, Spieker, & MacKain, 1982). To date, there is limited empirical evidence identifying the salient acoustic features of infant-directed speech that may contribute to language acquisition/learning; however, infants appear to use the prosodic characteristics of infant-directed speech to process linguistic information. By 7 months, typically developing infants are demonstrating a listening preference for words over nonwords when presented using the prosodic characteristics of infant-directed speech (Thiessen, Hill, & Saffran, 2005). Not only does the use of infant-directed speech appear to facilitate the infant's recognition of words, it also impacts on the infant's long-term memory of these words and the infant's ability to generalise the recognition of these words (Singh, Nestor, Parikh, & Yull, 2009). At this age, infants' listening preference shifts from infant-directed speech to native infant-directed speech (Hayashi, Tamekawa, & Kiritani, 2001), suggesting they are attuning to the linguistic content in the speech rather than just the acoustics. Kuhl, Conboy, Padden, Nelson, and Pruitt (2005) identified a predictive correlation between 7-month-old infants' preference for native language speech and language scores tested at 18 and 24 months. Features of infant-directed speech prosody continue to be used to aid in word recognition by toddlers (Song, Demuth, & Morgan, 2010). Although the authors acknowledge further research is required, there is evidence to indicate an associative relationship between infant-directed speech and typical language acquisition. This may also be the case for children with ASD.

A number of studies have identified a concurrent relationship between the time spent by individuals with ASD attending to infant-directed speech and measures of language performance (Kuhl, Coffey-Corina, Padden, & Dawson, 2005; Paul, Chawarska, Fowler, Cicchetti, & Volkmar, 2007; Watson, Baranek, Roberts, David, & Perryman, 2010). Paul et al. (2007) identified that the listening time of toddlers with ASD to infant-directed speech was correlated to their receptive language abilities 1 year later. Consistent with typical developmental pathways, the associative relationship between the prosodic characteristics of infant-directed speech and language acquisition appears to extend beyond the prelinguistic stage of development for individuals with ASD. Santarcangelo and Dyer (1988) found teenagers with ASD and a developmental age at or less than 3 years were more responsive to directives when the adult used speech prosody typical of infant-directed speech compared to conversational speech. The infant-directed speech prosody was less effective with individuals functioning at a developmental level older than 5 years. These findings would suggest a potential association between listening to infant-directed speech prosody and language learning in children with ASD that may be dependent on the individual's developmental age.

Although the research reviewed here suggests a relationship between infant-directed speech and language learning for children with ASD, these children are generally less responsive to infant-directed speech than their typically developing peers, and this may contribute to language acquisition difficulties (Kuhl, Coffey-Corina, et al., 2005; Paul et al., 2007; Watson, Roberts, Baranek, Mandulak, & Dalton, 2012).

Infant-directed song

Infant-directed speech shares many features in common with infant-directed song (Trainor, Clark, Huntley, & Adams, 1997), but to date, research has not investigated whether infant-directed song promotes language acquisition in the same way as infant-directed speech. Given the listening preference for music by children with ASD (Blackstock, 1978), the use of infant-directed songs may hold some promise as a means of facilitating language learning. Although there has been little research investigating infant-directed song and language acquisition, there is a broader body of research investigating the relationship between music and language acquisition. There is some indication that the use of melody may heighten the infant's awareness of the lyrics. Many play songs used with infants have the same tune but different linguistic content, suggesting the linguistic content is important (Trehub & Schellenberg, 1995). Also, as seen with infant-directed speech, mothers modify their singing to the linguistic developmental stage of their audience (Bergeson & Trehub, 1999). The addition of melody to lyrics assists in detecting changes in lyrics (Lebedeva & Kuhl, 2010) and enhances the learning of both lyrics and melody in typically developing infants (Thiessen

& Saffran, 2009). These findings support the theory of a bidirectional relationship between music and language (Peretz, Radeau, & Aguin, 2004). Further, this learning appears to be more effective when the melody is familiar or heard repeatedly, thereby creating familiarity, and this association improves long-term memory of words (Rainey & Larsen, 2002; Wallace, 1994). It is yet to be investigated whether combining lyrics and melody may be beneficial for individuals with ASD.

Communication interventions using music

A number of studies have reported that individuals with ASD demonstrate a preference for auditory stimuli over other stimuli when the auditory stimuli are presented in the form of music (Blackstock, 1978; Kolko, Anderson, & Campbell, 1980). In addition, some children with ASD find music compared to nonmusic conditions more engaging (Carnahan, Basham, & Musti-Rao, 2009; Simpson, Keen, & Lamb, 2013).

To date, there has been no research investigating the use of infant-directed singing with children with ASD, but there has been limited research incorporating music elements into communication interventions. In a recent paper, Simpson and Keen (2011) reviewed experimental studies incorporating music elements in interventions with children with ASD, and although they identified some support for the use of music to target skills in interventions, only two studies specifically targeted language learning (Buday, 1995; Simpson & Keen, 2010). Buday (1995) targeted expressive language by comparing song and spoken conditions to teach spoken and signed words to children (4–9 years) with ASD. Learning was more effective and the children were viewed as more attentive in the song compared to spoken condition. In Simpson and Keen's (2010) study, receptive skills were effectively targeted by teaching picture names to young children (3–4 years) with ASD, incorporating a familiar melody "Old MacDonald" into a computer-aided intervention. These skills were maintained at follow-up, but there was little generalisation to other contexts

(Simpson & Keen, 2010). After publication of the Simpson and Keen review, Lim (2010) compared the use of an accompanied song to spoken instruction and no instruction on verbal production in young children with ASD. Children's verbal production improved following both instruction conditions compared to the no instruction condition.

These studies provide some support for the use of music in communication interventions for children with ASD. However, they have involved small numbers with only one study directly comparing sung and spoken conditions, and maintenance and generalisation data are seldom reported. Given the challenge children with ASD experience in generalising learning from intervention conditions to other materials and environments (Koegel, Camarata, Valdez-Menchaca, & Koegel, 1998), this is an important consideration in determining the effectiveness of intervention strategies. No studies have been identified that directly compare the effects of infant-directed song and infant-directed speech on aspects of language learning in children with ASD. In view of the associative relationship between infant-directed speech and language acquisition, the reduced responsivity to infant-directed speech individuals with ASD demonstrate, and their listening preference for music stimuli, infant-directed song may provide a means for enhancing language learning.

The aim of this study was to determine if music elements facilitate receptive labelling in young children with ASD. Learning outcomes for participants were compared following two intervention conditions: infant-directed song and infant-directed speech. It was hypothesised that performance in the infant-directed song condition would be superior as the music elements would facilitate receptive labelling. Generalisation and maintenance conditions were also included to address the limitations of previous studies in this area.

The following research question guided the study: Do children with ASD and severe language delay demonstrate an increased ability to learn picture labels using infant-directed song compared to infant-directed speech? In this event, we may conclude that musical elements facilitated language learning with this group of children.

Methods

Participants

Twenty-two children participated in this study. Participants were identified through state-funded special schools across Queensland, Australia, located within a maximum distance of 50 km from the researchers' base. School principals identified children for the study based on the inclusion criteria: (a) a confirmed diagnosis of ASD from a paediatrician, (b) aged between 3.5 and 8 years old, and (c) identified by the class teacher as individuals with limited communication skills, requiring augmentative and alternative communication support. Ethics approval for the study was granted by the Human Research Ethics Committee of the Australian Catholic University (Q2010 66) and from the Queensland Government Department of Education and Training. Participants were part of a larger study being conducted by the authors to investigate the use of musical elements to support learning in children with ASD.

The Social Communication Questionnaire (SCQ; Rutter, Bailey, Berument, Lord, & Pickles, 2003) was used as a measure of ASD symptomatology. The SCQ is a parent-completed questionnaire that is used as a screening tool for ASD (Eaves, Wingert, Ho, & Mickelson, 2006). All participants in the study exceeded the ASD cut-off score on this tool (≥ 15). The mean chronological age of participants was 5.88 (*SD* = 1.60, range: 3 years 9 months–8 years 9 months). The ratio of males to females was 3.4:1. Participants displayed severe delays in expressive and receptive language skills based on assessment using the Expressive Vocabulary Test, Second Edition (EVT-2; Williams, 2007) and the Peabody Picture Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007). All participants scored at or below the 0.1 percentile on the EVT-2 and PPVT-4, with the exception of two participants who scored 0.5 and 6 on the EVT-2. Both the EVT-2 and PPVT-4 are norm-referenced instruments that demonstrate robust psychometric properties (Dunn & Dunn, 2007; Williams, 2007). These two tests assessed expressive and receptive vocabulary attainment of participants and allowed for comparison of attainment levels of individual participants to their peers and to other participants. The Scales of Independent Behavior Early Development Form (SIBR; Bruininks, Woodcock, Weatherman, & Hill, 1996) is a test used to assess adaptive behaviour skills and problem behaviour in children functioning below 8 years. All participants scored below the chronological age equivalent of 4 years. There was a wide range in scores for maladaptive behaviours, with participants' behaviour described as marginally serious to very serious, requiring frequent to pervasive support. The participants demonstrated competency in following the task instructions and demonstrated no prior knowledge of the picture symbols.

Design

A crossover design (Hills & Armitage, 1979) was used to compare the learning outcomes for participants in two intervention conditions: infant-directed song and infant-directed speech. The participants were randomly assigned to one of two groups. Group 1 received the sung followed by the spoken condition, and Group 2 received the spoken followed by the sung intervention condition. This design was used to compare the effect of the two intervention conditions, infant-directed song (sung) and infant-directed speech (spoken), on receptive labelling skills in young children with ASD.

Materials

Participants were required to label garden creatures depicted in pictorial form. Two comparable sets of materials were required so that performance could be compared across the two different intervention conditions. Each set of materials consisted of a contextual sentence (sung or spoken depending on the intervention condition), a request to touch a picture, and four creatures (see Table 1). The two sets of garden creatures were counterbalanced across intervention conditions.

Table 1. Material sets

Set	Contextual sentence	Instruction	Creature
1	In my garden come and find the little creature that I say.	Touch the …	Daddy-long-legs, dragonfly, stink bug, slug
2	Little creatures everywhere can you find the one I say?	Touch the …	Praying mantis, cockatoo, gecko, wasp

Note. Reproduced from Simpson et al. (2013).

The lyrics and instructions were delivered via a computer interface, and vocal recordings for each set were made by a professional female musician. They were recorded in the presence of a 4-month-old infant during a play session to provide naturalness in creating the prosodic characteristics facilitated by the infant's presence (Trainor, 1996). The sung lyrics were set to the melody "Twinkle, Twinkle Little Star."

Both sung and spoken recordings displayed moderately soft dynamics that were maintained at an even level and warm timbre. This is consistent with research identifying the emotional expression of warmth observed in infant-directed speech (Trainor, Austin, & Desjardins, 2000). The lyrics were delivered slightly faster during the spoken sound files (7 s) compared to the sung sound files (10 s). There was no sense of metre in the spoken conditions; in contrast, the sung files displayed a simple quadruple metre. Analysis of the sound recordings indicated the pitch range was in accordance with the previous research using speech and play songs directed to 4-month-olds (McRoberts, McDonough, & Lakusta, 2009; Trainor et al., 1997). Frequency recorded during the sung sound bites was 220 - 370 Hz compared to the spoken sound bites' 174.61 - 493.88 Hz, with the majority of the speech occurring between 277 - 311 Hz. The melodic contour in the sung recordings displayed a specific intervallic rise and fall compared to the mellifluous contour evident in the spoken recordings. Thus the sung and spoken recordings differed only in the musical elements of melody and rhythm.

The creatures were presented visually via the computer interface using coloured Picture Communication Symbols (PCS™) selected from Boardmaker® Version 6 (Mayer-Johnson, 2007). Consideration was given to equating the two sets of creature PCS on visual characteristics of colour and type of creature.

Settings

A touch screen monitor was used to reduce the dependency on eye–hand skills required when using a mouse (Huguenin, 2004; Light & Drager, 2007). In addition, the use of a computer interface can provide an engaging learning context for children with autism (Moore & Calvert, 2000). The monitor was a 40 cm touch screen with an adjustable stand that could be situated on a table and adjusted for the child. Creature PCS were presented on the screen using Microsoft Office PowerPoint® 2007. The use of a computer program to deliver the instructions and PCS allowed for a standardised presentation of the intervention in a consistent format, removing the potentially confounding effect of inconsistency of presentation (Goldsmith & LeBlanc, 2004; Panyan, 1984). The intervention sessions were conducted at the participants' schools. They occurred in a separate room to avoid disruptions to the class and data collection. The participant was positioned on a chair, facing the monitor, which was placed at eye level. The researcher was positioned adjacent to the participant. A video recorder was positioned on a tripod to the side of the participant to capture the participant's eye gaze.

Procedure

Teacher and parental/guardian permission was obtained prior to the commencement of the study. Prior to the start of each session, participant consent was obtained. Participants were shown a picture communication symbol of the computer and invited to attend the session.

Expression of consent was assumed if the participant willingly accompanied the researcher, while refusal was viewed as a withdrawal of consent for that session. The duration of each session was 3–6 minutes.

Each intervention condition consisted of baseline, training sessions, teaching sessions and generalisation training and assessment. This was conducted within a 5-week period. Maintenance was measured 3–5 weeks after the last teaching session. The second intervention condition commenced after maintenance data had been collected. Participants in Group 1 received the sung intervention condition first, followed by the spoken intervention condition; Group 2 received the intervention conditions in the reverse order. All other procedures were identical for both intervention conditions.

Baseline. During baseline, horizontal arrays of three creatures were displayed on the monitor. Data were collected on the number of times a participant touched the correct creature following the instruction to do so. Touching any creature activated the presentation of the next slide. Each creature was requested three times during the baseline session. The order of presentation of the creatures, the position of the creature PCS in the array, and the selection of the distracter PCS was randomly generated using the computer program described below (see Training and Teaching Sessions). No error correction or reinforcement procedures were used. Correct and incorrect responses were recorded and sessions were video-recorded.

Training sessions. Each participant received four training sessions. At this stage, only the named creature PCS was presented to the participant without any distracter symbols. In every session, each creature PCS was presented three times. A correct response involved touching the PCS on the screen within 10-s. If the participant did not make a correct response within a 10-s period, the researcher prompted a correct response following a least-to-most prompt hierarchy (gesture, gesture and physical, full physical; Duker, Didden, & Sigafoos, 2004). A correct response (unprompted or prompted) triggered the next slide, which displayed the correct creature PCS using the "spin" animation, and simultaneously the prerecorded name of the creature was presented in a way consistent with the intervention condition (sung or spoken). The order of presentation of creatures and the position of the PCS on the screen (left, centre, right) were randomised across the training sessions.

Teaching sessions. A block of 15 teaching sessions was conducted using a conditional discrimination task procedure (Green, 2001; Grow, Carr, Kodak, Jostad, & Kisamore, 2011). Creature PCS were presented in arrays of three: one target and two distracter PCS from the set. During each session a creature was presented three times. The time delay and prompt hierarchy were identical to the training session. The order of presentation of creatures, the position of the correct PCS in the array, and the selection of the two distracter PCS was generated by means of a program developed using Microsoft (.net 2008) C# program language. The random selection of arrays reduced the possibility of rote learning positional responses and avoided positional bias selection (Duker et al., 2004). A correct response was recorded by the researcher on unprompted correct responses.

Maintenance. Maintenance data were collected 3–5 weeks following the cessation of the last teaching session in each intervention condition. The maintenance session procedures replicated the intervention sessions. Data were collected on the participants' (correct/incorrect) responses.

Generalisation and assessment. Generalisation training occurred during the intervention time period and was conducted in the classroom. Black and white PCS of each creature were presented in a laminated 10 x 10 cm booklet. There was one PCS per page. The booklets were shared with the participants once a day on the days they attended school. The adult (teacher, teacher aide, volunteer, researcher) labelled the PCS using a regular speaking voice. Frequency of generalisation sessions did not vary for individual participants across the two conditions, but did vary between participants depending on their school attendance.

Generalisation assessment was conducted following the last teaching session for that intervention condition. Participants were presented with a laminated A4 booklet containing 12 pages, positioned on a vertical slope board. Each page displayed a horizontal array of three coloured creature symbols consistent with the teaching condition. The researcher, using her regular speaking voice, provided the prompt “touch the [creature name].” Each of the creatures was presented three times. No error correction or reinforcers were employed. The assessment was conducted in the same room as the intervention. Data were collected on the participants’ (correct/incorrect) responses and the session was videotaped to enable interrater reliability.

Data analysis

Data analysis was designed to compare the effect of the two intervention conditions, infant-directed song (sung) and infant-directed speech (spoken), across time, baseline (preintervention), Session 15 (postintervention), and maintenance and generalisation on receptive labelling skills in young children with ASD in Group 1 (sung, spoken) and Group 2 (spoken, sung). The analyses, using IBM SPSS Version 19, were as follows:

1. A 2 x 2 x 3 (Group [1, 2] x Condition [sung, spoken] x Time [baseline, Session 15, maintenance]) ANOVA was undertaken to determine whether either of the groups’ correct response scores were significantly different for the two intervention conditions and at each of the time periods, and to determine if there was an interaction between order of presentation and condition across time.
2. A two-way repeated measures ANOVA was undertaken on correct response scores in both Group 1 and Group 2 to determine if there was a significant

difference for the two intervention conditions, between Session 15 and generalisation. These analyses were conducted to determine whether learning demonstrated as a result of the intervention was generalised for either of the groups.

Interobserver reliability

The use of a computer interface for the intervention task allowed for infinite presentations without degradation of fidelity (Mineo, Ziegler, Gill, & Salkin, 2009). This also provided an inbuilt fidelity for recording correct responses for the intervention phase. Activation of the slide was dependent upon the participant's selection of the correct picture. Reliability was assessed on sessions where there was no inbuilt fidelity (baseline and generalisation). A test of reliability was conducted on a random sampling of 30% of the video recordings by a second observer trained in the data recording, but blind to the purpose of the study. Sampling was representative of participants in both baseline and generalisation phases. The kappa measurement of agreement value on the baseline data was .93 (occurrence = 87.5%, nonoccurrence = 100.0%); and on the generalisation data was .93 (occurrence = 94.1%, nonoccurrence = 98.7% incorrect).

Results

The mean correct response scores for Group 1 and Group 2 in both intervention conditions (sung and spoken) for baseline, Session 15 and maintenance, and generalisation are presented in Table 2.

Table 2. Group 1 and 2 correct response scores

	Group 1 (n = 11)		Group 2 (n = 11)	
	Sung *M* (*SD*)	Spoken *M* (*SD*)	Spoken *M* (*SD*)	Sung *M* (*SD*)
Baseline	1.91 (0.83)	1.55 (0.82)	2.55 (1.04)	1.91 (1.30)
Session 15	3.36 (3.50)	2.55 (2.77)	6.45 (4.03)	5.91 (5.28)
Generalisation	3.36 (3.50)	1.91 (2.34)	5.00 (4.47)	6.09 (5.45)
Maintenance	3.55 (3.21)	2.45 (2.58)	6.45 (3.83)	6.55 (5.13)

Comparison of group, condition, and time

A 2 x 2 x 3 mixed between-within subject ANOVA was applied to the results in Table 2 to compare the groups' correct response scores in the sung and spoken intervention conditions across three time periods. A significant main effect for group was obtained, $F(1, 20) = 5.96$, $p = .02$, $\eta^2 = .23$, identifying 23% of the variance in scores was due to the group. No significant main effect for condition was found, $F(1, 20) = 0.14$, $p = .71$, $\eta^2 = .01$, indicating a small effect (using Cohen, 1988, guidelines). Given the size of the effect, it is unlikely that increasing the sample size would have increased the probability of a significant effect. A significant main effect for time was obtained; however, as sphericity was violated for time, adjustment was made using the Greenhouse–Geisser *F* resulting in (1.39, 20) = 12.94, $p < .001$, $\eta^2 = .39$. There was no significant interaction for group and condition, group and time, condition and time, or for group, condition, and time. The effect sizes were small, < .05 (using Cohen, 1988, guidelines), with the exception of a large effect size for time and group, $\eta^2 = .15$, consistent with the variance between groups. In summary, this indicates there were no differences as a result of any unique interactions between group, condition, and time. Furthermore, the results suggest there was no effect for condition (sung or spoken); however, there were differences due to the group regardless of condition and there were differences due to time regardless of condition. Further analysis was conducted to determine where the difference lay across the three time periods. Post-hoc tests using the Bonferonni adjustment revealed a significant increase from baseline (M = 1.93, SD = 1.05) to Session 15 (M = 4.57, SD = 4.21), p = .005, and large effect, d = 0.86, and baseline (M = 1.93, SD = 1.05) and maintenance (M = 4.75, SD = 4.09), p = .002, and large effect, d = 0.94. There was no significant difference between maintenance and Session 15, p > .05. These results indicate that the level of correct responses significantly increased with intervention and this level was maintained at follow-up.

Generalisation

A two-way repeated measure ANOVA was applied to the results for Group 1 and Group 2 (Table 2) to compare correct response scores for the condition (sung/spoken) in Session 15 and the generalisation session. There was no statistically significant main effect for Group 1 on condition, $F(1, 10) = 0.95$, $p = .35$, $\eta^2 = .09$, a moderate effect size or session, $F(1, 10) = 0.10$, $p = .75$, $\eta^2 = .01$, small effect size. There was no statistically significant main effect for Group 2 on condition, $F(1, 10) = 0.86$, $p = .77$, $\eta^2 = .01$, small effect size or session, $F(1, 10) = 0.54$, $p = .48$, $\eta^2 = .05$, a small effect size. There was no significant interaction between condition and session for Group 1 or Group 2. These results indicate participants' learning was generalised and generalisation was not better following one intervention condition over the other.

Discussion

This study found that there was no significant difference in the use of a sung over a spoken presentation to facilitate the learning of picture labelling for children with ASD. Participants maintained and generalised the receptive labelling responses gained during intervention following both conditions. There was, however, a significant difference between the level of correct responses between Group 1 and Group 2. These findings are discussed in terms of potential differences between the groups, similarities between the conditions with reference to implications for practice, future research, and study limitations.

Group comparison

Participants were randomly assigned to the groups to reduce the effect of any participant variability. In addition, the two groups demonstrated no differences on the screening measures in terms of chronological age, receptive and expressive language, and adaptive functioning and behaviour. Despite this, there was a significant difference in the levels of correct responses between groups in both conditions, suggesting that there may have been between-group differences that were not detected. This may have been due to a lack of sensitivity in the screening measures used or differences among participants in a characteristic not assessed in this study.

Participants' expressive and receptive language skills were assessed used the EVT-2 and PPVT-4 for descriptive purposes. These norm-referenced tests commence at the developmental age of 2 years 6 months. There were only two participants who scored above the 0.1 percentile on the EVT-2. These participants were both in Group 2, which may have resulted in noted differences between these groups. Although there was no significant difference between groups on this measure, learning outcomes may have been particularly sensitive to a very small difference in the level of preintervention expressive language skills. These two participants scored below 0.1 percentile on the PPVT-4, indicating that their receptive language ability may be more impaired than expressive language ability, consistent with the findings of Hudry et al. (2010) and highlighting the importance of addressing receptive language learning. In any case, expanding the range of measures used to identify and describe participants and employing more sensitive measures, particularly of expressive language skills, would be important for any future studies in this area.

Sung and spoken conditions

Both the sung and spoken interventions increased the participants' receptive labelling responses, consistent with previous findings that identified language production increased in young children with ASD following both music and speech interventions (Lim, 2010). Unlike earlier research (e.g., Buday, 1995), song did not lead to superior learning when compared to speech. The infant-directed song used in this study involved a solo voice without instrumental accompaniment, as is typically observed in caregiver singing to infants. Song can also include additional musical elements not included in this study, such as harmony and texture. The potential for these elements to facilitate learning in children with ASD has yet to be investigated. It is possible that the learning in this study was influenced by commonalities between the two intervention conditions, identified as similarities in the auditory stimuli, teaching strategies, and the use of a computer interface.

Common to both the sung and spoken conditions was elevated pitch. That pitch may be a salient feature for language learning is supported by previous research that has found individuals with autism have preserved or enhanced processing ability of linguistic pitch (Heaton, Hudry, Ludlow, & Hill, 2008; Järvinen-Pasley, Pasley, & Heaton, 2008) and musical pitch (Heaton, 2005; Mottron, Peretz, & Ménard, 2000). In addition, although the sung and spoken stimuli differed in terms of rhythm, a rhythmic pattern was apparent in both the sung and spoken conditions. In the sung condition, the use of metre was evident by the temporal regularity of the music, whereas the spoken stimuli displayed a sense of rhythmic quality not tied to a metre but produced by regularly stressed syllables. The rhythmic pattern provided a structure for the lyrics that could have potentially aided information encoding and word detection. This is consistent with research that has shown that recognition of word order by young infants is aided by the organisational framework of prosodic sentences compared to the use of fragmented or part sentences (Mandel, Kemler Nelson, & Jusczyk, 1996). The commonality of the auditory input between the two conditions may have contributed to the maintenance and generalisation of the learning. Previous research has found song aids long-term recall (Rainey & Larsen, 2002); this study showed learning was maintained in the sung condition. Learning was also maintained following the spoken condition. Further research is required to gain a more sophisticated understanding of the impact of the musical elements of pitch and rhythm on the language processing of children with ASD.

The intervention procedures in both conditions used strategies based on applied behaviour analysis that have a strong evidence base, including time delay and prompting. A number of these strategies have also been identified as features of typical language learning, such as the use of a familiar frame to cue the target word (Fernald & Mazzie, 1991; Fernald, McRoberts, & Swingley, 2001), reinforcing the target word in a singular format (Lew-Williams, Pelucchi, & Saffran, 2011), and the repetitive nature of the utterances (McRoberts et al., 2009). Furthermore, this intervention was presented using a computer interface. Given the responsiveness to computer instruction by children with autism (Moore & Calvert, 2000) the use of the computer interface may have contributed to the intervention outcomes in this study. The impact of using effective practices in conjunction with the infant-directed stimuli on language learning in young children with ASD requires further consideration.

Limitations

There were a number of limitations to the study that should be considered when interpreting these results. The crossover design and sample size helped to control for variability in performance across participants. However, no one intervention has been found to be effective for all children with ASD (Magiati, Tay, & Howlin, 2012), and caution is warranted in generalising any intervention results to all children with ASD.

A further limitation of this study was that the number of intervention sessions was dictated by the school calendar. Intervention studies frequently teach to criteria, even if that requires a large number of intervention sessions (see Grow et al., 2011). In this study, participants did not have the opportunity to reach predetermined learning criteria. As research has shown that music stimuli are engaging for children with autism, it would be interesting to determine if music could act as a mediating or moderating variable on performance in relation to intervention duration (Blackstock, 1978; Buday, 1995; Carnahan et al., 2009; Simpson et al., 2013). Furthermore, the school calendar limited the opportunity for extended follow-up, and further research to determine if a particular intervention condition showed a differential long-term effect would be valuable.

This study compared sung and spoken interventions to teach receptive labelling skills to young children with ASD. As such, it provides an initial investigation into the use of musical elements to facilitate receptive labelling skills in children with ASD. The results did not support the use of sung over spoken conditions to facilitate language learning in this context. Learning occurred in both conditions and natural forms of vocal stimuli used in the intervention and that occur during the prelinguistic stage of typical development may contribute to language learning in children with ASD.

Author note

This review is part of the first authors' PhD study, which was supported by a PhD scholarship - The Australian Catholic University Postgraduate Award. There is no conflict of interest for any of the authors.

Acknowledgements

We would like to acknowledge the support of the principals, teachers, parents, and students from the participating State Special Schools. In addition, we thank Jacqui Cuny for the contribution of her musical skills in the development of the auditory material.

References

Bergeson, T. R., & Trehub, S. E. (1999). Mothers' singing to infants and preschool children. *Infant Behavior & Development*, *22*, 51–64. doi:10.1016/S0163-6383(99)80005-8

Blackstock, E. G. (1978). Cerebral asymmetry and the development of early infantile autism. *Journal of Autism and Childhood Schizophrenia*, *8*, 339–353. doi:10.1007/BF01539636

Bruininks, R., Woodcock, R., Weatherman, R., & Hill, B. (1996). *Scales of Independent Behavior – Revised: Comprehensive manual*. Rolling Meadows, IL: Riverside.

Buday, E. M. (1995). The effects of signed and spoken words taught with music on sign and speech imitation by children with autism. *Journal of Music Therapy*, *32*, 189–202. doi:10.1093/jmt/32.3.189

Carnahan, C., Basham, J., & Musti-Rao, S. (2009). A low-technology strategy for increasing engagement of students with autism and significant learning needs. *Exceptionality*, *17*, 76–87. doi:10.1080/09362830902805798

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, *30*, 213–236. doi:10.1017/S0305000902005482

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Duker, P., Didden, R., & Sigafoos, J. (2004). *One-to-one training: Instructional procedures for learners with developmental disabilities*. Austin, TX: Pro-Ed.

Dunn, L. M., & Dunn, D. M. (2007). *The Peabody picture vocabulary test manual* (4th ed.). Minneapolis: NCS Pearson.

Eaves, L. C., Wingert, H. D., Ho, H. H., & Mickelson, E. C. R. (2006). Screening for autism spectrum disorders with the social communication questionnaire. *Journal of Developmental & Behavioral Pediatrics*, *27*, S95–S103. doi:10.1097/00004703-200604002-00007

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, *8*, 181–195. doi:10.1016/S0163-6383(85)80005-9

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*, 209–221. doi:10.1037/0012-1649.27.2.209

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in recognizing and understanding words in fluent speech. In J. Weissenborn & B. Höhle (Eds.), *Approaches to bootstrapping: Volume I: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (pp. 97–123). Amsterdam, the Netherlands: Benjamins.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*, 104–113. doi:10.1037/0012-1649.20.1.104

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*, 477–501. doi:10.1017/S0305000900010679

Goldsmith, T. R., & LeBlanc, L. A. (2004). Use of technology in interventions for children with autism. *Journal of Early and Intensive Behavior Intervention*, *1*, 166–178.

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, *16*, 72–85. doi:10.1177/108835760101600203

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *44*, 475–498. doi:10.1901/jaba.2011.44-475

Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research*, *44*, 1189–1200. doi:10.1044/1092-4388(2001/092)

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders, 35,* 787–793. doi:10.1007/s10803-005-0024-7

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, *25*, 771–782. doi:10.1080/02643290802336277

Hills, M., & Armitage, P. (1979). The two-period cross-over clinical trial. *British Journal of Clinical Pharmacology*, *8,* 7–20. doi:10.1111/j.1365-2125.1979.tb05903.x

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., … PACT Consortium. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders*, *45,* 681–690. doi:10.3109/13682820903461493

Huguenin, N. H. (2004). Assessing visual attention in young children and adolescents with severe mental retardation utilizing conditional-discrimination tasks and multiple testing procedures. *Research in Developmental Disabilities*, *25*, 155–181. doi:10.1016/j.ridd.2004.01.001

Järvinen-Pasley, A., Pasley, J., & Heaton, P. (2008). Is the linguistic content of speech less salient than its perceptual features in autism? *Journal of Autism and Developmental Disorders, 38,* 239–248. doi:10.1007/s10803-007-0386-0

Koegel, L. K., Camarata, S. M., Valdez-Menchaca, M., & Koegel, R. L. (1998). Setting generalization of question-asking by children with autism. *American Journal on Mental Retardation*, *102*, 346–357.

Kolko, D. J., Anderson, L., & Campbell, M. (1980). Sensory preference and overselective responding in autistic children. *Journal of Autism and Developmental Disorders*, *10,* 259–271. doi:10.1007/BF02408285

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, *8*, F1–F12. doi:10.1111/j.1467-7687.2004.00384.x

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period”. *Language Learning and Development*, *1*, 237–264. doi:10.1080/15475441.2005.9671948

Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: Infants’ perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior & Development*, *33*, 419–430. doi:10.1016/j.infbeh.2010.04.006

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, *14*, 1323–1329. doi:10.1111/j.1467-7687.2011.01079.x

Light, J., & Drager, K. (2007). AAC technologies for young children with complex communication needs: State of the science and future research directions. *Augmentative and Alternative Communication*, *23*, 204–216. doi:10.1080/07434610701553635

Lim, H. A. (2010). Effect of “developmental speech and language training through music” on speech production in children with autism spectrum disorders. *Journal of Music Therapy*, *47*, 2–26. doi:10.1093/jmt/47.1.2

Luyster, R., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language*, *34*, 623–654. doi:10.1017/S0305000907008094

Magiati, I., Tay, X. W., & Howlin, P. (2012). Early comprehensive behaviorally based interventions for children with autism spectrum disorders: A summary of findings from recent reviews and meta-analyses. *Neuropsychiatry*, *2*, 543–570. doi:10.2217/NPY.12.59

Maljaars, J., Noens, I., Scholte, E., & van Berckelaer-Onnes, I. (2012). Language in low-functioning children with autistic disorder: Differences between receptive and expressive skills and concurrent predictors of language. *Journal of Autism and Developmental Disorders*, *42*, 2181–2191. doi:10.1007/s10803-012-1476-1

Mandel, D. R., Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development*, *11*, 181–196. doi:10.1016/S0885-2014(96)90002-7

Mayer-Johnson. (2007). Boardmaker (Version 6) [Computer software]. Solana Beach, CA: Author LLC.

McRoberts, G. W., McDonough, C., & Lakusta, L. (2009). The role of verbal repetition in the development of infant speech preferences from 4 to 14 months of age. *Infancy*, *14*, 162–194. doi:10.1080/15250000802707062

Mineo, B. A., Ziegler, W., Gill, S., & Salkin, D. (2009). Engagement with electronic screen media among students with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 172–187. doi:10.1007/s10803-008-0616-0

Moore, M., & Calvert, S. (2000). Brief report: Vocabulary acquisition for children with autism: Teacher or computer instruction. *Journal of Autism and Developmental Disorders, 30,* 359–362. doi:10.1023/A:1005535602064

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry, 41*, 1057–1065. doi:10.1111/1469-7610.00693

Panyan, M. V. (1984). Computer technology for autistic students. *Journal of Autism and Developmental Disorders*, *14*, 375–382. doi:10.1007/BF02409828

Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior & Development*, *14*, 415–440. doi:10.1016/0163-6383(91)90031-M

Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F. (2007). "Listen my children and you shall hear": Auditory preferences in toddlers with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 50,* 1350–1364. doi:10.1044/1092-4388(2007/094)

Peretz, I., Radeau, M., & Aguin, M. (2004). Two-way interactions between music and language: Evidence from priming recognition of tune and lyrics in familiar songs. *Memory & Cognition*, *32*, 142–152. doi:10.3758/BF03195827

Rainey, D. W., & Larsen, J. D. (2002). The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception, 20,* 173–186. doi:10.1525/mp.2002.20.2.173

Rutter, M., Bailey, A., Berument, S., Lord, C., & Pickles, A. (2003). *Social communication questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Santarcangelo, S., & Dyer, K. (1988). Prosodic aspects of motherese: Effects on gaze and responsiveness in developmentally disabled children. *Journal of Experimental Child Psychology*, *46*, 406–418. doi:10.1016/0022-0965(88)90069-0

Simpson, K., & Keen, D. (2010). Teaching young children with autism graphic symbols embedded within an interactive song. *Journal of Developmental and Physical Disabilities*, *22*, 165–177. doi:10.1007/s10882-009-9173-5

Simpson, K., & Keen, D. (2011). Music interventions for children with autism: Narrative review of the literature. *Journal of Autism and Developmental Disorders*, *41*, 1507–1514. doi:10.1007/s10803-010-1172-y

Simpson, K., Keen, D., & Lamb, J. (2013). The use of music to engage children with autism in a receptive labelling task. *Research in Autism Spectrum Disorders*, *7*, 1489–1496. doi:10.1016/j.rasd.2013.08.013

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, *14*, 654–666. doi:10.1080/15250000903263973

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*, 389–400. doi:10.1121/1.3419786

Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, *18*, 727–735. doi:10.1037/0012-1649.18.5.727

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*, 53–71. doi:10.1207/s15327078in0701_5

Thiessen, E. D., & Saffran, J. R. (2009). How the melody facilitates the message and vice versa in infant learning and memory. *Annals of the New York Academy of Sciences, 1169*, 225–233. doi:10.1111/j.1749-6632.2009.04547.x

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior & Development*, *19,* 83–92. doi:10.1016/S0163-6383(96)90046-6

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of vocal expression of emotion? *Psychological Science*, *11*, 188–195. doi:10.1111/1467-9280.00240

Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior & Development*, *20*, 383–396. doi:10.1016/S0163-6383(97)90009-6

Trehub, S. E., & Schellenberg, E. G. (1995). Music: Its relevance to infants. In R. Vasta (Ed.), *Annals of child development* (Vol. 11, pp. 1–24). New York, NY: Kingsley.

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1471–1485. doi:10.1037/0278-7393.20.6.1471

Watson, L. R., Baranek, G. T., Roberts, J. E., David, F. J., & Perryman, T. Y. (2010). Behavioral and physiological responses to child-directed speech as predictors of communication outcomes in children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *53*, 1052–1064. doi:10.1044/1092-4388(2009/09-0096)

Watson, L. R., Roberts, J. E,, Baranek, G. T., Mandulak, K. C., & Dalton, J. C. (2012). Behavioral and physiological responses to child-directed speech of children with autism spectrum disorders or typical development. *Journal of Autism and Developmental Disorders*, *42*, 1616–1629. doi:10.1007/s10803-011-1401-z

Williams, K. T. (2007). *Expressive vocabulary test*, (2nd ed.). Minneapolis, MN: NCS Pearson.