2007-12-13

The Effect of "Developmental Speech-Language Training through Music" on Speech Production in Children with Autism Spectrum Disorders

Hayoung Audrey Lim
University of Miami, hayoung_l@hotmail.com

Follow this and additional works at: http://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

THE EFFECT OF “DEVELOPMENTAL SPEECH-LANGUAGE TRAINING THROUGH MUSIC” ON SPEECH PRODUCTION IN CHILDREN WITH AUTISM SPECTRUM DISORDERS

By

Hayoung Audrey Lim

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2007
THE EFFECT OF “DEVELOPMENTAL SPEECH-LANGUAGE TRAINING THROUGH MUSIC” ON SPEECH PRODUCTION IN CHILDREN WITH AUTISM SPECTRUM DISORDERS

Hayoung A. Lim

Approved:

___________________                              _____________________
Dr. Shannon K. de l’Etoile                            Dr. Terri A. Scandura
Associate Professor of Music Therapy                   Dean of the Graduate School

____________________                             ______________________
Dr. Edward P. Asmus                                 Dr. Teresa L. Lesiuk
Associate Dean of Graduate Studies                   Assistant Professor of Music Therapy
Professor of Music Education

_____________________                           _______________________
Mr. Ross T. Harbaugh                               Dr. Peter C. Mundy
Professor of Violoncello                            Professor of Psychology, Child Division
                                                   Director of Center for Autism &
                                                   Related Disabilities (CARD)
Children with Autism Spectrum Disorders demonstrate deficits in speech and language, with the most outstanding speech impairments being in comprehension, semantics, prosody, and pragmatics. Perception and production of music and speech in children with ASD appear to follow the same principles of Gestalt pattern perceptual organization. In addition, common neuroanatomical structures and similar patterns of cortical activation mediate the perception and production of speech and music. Therefore, the purpose of this study was to explore how the perception of musical stimuli would impact the perception and production of speech and language in children with ASD. The study examined the effect of developmental speech-language training through music on the speech production of children with ASD. The participants were 50 children with ASD, age range 3 to 5 years, who had previously been evaluated on standard tests of language and level of functioning. The children completed the pre-test, six sessions of training, and the post-test. The pre-and post-tests consisted of the Verbal Production Evaluation Scale (VPES) and measured each participant’s verbal production including semantics, phonology, pragmatics, and prosody, of 36 target words. Eighteen participants completed music training, in which they watched a music video containing six songs and pictures of the 36 target words. Another group of eighteen participants completed speech
training, in which they watched a speech video containing six stories and pictures of target words. Fourteen participants were randomly assigned to a no-training condition. Results of the study showed that participants in both music and speech training significantly increased their scores on the VPES from the pre-test to the post-test. Both music and speech training were effective for enhancing participants’ speech production including semantics, phonology, pragmatics, and prosody. Participants who received music training made greater progress on speech production than participants who received the speech training; however, the difference was not statistically significant. Results of the study also indicated that the level of speech production was influenced by the level of functioning in children with ASD. An interaction between level of functioning and training conditions on speech production approached significance. The results indicate that both high and low functioning participants improved their speech production after receiving either music or speech training; however, low functioning participants showed a greater improvement in speech production after the music training than after the speech training. Collectively, music training was more effective for speech production in low functioning children with ASD than was speech training. The study suggests that the superior performance in speech production in children with ASD who received music training might be generated from music stimuli which were organized by the Gestalt laws of pattern perception. In conclusion, children with ASD appear to perceive important linguistic information (i.e., target words) embedded in music stimuli, and can verbally produce the words as functional speech. These results provide evidence for the use of music as an effective way to enhance speech production in children with ASD.
ACKNOWLEDGEMENTS

I wish to thank my committee members for their invaluable support for this dissertation, especially the chairman of the committee, Dr. Shannon K. de l’Etoile, for her zealous dedication to the completion of this study. I would also like to express my gratitude to Dr. Edward Asmus for his thoughtful consultation, and Dr. Teresa Lesiuk for her meaningful feedback. I would like to thank Professor Ross Harbaugh for his sincere support and care during the dissertation process, as well as for the direction for my cello performance. Many thanks also to Dr. Peter Mundy for his valuable contributions and knowledge regarding children with Autism Spectrum Disorders and the methodology of the present study.

I owe many thanks to parents and children who agreed to participate in my study and showed wonderful support throughout the research procedures. Special thanks to the directors, staff, and teachers of The Victory School for Children with Autism; The Kendall Speech and Language Center; The Children’s Resources Preschool; The United Cerebral Palsy Association of Miami; The Dave and Mary Alper Jewish Community Center; The Creative Children Therapy; and The University of Miami Center for Autism and Related Disorders; for their countless help and assistance. My sincere appreciation extends to Barbara Soliday for her great editorial work for this dissertation. In addition, my special expression of gratitude goes to Jodi Silver who presented songs and stories for the present study.

This dissertation would not have been possible without the dedication and commitment of my parents. Without them, the education I have received in the U.S.A., and the opportunities I have now would have never been possible; so I owe a lifetime of
gratitude to my parents. I would also like to thank to my parents-in-law for their continuous support and prayers. They have encouraged me consistently and showed great love throughout my time of the study. In addition, I would like to thank my sisters and brothers in Miami Onnuri Church for their earnest prayers and supports.

Finally, I wish to express my utmost gratitude and love to my husband, Timothy, for his dedication and unlimited encouragement, trust, and prayer. This dissertation and my career as music therapist would never been possible without his love and commitment.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
</tbody>
</table>

## Chapter

### I INTRODUCTION       1

- Statement of the Problem       1
- Need for the Study            10
- Purpose of the Study          14

### II REVIEW OF RELATED LITERATURE  16

- Perception and Production of Speech in Children with ASD       16
  - Speech and Language Impairments       16
  - Abnormal Auditory Cortical Processing in Children with ASD   20
  - Language Acquisition in Children with ASD       23
  - Gestalt Styles of Language Acquisition in Children with ASD 29
- Perception and Production of Music in Children with ASD       38
  - Music Perception       38
  - Similarities between Music and Speech/Language Perception  40
  - Neuro-anatomical Commonalities between Music and Speech      42
  - Perception of Music in Children with ASD       47
  - Pattern Perception of Music and Speech in Children with ASD 49
  - Musical Behaviors in Children with ASD       54
The Effect of Music on Speech and Language in Children with ASD...58

Summary of the Literature Review ........................................64

III METHOD ...........................................................................67

Participants .................................................................67

Materials ........................................................................74

Procedure .......................................................................81

IV RESULTS .........................................................................87

Research Question # 1 .....................................................87

Research Question # 2 .....................................................91

Research Question # 3 .....................................................96

Analysis of Training Condition and Level of Functioning ........101

Research Question # 4 .....................................................105

Target Word Analysis .....................................................107

V DISCUSSION .................................................................113

Discussion of the Research Questions .............................113

Discussion of Target Words ............................................129

Discussion of Music Stimuli and Speech Stimuli ...............133

Limitations of the Present Study ......................................137

Theoretical Implications ................................................139

Clinical Implications .....................................................141

Recommendations for Future Research .........................145

Conclusions ..............................................................148

References ......................................................................150

vi
Appendices

Appendix A: Texts and Notation for Speech and Music Stimuli………………159
Appendix B: Verbal Production Evaluation Scale …......................................167
Appendix C: Letter to Parents …...............................................................168
Appendix D: Informed Parental Consent Form …..........................................169
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mean Change Scores by Training Condition</td>
<td>88</td>
</tr>
<tr>
<td>2.</td>
<td>Mean Change Scores by Level of Functioning</td>
<td>94</td>
</tr>
<tr>
<td>3.</td>
<td>The Effects of Training Condition and Level of Functioning on Mean Changes in Verbal Production</td>
<td>104</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participants’ Age, Level of Functioning, Language Age, Echolalia and Gender</td>
<td>75</td>
</tr>
<tr>
<td>2. Criteria for Vocabulary Word Selection and Target Words</td>
<td>78</td>
</tr>
<tr>
<td>3. Means and Standard Deviations for Pre-test, Post-test, and Change Scores in Verbal Production Evaluation Scale by Training Condition</td>
<td>87</td>
</tr>
<tr>
<td>4. Analysis of Training Conditions on the Change Scores</td>
<td>90</td>
</tr>
<tr>
<td>5. Cohen’s d Effect Size between Training Conditions</td>
<td>91</td>
</tr>
<tr>
<td>6. Means and Standard Deviations for Post-test Score and Change Scores by Level of Functioning and Echolalia</td>
<td>92</td>
</tr>
<tr>
<td>7. Pearson Product-Moment Correlation Coefficient for Type of Training, Level of Functioning, Echolalia, Language Age, and Pre-test Score</td>
<td>95</td>
</tr>
<tr>
<td>8. Descriptive Results for Post-test by Music Training Condition, Echolalia, and Level of Functioning</td>
<td>97</td>
</tr>
<tr>
<td>9. Descriptive Results for Post-test by Speech Training Condition, Echolalia, and Level of Functioning</td>
<td>98</td>
</tr>
<tr>
<td>10. Descriptive Results for Post-test by No-Training Condition, Echolalia, and Level of Functioning</td>
<td>99</td>
</tr>
<tr>
<td>11. Mean Change scores and Standard Error by Training Condition and Level of Functioning</td>
<td>102</td>
</tr>
<tr>
<td>12. ANOVA for Between Subjects Effect of Training Condition and Level of Functioning on Pre to Post-Test Change</td>
<td>105</td>
</tr>
<tr>
<td>14. Target Words for Song-Story # 1 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test</td>
<td>108</td>
</tr>
<tr>
<td>15. Target Words for Song-Story # 2 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test</td>
<td>108</td>
</tr>
</tbody>
</table>
16. Target Words for Song-Story # 3 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test .................................109

17. Target Words for Song-Story # 4 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test .................................109

18. Target Words for Song-Story # 5 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test .................................110

19. Target Words for Song-Story # 6 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test .................................110
CHAPTER ONE

Introduction

Statement of the Problem

Autism Spectrum Disorder (ASD) is identified as a behaviorally defined syndrome with a broad range of severity, resulting from brain dysfunction (Rapin & Dunn, 2002). The diagnostic schemes and description of the underlying deficits in ASD have evolved and changed since Leo Kanner’s first description of infantile autism in 1943 (Prizant & Wetherby, 2005). One criterion for diagnosis has remained constant however, which is difficulty in the development of social communication (American Psychiatric Association [APA], 1994). In particular, speech and language impairment has been regarded as one of the most significant deficits in ASD (Lord & Paul, 1997; Rapin & Dunn, 2002). Effective intervention therefore is needed in the development of social communication including speech and language.

Speech/language impairment in ASD. Social communication requires the acquisition and use of conventional and socially appropriate means to communicate for a variety of purposes across social contexts (Prizant & Wetherby, 2005). Communicative means may include verbal and nonverbal behaviors, such as language expressed through speech, identifying pictures, gestures, or manual signs. In particular, communicative language competence may primarily determine the extent to which individuals with ASD can develop relationships with others and participate in daily activities at school, at home, and in the community. Furthermore, social communication ability, particularly the use of language, is critical for individuals with ASD in order to lead independent and productive lives.
The level of communicative language competence achieved by children with ASD is closely related to their development of cognitive, emotional and social behaviors (Lord & Paul, 1997; Prizant & Wetherby, 2005). In addition, gains in communication skills are directly related to the prevention and reduction of problem behaviors in children with ASD (Reichle & Wacker, 1993). Furthermore, a delay in the development of speech and language is one of the most distinguishable deficits in children with ASD, compared to typically developing children. Parents of children with autism have identified speech/language impairment as one of the most significant stresses they experience with their children in the preschool and school years (Bristol & Schopler, 1984). In summary, a deficit in the communicative use of language is a defining feature of ASD, and it directly reflects the core developmental difficulties of children with ASD.

Social communication ability is the key for success in everyday activities for children with ASD. Therefore, providing effective intervention to improve communication skills and language is a high priority in efforts to treat children with ASD (Prizant & Wetherby, 2005). Researchers suggest that music could be an effective part of intervention for children with ASD for these reasons: (1) some children with ASD have musical sensitivity, (2) some children with ASD may have a perceptual preference for music, and (3) some children with ASD are able to produce musical patterns. In fact, because of autistic children’s positive musical responses, some clinicians have utilized music for development in various areas such as: fine and gross motor coordination, attention span, social skills, concept of self, and verbal and non-verbal communication (Thaut, 1999).
Musical sensitivity in children with ASD. Musical sensitivity refers to a particular capacity of responsiveness to musical stimuli (Radocy & Boyle, 2003). A distinctive sensitivity and attention to music has been frequently mentioned in the literature on children with ASD (Thaut, 1999). Musical sensitivity in children with ASD includes auditory acuity such as superior auditory detection or discrimination abilities and strong memory for melodies. Early studies on music and autism focused on the unique ways that children with ASD respond to music. In 1953, Sherwin reported three case studies with autistic boys who showed unusual auditory sensitivity, perfect pitch, strong melodic memory, and great interest in playing musical instruments (Sherwin, 1953). Other autism researchers have reported exceptional ability in piano performance of children with autism (Applebaum, Egel, Koegel, & Imhoff, 1979; O’Connell, 1974). These reports indicate that some children with autism have an unusual and superior musical sensitivity.

Not all children with autism, however, possess such musical sensitivity or a particular musical ability. Furthermore, it is not known that musical ability transfers to functional behaviors for children with autism. In fact, cognitive or functional ability does not relate to their superior auditory acuity and musical sensitivity (Berger, 2002; O’Connell, 1974; Radocy & Boyle, 2003). Some children with ASD who have perfect pitch showed limited expressive language (Berger, 2002; O’Connell, 1974). In addition, they showed surprising and inappropriate behaviors, such as screaming when a piece of music happened to be played in a different key, not the original key which was first presented (Berger, 2002).
Perceptual preference of music in children with ASD. Another reason why music may play an effective role in speech and language development of children with ASD is their perceptual preference for music. Researchers have suggested that autistic children are more likely to attend to an auditory than to a visual stimulus, especially when the auditory stimulus is musical (Kolko, Anderson & Campbell, 1980; Thaut, 1987). Thaut (1987) found that when given a choice, ASD children showed a slight preference for a musical stimulus over a visual stimulus. Children with autism spent a significantly longer amount of time listening to the music than watching the visual stimulus. A group of non-autistic children matched by approximate chronological age and developmental age, however preferred the visual stimulus over the music.

Another earlier study on music and autism compared the auditory preferences of autistic children with non-autistic children. The typical children showed no preference, whereas the autistic children preferred music when given a listening choice between verbal and musical selections (Blackstock, 1978). These findings might indicate that there is a marked difference in preference and response to music between autistic children and non-autistic children (Thaut, 1987).

Production of musical patterns by children with ASD. Some studies have indicated that although children with ASD have sensory and cognitive deficits, they are able to perceive and produce musical patterns (Edgerton, 1994; Thaut, 1998). Many children with autism perform well musically in comparison with other areas of their behaviors, as well as in comparison with children without autism. One study compared sequences of either musical tones or colors spontaneously produced by autistic, mentally retarded or typical children (Frith, 1972). Musical tone sequences produced by autistic
children were more complex and varied to a significant degree, and thus superior to non-autistic children’s musical tones as well as to their color sequences (Frith, 1972).

In a more recent study, improvised musical tone sequences produced by children with autism were compared with musical improvisations by typical and mentally retarded children (Thaut, 1988). Autistic children’s improvised melodies were similar to those of normal children. Children with autism scored significantly higher than mentally retarded children on rhythm production, restriction (i.e., maintaining musical rules), originality, and total performance score. The author reported that on the given musical performance test, children with autism did not perform significantly lower than normal children (Thaut, 1988).

In summary, the reviewed studies demonstrate that many children with ASD have superior musical sensitivity and a perceptual preference for music. Research has shown that sometimes these children respond more favorably and appropriately to music than other sensory stimuli (Thaut, 1999). Furthermore, children with ASD are able to perceive and produce well-organized musical patterns, such as melodic and rhythmic patterns. Development of speech and language also depends largely on pattern perception and production. Therefore, ASD children’s intact musical abilities might influence the perception and production of patterns in speech/language. It is worthwhile to examine the effects of music on speech and language development in children with ASD. The following section will discuss the effect of music therapy on the non-musical skills of children with ASD, in particular the therapeutic effect of music on communicative behaviors including speech and language.
General effects of music therapy on autism. Since 1965, a number of researchers have examined the effect of musical activities on cognitive, social and sensory development of children with ASD. Recently, one meta-analysis and one descriptive analysis explored the general effects of music therapy interventions on individuals with ASD (Kaplan & Steele, 2003; Whipple, 2004). One researcher analyzed nine quantitative studies comparing music to non-music conditions during the treatment of children and adolescents with autism (Whipple, 2004). Results indicated that all types of music intervention, including singing, background music, social stories set to music, and directions followed with music, have been effective for children and adolescents with autism. According to the analysis, benefits from the use of music therapy included increased appropriate social behaviors such as engagement with others, and decreased stereotypical and self-stimulatory behaviors. In particular, the use of music interventions was effective for increased communicative behaviors including vocalization, verbalization, gestures, vocabulary comprehension, echolalia with communicative intention (Whipple, 2004).

Kaplan and Steele (2003) conducted a descriptive study of 40 autistic patients treated over a two-year period. Researchers found that the primary goal areas in music therapy intervention for children with ASD were language and communication (Kaplan & Steele, 2003). In addition, the most frequently utilized interventions were interactive singing and interactive instrument playing. The most common type of session for music therapy interventions was individual, which utilized direct interaction between therapist and client. The analysis also revealed that a variety of music therapy interventions was effective across a number of different treatment goals, such as behavioral/psychosocial,
language/communication, perceptual/motor, cognitive and musical goals (Kaplan & Steele 2005). In other words, all types of interventions were equally effective for general enhancement of functioning in individuals with ASD. Moreover, parents and caregivers of children with ASD reported that their children applied and generalized skills or responses acquired in music therapy to non-music therapy environments (Kaplan & Steele, 2005). The researchers, however, did not provide an explanation of how those skills or responses transferred to non-musical functions in children with ASD.

While these analyses indicated that music therapy is effective across a number of functioning domains for individuals with ASD, several critical limitations can be noted in this body of research (Kaplan & Steele, 2003; Whipple, 2004). The benefits of music were not differentiated by treatment intervention, age of participants, selection of music or goal areas (Kaplan & Steele, 2005; Whipple, 2004). These studies emphasized the general effects of music on autism. The method or outcome for the use of music in treating ASD, however, was not clearly discussed. The studies also mentioned a strong beneficial effect of music on speech and language development of children with ASD; however, they did not indicate a theoretical paradigm and did not explain how music impacts communicative behaviors in individuals with ASD. Results of these studies did not provide a rational connection between music therapy interventions and the positive outcomes.

*The positive effect of music on speech/language development.* Some of the early music therapy literature focused specifically on the effect of music on communication and language improvement in children with speech/language delays, including children with ASD. Stevens and Clark (1969) examined the effect of music therapy on socially
adaptive behaviors and communication among children with autism. Results indicated that music therapy techniques such as instrument playing, singing, or action songs were found to be significantly effective in improving social and communicative behaviors of autistic children including interpersonal verbal or non-verbal communication; involvement; and drive for mastery (Stevens & Clark, 1969).

Other researchers found that the use of contingent music and corresponding audio-visual stimuli significantly increased the frequency of verbal production including general intelligible and spontaneous speech in children with speech/language delay (Seybold, 1971; Walker, 1973). Listening to music and playing rhythm instruments appeared to have reward value for the participants. The authors explained that when pleasurable and varied stimuli are paired with appropriate verbal behavior, learning may be enhanced for children with speech/language impairments (Walker, 1973). Seybold (1971) suggested that music activities in speech therapy provided the necessary stimulation for speech delayed children to use spontaneous speech. Researchers also speculated that the positive effect of music on increased spontaneous speech might be the result of the pleasurable experience associated with music and a more comfortable atmosphere created by the musical experience (Seybold, 1971).

In more recent music and autism literature, similar results have been reported regarding the positive effect of music on speech/language in children with ASD. Antiphonal singing with picture cards resulted in a significant improvement in expressive and receptive language in children with autism (Hoskins, 1988). A combination of manual signs and singing elicited a significantly higher number of signs and spoken words imitated by children with autism, in comparison to signs and spoken words
imitated during a speech only condition (Buday, 1995). Brownell (2002) reported that the use of a musically adapted social story was an effective and viable treatment option for modifying behaviors in children with autism, such as excessive delayed echolalia, following directions, or using a quiet voice in the classroom. These findings suggest a link between music and speech/language development in children with ASD.

Conclusions from previous studies. The reviewed studies investigated more specific outcomes of music’s effect on communication and speech/language development in children with ASD. The positive effects for music in these studies were largely attributed to increased attention, enjoyment, and optimal social context (Brownell 2002; Buday, 1995; Seybold, 1971; Walker, 1973). The previous studies, however, did not describe the scientific mechanisms by which music can impact speech/language of children with ASD. Furthermore, methods and outcomes were not clearly stated in certain studies (Bettison, 1996; Cartwright & Huckby, 1972). Research is needed to explore the effect of a systematic intervention on speech/language skills in children with ASD. Such an intervention should be based on scientific evidence regarding the unique ways in which children with ASD perceive and produce music.

In conclusion, speech and language impairments have been regarded as one of the most pervasive developmental deficits in children with ASD. A number of studies have suggested that music can positively influence treatment of these deficits, due to autistic children’s musical sensitivity, preference, and abilities. The body of research examining how music affects speech production in children with ASD, however, is limited.
Need for the Study

Research regarding the influence of music on speech/language development in ASD should describe the mechanisms of music perception and production in this population. In addition, such research should explore similarities between perception and production of music and speech/language in ASD. This kind of research will explore how music can mediate change in the communicative behaviors in children with ASD.

Theoretical contribution. Music can be an effective tool for development of language and speech skills as well as non-verbal communication skills, because music is closely related in human beings to speech and language, both neurologically and developmentally (Gfeller, 1999; Michel & Jones, 1992; Thaut, 2005). The findings of the present research may augment the understanding of the perception and production of both music and speech. The link between music and speech may be verified as a result of this investigation, and the common principles and the mechanisms of both music and speech production might be explained. Furthermore, the findings of the present study might support the previous studies that have suggested the significance of integrating the two domains, music and speech, in early childhood development. As part of the same developmental sequence, music and speech are complementary and may reciprocate in many ways.

Researchers have theorized that musical skill and speech develop in a parallel fashion from adjacent areas of the brain (Michel & Jones, 1992). Some researchers have indicated a parallel neural process for both music and speech production (Brown, et al., 2004; Patel, Peretz, Tramo & Labreque, 1998). This parallel in neuro-anatomy underlies the fact that both music and speech are aural forms of communication. Music and speech
share the same acoustic and auditory parameters, including frequency, intensity, waveforms and timbre, duration, rate, contour, rhythm, and cadential factors (Thaut, 2005). The present study would support these theories by exploring basic similarities between music and speech production mechanisms. In particular, researchers have agreed that music is perceived and produced in patterns, such as pitch, melodic contour, rhythm and form. Music perception and production follow the principles of Gestalt perceptual organization (Radocy & Boyle, 2003). Pattern perception and production are also a common phenomenon for speech/language (McMullen & Saffran, 2004; Patel, 2003). Therefore, the present study examines whether perception and production of elements of music is similar to perception and production of linguistic information in speech.

Music and speech serve different functions in aural communication and expression as embedded in the neuro-anatomical structures of the human brain (Thaut, 2005). They might arise in parallel, but differentiate into what are called speech language and music language (Michel & Jones, 1992). Speech is more functional and concrete, music is more aesthetic and abstract. The results of the study might indicate that perception and production of musical elements through listening and singing can influence speech production by activating the common mechanisms involved in both music and speech. Ultimately, the present study might describe how functional vocabulary words can be produced with a combination of musical elements. Functional and concrete speech (i.e., words) might be integrated into singing, and the musical elements in songs may facilitate speech production. The present study could support a
developmental speech and language training tool through music and expand the functions of musical elements in enhancing speech.

Clinical contribution. Investigation exploring the use of music in treating ASD has to be based on scientific evidence regarding how children with ASD perceive music. Thaut (1999) indicated that the most promising explanation for musical behaviors in autism may lie in the knowledge of brain function and perceptual processes of children with ASD. Unfortunately, research exploring the reasons for the musical responsiveness of children with autism is limited. Therefore, empirical mechanisms of music perception and production in children with ASD as well as the theoretical foundations for the use of music in treating speech-language problems due to autism will be investigated in the present study.

With the many similarities between music and speech, educators, therapists, or parents often assume that music, and especially singing, is a valuable tool for the treatment of speech disorders (Thaut, 2005). The present study may provide evidence of the close link between music and language development in children with ASD, and explain the similar mechanisms in their perception and production of music and speech. Once this rationale is established, the practical use of music to enhance speech and language skills in children with ASD might be possible. The principle of music perception and production, and scientific evidence of music’s beneficial effect may justify the use of music in treating speech/language impairments in children with ASD.

The findings of the present study might be useful for music therapists who specialize in treating children with ASD to implement interventions for enhancing social communicative functions of their patients. In addition, the present study will enhance the
collaborative efforts of speech/language pathologists, special education teachers, and music therapists who practice speech-language training for children with ASD. For example, the present investigation will inform the selection of materials and interventions based on scientific evidence, and suggest more systematic implementation of a speech/language training tool through music. Therefore, the therapists or teachers can produce more consistent outcomes and follow procedures indicated by best practice.

Furthermore, the present study might provide an example of utilizing musical elements in teaching functional vocabulary words for children with ASD. Singing has been used to teach pre-academic concepts such as the alphabet, numbers or colors for typically developing children, or children with learning disabilities, including children with ASD (Gfeller, 1999). The present study might expand the functions of singing for teaching new vocabulary words and improving overall aspects of speech production in children with ASD including semantics, pragmatics, phonology and prosody. The findings might indicate that the use of musical elements is then related to improvement in speech production of children with ASD.

Throughout the present study, children with ASD can learn functional vocabulary words that they can use effectively in everyday interactions. In addition, the children might improve various speech elements such as semantics, pragmatics, phonology and prosody. These learning experiences and positive changes in speech production might result in improved communication skills, more conversations and enhanced interactions with others. Improved relationships with others and increased social-communicative experiences might help the children to gain more cognitive stimulation for optimal
development and independence. Furthermore, improved semantic or phonologic skills might be helpful to acquire reading skills for the children.

Parents of children with ASD may find the results of this study pertinent to their children’s needs; enhancing the children’s social communicative behaviors, particularly, the conventional use of speech with their parents or caregivers. Parents might use the developmental speech/language training tool from the present research in order to teach more vocabulary words at home. In conclusion, findings of this research could be applied to the treatment of children who have speech-language impairments due to ASD.

Purpose of the Study

The purpose of the present study was to explore how the perception and production of musical patterns in singing would impact the perception and production of speech and language. The study examined the effect of music as part of developmental speech-language training in the speech production of children with ASD. Specific research questions for the study:

1. Does level of speech production in children with ASD differ by training conditions: music versus speech versus no-training?

2. Does level of speech production in children with ASD differ by levels of functioning, language age, or presence of echolalia?

3. Does any interaction exist between type of training condition, levels of functioning, language age, presence of echolalia, and overall speech production in children with ASD?

4. Does any relationship exist between training conditions: music versus speech and various aspects of speech production, including: semantics, phonology, pragmatics, and prosody, in children with ASD?
The null hypotheses for these research questions are following:

1. There is no difference between the training conditions; that is, type of training condition has no effect on speech production in children with ASD.

2. There are no differences in speech production in children with ASD due to the different levels of functioning or the different language age.

3. There is no difference between presence of echolalia and absence of echolalia on speech production in children with ASD.

4. There is no interaction between type of training condition, levels of functioning, language age, presence of echolalia, and speech production in children with ASD. All mean differences between speech productions in children with ASD are explained by the main effects of training condition and/or range of symptoms.

5. There is no relationship between two training condition: music versus speech and various aspects of speech production, including: semantics, phonology, pragmatics, and prosody, in children with ASD.
CHAPTER TWO

Review of Related Literature

This chapter will review research literature that is pertinent to understanding the perception and production of speech and music in children with autism spectrum disorders (ASD). The literature review is divided into three sections. The first section will focus on the perception and production of speech in children with ASD. This section will review the main speech impairments in ASD, the connection between echolalia and development of functional speech, the Gestalt style of language acquisition in ASD, and the utilization of Gestalt principles in the development of functional speech. The second section will focus on the perception and production of music in children with ASD. This section will discuss Gestalt processing in general music perception, music perception in children with ASD, and music production in children with ASD. The second section will also explore similarities between music and language perception, as well as neuro-anatomical structures implicated in both music and speech. Finally, the third section will discuss the effects of music on speech and language development in children with ASD.

Perception and Production of Speech in Children with ASD

Speech and Language Impairments

Inadequate use of language is a defining feature of ASD. Most parents of autistic children first begin to be concerned that something is not quite right in their child’s development when early delay or regression occurs in the development of speech (Lord & Paul, 1997; Short & Schopler, 1988). Many studies which focused on the language of verbal children with ASD identified aberrant speech features, such as unusual word choice, pronoun reversal, echolalia, incoherent discourse, unresponsiveness to questions,
aberrant prosody, and lack of drive to communicate (Rapin & Dunn, 2003). Persistent
lack of speech of some individuals with ASD was attributed to the severity of their
disorders, and attendant mental retardation, rather than the possible inability to decode
auditory language. These studies determined that production and comprehension of
speech was more severely compromised in children with ASD than in mentally retarded
children matched for non-verbal cognitive level (Lord & Paul, 1997).

Findings of recent studies support the view that the language disorders of
children on the autistic spectrum have much in common with those of children with
developmental language disorders or specific language impairment (SLI) (Tager-Flusberg,
2003). More specifically, there was a subgroup of children with autism who also had
SLI. Language impairment is a co-occurring disorder in some, but not all, children with
autism. The subgroup of children with autism shared the same phenotypic features of
language impairment that characterize SLI, such as similar patterns of atypical brain
asymmetry. The language disorders with autism, however, differ strikingly by their
universally impaired pragmatics, including problems with conversational use of language,
and comprehension of discourse (Rapin & Dunn, 2003; Tager-Flusberg, 2003).

Tager-Flusberg (1981, 2003) indicated that children with ASD did not differ from
children with mental retardation in phonology or syntax, but that they had more severe
comprehension and pragmatic deficits than children with developmental language
disorders. Rapin and Dunn (2003) indicated that children with ASD also differ by the
prevalence of higher order processing disorders such as lexical or syntactic impairments,
and of impaired semantic classification of words. Rapin and Dunn suggested that these
deficits may be ascribed to problems with semantic organization, rather than to delayed
maturation of phonology and syntax. Children with ASD also have word retrieval deficit, which is semantic in nature due to impaired comprehension and formulating of discourse (Rapin & Dunn, 2003). In addition, many children with ASD are quite devoid of speech affect or prosody (Tager-Flusberg, 1997). In summary, children with ASD frequently have various kinds of language deficits, however, semantic, pragmatic and prosodic deficits are the most pervasive.

In spite of these deficits, more than half of all children with ASD show an intact ability to perceive and produce speech sounds, and do develop some level of functional speech (Schuler, 1995; Tager-Flusberg, 1997). Researchers have agreed that children with ASD clearly have language impairments; however, they also have certain components of language that are intact and spared (Tager-Flusberg, 1997; Prizant, 1995; Paul & Sutherland, 2005). Studies have shown that autistic children who do develop some functional language have relatively little difficulty acquiring the formal, rule-governed components of language, such as phonology and syntax. By contrast, certain pragmatic aspects of language, those that entail an understanding of other’s minds, are specifically and uniquely impaired in autism (Tager-Flusberg, 1997).

Tager-Flusberg (1997) identified examples of spared and impaired components of language and communication in autism. Speech sounds (i.e., phonology), linguistic form (i.e., syntax), questions for requests, turn-taking abilities, and communicative gestures are usually spared in children with ASD. However, intonation/vocal quality (i.e., prosody), linguistic function (i.e., pragmatics), semantics, questions seeking information, adding to conversational topic, and expressive gestures are usually impaired.
Early clinical descriptions of autistic language focused on its atypical characteristics, including echolalia, pronoun reversals, use of stereotyped language, and unusual meanings (Kanner, 1946). Some early researchers went on to argue that even the process of communication development was different in autism; that is, children with autism who acquire language do not follow the same stages or developmental patterns as do other children (as cited in Tager-Flusberg, 1997). Recent studies, however, have demonstrated that children with ASD exhibit much greater similarity to other children without autism. They acquire more of the computational and semantic aspects of language than was previously thought.

Tager-Flusberg and colleagues (1990) conducted a longitudinal study with six children with ASD and found that they followed the same developmental path as children with Down syndrome, and as typically developing children drawn from the extant literature. In looking at their grammatical development, both the autistic and Down syndrome children showed similar growth curves in the length of their utterances. For most of the children with autism, however, the rate of growth was slower than in typically developing children. The researchers also found that children with ASD and Down syndrome acquired grammatical structures in the same order as typically developing children (Tager-Flusberg, et al., 1990; Tager-Flusberg & Calkins, 1990).

More detailed analysis of the development patterns and a comparison of spontaneous and imitative (or echolalic) utterances suggested that even the processes involved in grammatical development in the children with ASD were similar to those of typically developing children (Tager-Flusberg & Calkins, 1990). Other longitudinal studies indicated that the lexical growth (i.e., vocabulary learning) and semantic
representations of children with ASD also showed a developmental pattern similar to typically developing children (Tager-Flusberg, 1985; 1986). The development of word use is also similar to children without autism, but it is slower (Travis & Sigman, 2001). In summary, children with autism who acquire speech follow the same pattern that typically developing children, but at a reduced rate (Paul & Sutherland, 2005).

Studies have shown that children with ASD assign words to the same conceptual categories as typically developing children do (Paul & Sutherland, 2005; Lord & Paul, 1997). For example, children with ASD recognize objects they sit on as “chairs.” Studies also showed that high-functioning autistic children use semantic groupings (e.g., bird, boat, food) in ways that are very similar to those used by non-autistic children to categorize and to retrieve words (Boucher, 1988; Lord & Paul, 1997; Minshew & Goldstein, 1993; Tager-Flusberg, 1985). In fact, some high-functioning individuals with ASD show particular interest in words. Unusual and idiosyncratic word use, such as inventing words, is frequently seen (Volden & Lord, 1991). Tager-Flusberg (1991) also suggests that autistic children are able to represent word meanings in memory, but fail to use these meanings in a conventional way in retrieval or organizational tasks.

*Abnormal Auditory Cortical Processing in Children with ASD*

Some researchers suggest that delayed speech development in children with ASD is due to problems or deficits in the neural mechanisms for the perception and production of sounds (Rapid & Dunn, 2003). Autism may be associated with abnormal auditory behavior, and abnormal perception of speech-like sounds in childhood may account for inadequate responses to sounds, and thus for language impairments typical of ASD (Boddaert, et al. 2004).
Rapin and Dunn (2003) found that autistic children with communication or language deficits had decreased glucose metabolism in the lateral temporal gyri bilaterally compared to typically developing children. The lateral temporal gyri are cortical areas crucial to auditory and language processing. Recent cortical evoked response studies provide evidence for more consistent dysfunction in the lateral surface of the superior temporal gyrus. In summary, a persistent abnormality in the secondary auditory cortex (i.e., auditory associative area) has been found in children with ASD (Rapin & Dunn, 2003).

Boddaert and colleagues (2003) studied auditory cortical processing in adults with autism using complex speech-like sounds. They found less activation of the left hemisphere temporal word processing network in autistic adults than in healthy comparison participants. The results obtained for autistic adults suggest a dysfunction of specific temporal regions specializing in the perception and integration of complex sounds. Consequently, such auditory stimuli are rarely recognized as speech. The abnormal pattern of activation found in autistic adults could reflect basic anomalies of paralinguistic auditory processing, rather than a consequence of abnormal language development.

More recently, Boddaert and colleagues (2004) conducted a positron emission tomography (PET) auditory activation study to investigate whether the abnormal cortical-auditory processing is also present in children with ASD. They found significant activation of the auditory cortex in the bilateral superior temporal gyrus in both children with ASD and children with mental retardation, while children were listening to speech-like stimuli. However, the activation pattern was different for the two groups. The
children with mental retardation showed activation of the superior temporal cortex bilaterally with left-biased asymmetry, as the investigators have previously observed in normal adults. This left hemisphere dominance, however, was not observed in the children with ASD. In addition, autistic children had additional significant activation outside the auditory cortex, including the left temporal pole, the bilateral cingulum, the bilateral posterior parietal, the cerebellar hemispheres, and the brain stem. The direct comparison between the two groups of children revealed less activation in left speech-related areas, including Wernicke’s area, in autistic children. The abnormal cortical activities in children with ASD are similar to those previously described in autistic adults (Boddaert et al., 2004).

The researchers concluded that listening to complex speech sounds induced an abnormal cortical activation including an aberrant functioning network in children with ASD (Boddaert et al., 2004). This abnormal auditory-cortical activation might be associated with a dysfunction of specific temporal regions specialized in the perception and integration of complex sounds. The areas found to be less activated by complex sounds in children with ASD are thought to be auditory association areas that are involved in word processing and are also presumed to act as an interface between word perception and long-term representation of familiar words in memory (Price, et al., 1996; Wise, et al., 2001). In the dominant hemisphere for language, these areas play a critical role in the ability to understand and produce meaningful speech. Thus, a dysfunction of left-speech-related cortical areas could contribute to the language developmental impairments observed in autism.
Many previous studies suggested that speech/language impairments in autism are behavior problems or learning disabilities. The reviewed literature regarding abnormal auditory cortical processing in children with ASD; however, indicates that the origin of speech/language deficits may be brain dysfunction. Despite these brain dysfunctions, a number of studies about speech and language development in ASD indicate that children with ASD can still learn language and develop their communication skills through speech. Researchers suggest focusing on the children’s spared capacity for language development and exploring ways to facilitate their functional speech (Lord & Paul, 1997; Prizant & Wetherby, 2005). Therefore, the following section will discuss ways to enhance speech and language development in children with ASD.

Language Acquisition in Children with ASD: Development of Functional Speech.

Echolalia. One of the most salient examples of deviant speech in autism is echolalia, the repetition of utterance with similar intonation of words or phrases that someone else has said. Echolalia may occur immediately after or significantly later than the original production of an utterance (Lord & Paul, 1997; Prizant & Wetherby, 2005; Rydell & Prizant, 1995). Immediate echolalia is produced either following immediately, or within two turns of original production, and involves exact repetition (i.e., pure echolalia) or minimal structural changes (i.e., mitigated immediate echolalia). Immediate echolalia has been considered a necessary stage of language development for verbal autistic children (Prizant & Duchan, 1981). Delayed echolalia is repeated at a significantly later time (i.e., at least three turns following original utterance, but more typically hours, days, or even weeks later), and also involves exact repetition, or minimal structural changes. The production process of delayed echolalia involves retrieval of
information from some type of long-term memory, while for immediate echolalia, short-term memory or working memory is most often implicated.

Echolalia is characteristic of at least 85 percent of children with ASD who acquire speech (Prizant, 1987; Rydell & Prizant, 1995), and once viewed as an undesirable, nonfunctional communication behavior (Koegel, Lovaas, & Schreibman, 1974; Lovaas, 1977). Some researchers considered echolalia a type of communication disorder and therefore advocated for the extinction or replacement of echolalic behaviors through the use of behavior modification procedures (Lovaas, 1977). Previous researchers defined echolalia as a pathological behavior that could interfere with cognitive and linguistic growth (Coleman & Stedman, 1974; Schreibman & Carr, 1978). More recent clinical researchers, however, beginning with Fay (1969) and elaborated by Prizant and colleagues (1981; 1983), have emphasized that immediate and delayed echolalia have functions for children with ASD (Lord & Paul, 1997; Paul & Sutherland, 2005; Prizant, et al., 2005; Prizant & Wetherby, 1993; Sundberg & Partington, 1998). These researchers describe echolalia as serving communicative functions, and explain echolalic behavior within the context of a child’s cognitive and linguistic development.

Prizant and Duchan (1981) conducted a study to discover how immediate echolalia functioned for children with ASD in interacting with familiar adults. After conducting a multilevel analysis of over 1,000 utterances from four autistic children with echolalic speech, the researchers concluded that immediate echolalia is not a meaningless behavior or undesirable symptom. The results indicated that for all four children in the study, regardless of individual communicative functioning levels, the percentage of echolalia produced with evidence of comprehension (62.75%) was significantly greater.
than the percentage produced with no evidence of comprehension (37.25%). The researchers identified six communicative functions of immediate echolalia, including: turn-taking, assertions, affirmative answers, requests, rehearsal to aid processing, and self-regulation. In summary, echolalia has been reevaluated as a developmental behavior that might accomplish communicative purposes.

Prizant (1983) further hypothesized that echolalic behavior may play a role in the acquisition of linguistic function and structure of speech among children with autism. Tager-Flusberg (1985) stated, “Echolalia and stereotyped language are now seen as primitive strategies for communicating, especially in the context of poor comprehension” (p.72). Echolalia is regarded as a speech imitation skill, and this particular speech production may predict further speech-communication development in children with ASD.

Goldstein (2002) conducted a meta-analysis of studies on communication interventions for children with ASD. The participants who benefit most from the interventions (e.g., modeling, prompting, fading or use of visual modality) seem to have better verbal imitation skills. Furthermore, children with good verbal imitation skills (i.e., echolalia) are more likely to demonstrate speech production in addition to or in parts of sign production than children with poor verbal skills. These findings suggest that a positive correlation might exist between imitation skills of children with ASD and their communication development. Children with ASD may learn communication behavior including functional use of speech through imitation.

Several studies indicate that imitating or copying patterns in a partner’s gesture, sequential movements, and behaviors is a key predictor of language outcome in children
with ASD (Charman & Baron-Cohe, 1994; Charman et al., 2003; Paul & Sutherland, 2005). Such researchers have suggested that imitation may be a developmental precursor of communicative behaviors for children with autism. Charman and Baron-Cohen (1994) found that children with autism have an intact ability to imitate a basic level of gestures and procedures (i.e., sequential behaviors). Their performance did not differ from that of children with mental retardation, and suggests that children with ASD have some ability to imitate patterns.

Charman and colleagues (2003) conducted a longitudinal study with nine children with autism and nine children with pervasive developmental disorders (PDD) to examine associations between diagnosis, joint attention, play and imitation abilities, and language outcome. The researchers demonstrated that imitation of actions on objects at age 20 months was associated with language ability at four years. Results confirmed that both joint attention and imitation were longitudinal predictors of later language ability, both receptively and expressively. These findings suggest that some children with ASD have an intact ability to recognize and imitate patterns in other people’s gestures, sequential movements, and social-communicative behaviors. Moreover, the degree of their ability to imitate can predict subsequent language development.

**Expressive language and functional speech.** From the findings of the reviewed studies, it is possible to postulate that some children with ASD have an intact ability to recognize and imitate patterns heard in other’s speech (i.e., echolalia) and other communicative behaviors (i.e., gesture, and sequential movements). Researchers suggest that 50% of individuals with autism and PDD develop communicative and language-
related cognitive abilities; that is, they can communicate through language (Prizant, Schuler, Wetherby & Rydell, 1997).

These individuals experience two general levels of language development: the emerging language level and the advanced language level (Prizant, et al., 1997). At the first level, which is referred to as the emerging and early language level, individuals show evidence of the acquisition of a conventional symbolic system for communication. They might have expressive abilities ranging from the emergence of a stable, core vocabulary of single words used with comprehension and intent, to the production of early multiword utterances or sign-symbol combinations that demonstrate the acquisition of early semantic-syntactic knowledge. At this level, echolalia and other forms of unconventional verbal behaviors (UVB) may comprise a significant portion of expressive utterances, and may be used along with single and multiword utterances to serve different communicative functions (Prizant, et al., 1997).

At the second level, which is referred to as the more advanced language level, individuals show abilities beyond emerging and early states of language acquisition. Their expressive language abilities range from the production of more grammatically complete simple utterances and different sentence types, to the use of language as part of conversational and narrative discourse. They use language both expressively and receptively as a primary mode of acquiring and conveying information, and of expressing needs and desires to others. Language production at more advanced language levels may reflect relatively sophisticated knowledge of linguistic structure (e.g., production and comprehension of different sentence types, including declaratives, questions, negatives,
and even complex sentence forms), although significant difficulties in pragmatics or social use of language may still remain (Prizant, et al., 1997; Tager-Flusberg, 1997).

Collectively, emerging and advanced language levels include individuals with autism or PDD who have achieved at least a single-word utterance state in expressive linguistic ability, as well as those who are able to use language in conversations. Many autistic individuals within this range of ability also produce echolalia and other forms of unconventional verbal behavior, or have progressed through periods of echolalia in language development (Prizant & Rydell, 1993). Prizant and colleagues (1997) demonstrated that children with ASD at the emerging language level might communicate primarily through early language forms and echolalia while children at more advanced linguistic levels might have the ability to communicate more consistently through generative and creative linguistic means. The researchers indicated that there is a continuum of communicative ability in children with ASD at language states, rather than a clear dichotomy of language state. Progressing to advanced language levels from a single-word utterance, however, is a challenging developmental process.

At the emerging language level and even at the more advanced language level, children with ASD may produce unconventional verbal behaviors (UVB) for communicative as well as non-communicative purposes (Prizant & Rydell, 1993). Researchers pointed out that the presence of echolalia is critical for enhancing language and communication development in ASD (Prizant, et al., 1997). Within the category of UVB, there is a continuum of linguistic behavior ranging from highly unconventional and non-interactive speech patterns to speech patterns that are produced with clear intent and are close enough to conventional forms that most listeners understand the speaker’s
purpose and meaning. Therefore, the communicative success of individuals at emerging language levels will vary greatly, depending on the presence of UVB and the degree of conventionality and intentionality achieved (Prizant et al., 1997).

**Gestalt Styles of Language Acquisition in Children with ASD**

Echolalia is the most common form of UVB and is the most frequently discussed characteristic of children with ASD who acquire speech (Prizant, 1983). Echolalia and other UVB have been offered as evidence of Gestalt processing of language in autism (Frith, 1989; Prizant, 1983; Rydell & Prinzant, 1995). Based on a study of typical language acquisition (Peters, 1983), researchers have referred to echolalia as a Gestalt language form (Prizant, 1983; Prizant, et al. 1997).

Gestalt language forms are unanalyzed units of speech. They refer to multiword utterances that are learned as memorized forms or whole units, but may appear to be the result of productive linguistic processes or the application of combinatorial rules (Prizant, 1983). Gestalt language forms are directly related to a Gestalt style of language acquisition, in which early utterances are comprised largely of Gestalt forms (i.e., structures or patterns). In addition, growth and progress in the acquisition of a flexible and generative language system depends on analysis and segmentation of Gestalt forms for rule induction.

In the language development of typical children, two dominant styles of acquisition have been discussed. The first style, referred to as “analytic,” is one in which children in early stage language development emphasize single words for primarily referential functions and acquire more complex language by combining elements into multiword utterances using productive rules (Prizant, 1983). In the Gestalt style of
language acquisition, children produce unanalyzed language forms or unanalyzed “chunks” with little appreciation of their internal structure or specific meaning, although the utterances may be used somewhat appropriately in communicative interactions (Prizant, 1983). Peters (1980) claimed that typical children’s use of unanalyzed chunks or deferred imitations served important functions in ongoing communicative interactions as well as in the language acquisition process.

Gestalt and analytic styles of language acquisition are not necessarily exclusive of one another. Researchers have agreed that most typical children use the analytic style, or may use both analytic and Gestalt forms (Peters, 1977; 1983; Prizant, 1983; Eysenck, 2001; Barrett, 1999). Some researchers have argued that typical children may show elements of each style to varying degrees in their language development, and alternate at different points along the continuum between primarily Gestalt and primarily analytic processing (Peters, 1977; 1980; Prizant, 1983).

Krashen and Scarcella (1978) also identified two types of linguistic patterns; prefabricated routines and prefabricated patterns, as primary strategies of Gestalt style language acquisition for typical first and second language learners. Prefabricated routines refer to memorized whole utterances or phrases which a speaker may use without any knowledge of their internal structure. Prefabricated patterns refer to partly creative and partly memorized wholes, such as memorized sentence frames with an open slot for a phrase (e.g., I want ________; This is a ________). Prefabricated routines and prefabricated patterns appear to resemble delayed echolalia and mitigated delayed echolalia, respectively (Prizant, 1983).
Researchers who study Gestalt styles and Gestalt language forms in typical children acquiring language consider such linguistic patterns to be important, if not essential, to language acquisition and social interactive growth. Use of Gestalt forms of language acquisition, such as formulaic utterances, unanalyzed chunks, and prefabricated routines, provide children with a framework for developing more complex communicative skills. These findings can be applied to investigations of language acquisition in children with ASD and used to understand characteristics and deficits in autistic children’s communicative behaviors.

Echolalic speech is mostly comprised of unanalyzed units of speech. As noted earlier, the language patterns of immediate echolalia and delayed echolalia, and interactive inflexibility are the most striking and prevalent features of communication of verbal autistic children. Each of these characteristics can be better understood as manifestations of Gestalt processing. A child with ASD who demonstrates immediate echolalia seems to be treating each repeated utterance as a unit, with a lack of appreciation of its internal constituent structure (Fay, 1973; Prizant, 1983). Even if the child demonstrates some comprehension of an utterance he or she echoes, such understanding is extremely limited (Prizant & Duchan, 1981). The child’s major strategy seems to be repeating utterances that are beyond his processing capacities, even though parts of the utterance may be recognized. This strategy is best achieved by a reproduction of the whole acoustic form, or the last section of a form, depending upon short-term memory limitations. The reproduction of the (whole) speech pattern has been associated with Gestalt or holistic modes of language processing (Prizant, 1983).
Delayed echolalia is another good example of Gestalt processing because it seems to be an effort to bring forth whole forms that were heard previously in similar situations (Prizant, 1983; Prizant, et al., 1997). Therefore, delayed echoic patterns may be manifestations of Gestalt processing at both situational (i.e., context) and linguistic levels (Lord & Paul, 1997; Prizant, 1983). Children with ASD may produce multiword utterances as whole units, with little if any knowledge or understanding of the internal structure. Subsequently, children may produce such unanalyzed units as a partial fulfillment of a situational cue or context, in which a child attempts to replicate a previous situation. Utterances in delayed echolalia might not refer to prior events, but might be a reproduction of portions or elements of the prior contexts that were retained in episodic memory (Prizant, 1983).

In addition to patterns of immediate and delayed echolalia, inflexibility in social interactive patterns (i.e., adherence to routines, stereotypic conversational openers, or incessant questioning) also provides evidence for Gestalt processing in children with ASD (Prizant, 1983). Verbal children with ASD may acquire knowledge of the structure of social interactions (e.g., “It is time to go”), but they demonstrate incompetence in handling the subtle adjustments and modifications necessary for an efficient exchange of information (e.g., “It is time to say good bye”). Such observations suggest that children with ASD may be preoccupied with the predictability of the structure of interactive exchanges or their external framework (i.e., pattern), rather than with internal content (i.e. the information shared). Prizant (1983) claimed that these rigid communicative or interactive patterns might be caused by an extreme form of Gestalt processing.
Gestalt language acquisition in children with ASD has been explained by the perception and production of patterns in their speech. Speech/language patterns of autistic children are characterized by repetition of unanalyzed forms that may be non-communicative or may be used as a means to express communicative intent. Such expressive patterns in autism may reflect Gestalt language acquisition, thus indicating an inability to analyze or segment others’ utterances and recognize their internal structure (i.e. semantic-syntactic processing). Because some people with autism appear to be limited to an extreme style of Gestalt processing, the process of language acquisition, even for higher functioning autistic children, is truly challenging. Furthermore, those who may remain primarily echolalic might not move along the continuum toward an analytic processing of language due to their cognitive limitations (Prizant, 1983).

**Echolalia and functional speech.** Greater perceptual and cognitive development in ASD probably allows some movement from a period of primarily Gestalt processing toward an analytic approach in language acquisition (Lord & Paul, 1997; Prizant, 1983). Defining this developmental process is necessary for enhancing language acquisition and speech production for children with ASD. How does the reproduction of memorized multiword units (i.e., echolalia) become the creative and generative linguistic process typically associated with the spontaneous production of multiword utterances? Prizant (1983) proposes a four-stage process of language acquisition from echolalia to spontaneous language. The four stages were established from the researcher’s observations of sequential language development in children with ASD who demonstrate echolalia. Prizant and colleagues (1997) suggested the appropriate interventions for language acquisition according to the proposed four stages of language acquisition.
The four stage process is best understood as continuous, without clear points of delineation (Prizant, 1983). In the first stage, utterances are predominantly echolalic and may fulfill a conversational “turn-taking,” while other utterances may be produced for a self-stimulatory effect. According to Prizant, most accounts of speech in this early stage show little evidence of comprehension or communicative functions. The language enhancement strategies in the first stage might include: modifying situations or contexts in which echolalia or UVB are found to be correlated; simplifying language input; and varying interaction styles. At the end of the first stage, echolalia may represent a transitional phase to more conventional communication (Prizant, et al., 1997).

In the second stage, a child’s growth of general knowledge and relationships within the environment may exceed linguistic growth, and language remains predominantly echolalic (Prizant, 1983). Due to cognitive growth and experiences in social interaction, a greater variety of functions will be served by echolalia. A child might attempt to express intentions and to comment on relationships within the environment by using echolalia. In addition, echolalia may serve as a means of behavioral self-regulation and as a rehearsal strategy. Toward the end of the second stage, a child might also apply particular strategies to break down echolalic utterances and acquire one-and two-word utterances. These strategies allow for an increased understanding of the constituent structure of utterances (i.e., speech patterns) and of the semantic relationships encoded in the utterances. Introduction of conventional, non-speech augmentative communicative means, such as pictures, gestures, or signs, may be an important strategy at this time (Prizant, et al., 1997).
The third stage is mainly characterized by increasing flexibility in language structure, and simultaneous use of both spontaneous and echolalic utterances. This stage includes a period in which similar communicative functions are served by spontaneous and echolalic utterances, followed by a decrease in echolalia with a concomitant increase in spontaneous speech. As pure echolalic speech decreases, speech repetition becomes more flexible and less rigidly produced. In other words, echolalia becomes mitigated with increased modifications and with decreased exact reproductions.

In this third stage, a child might acquire linguistic forms governed by knowledge of early semantic and syntactic rules. A further segmentation or combination of echolalic utterances also occurs, which may be part of the process of acquiring more spontaneous forms of speech patterns. More creative speech patterns begin to emerge as well. Therefore, modeling language in a context of active involvement and in synchrony with relevant action patterns is a powerful teaching strategy. Replacing pure echolalia with more conventional means to communicate, such as intra-verbal communication or verbal cueing, could be an effective strategy at this stage.

In the final stage of language acquisition, more spontaneous and flexible language is demonstrated, thus reflecting an increasing knowledge of semantic-syntactic and morphological rules. A more rule-governed and generative linguistic system might be developed in the final stage. Echolalia no longer serves primary communicative functions, because the child has become capable of operating the language system. As a result, the child is able to process language and regulate communication. Communicative functions are now served primarily by creative, spontaneous utterances. In this stage, it is essential that the language environment includes stimuli that are relevant and meaningful
in context. Decision making and choice making in daily activities might provide a motivating context for the child to learn to use more conventional forms of verbal behavior with clear consequences (Prizant, et al., 1997).

In conclusion, echolalic speech can be used to acquire functional language in children with ASD. According to Prizant’s four-stage developmental process of language acquisition, echolalia is the starting point, and therefore should be encouraged at the first stage. At the second stage, the link between echolalic response and other experiences and stimuli should be established. At the third stage, echolalia needs to be shaped by modification of speech patterns toward more flexible and spontaneous speech. Certain behavior modification strategies such as modeling, cuing and reinforcement might be helpful in this stage. At the final stage, functional and spontaneous use of language is accomplished.

Eventual segmentation of Gestalt forms and echolalic speech patterns with rule induction allow for greater creativity and generativity in the speech production of children with ASD (Prizant, 1983; Prizant, et al., 1997). Expansion and reduction of Gestalt forms of speech patterns have been shown to be effective strategies in helping children with ASD to produce more functional, rule-governed utterances (Scherer & Olswang, 1989; Sundberg & Partington, 1998). Modifying UVB or echolalia and teaching intra-verbal behaviors will enhance speech development in children with ASD (Prizant, et al, 1997; Sundberg & Partington, 1998). Gestalt style language acquisition, including pattern perception, anticipation and planning of speech production, may eventually elicit spontaneous speech.
A major achievement in language development for children with ASD is movement from single-unit communication to a combination of units, reflecting knowledge of early semantic-syntactic relationships. This goal may be approached through clear modeling of multiword utterances, in natural as well as contrived contexts, and based on both developmental and functional needs. Therefore, expansion of single-unit language in contrived and naturalistic learning settings is a necessary process in speech development for children with ASD (Prizant et al., 1997; Prizant & Wetherby, 2005; Paul & Sutherland, 2005).

Many children with ASD demonstrate progressive changes in their echolalic utterances, leading to the emergence of some degree of more creative, rule-governed linguistic behavior (Prizant, et al., 1997). Echolalia can eventually serve a variety of cognitive and social-communicative functions and may become the vehicle by which children with ASD acquire more conventional language forms (Lord & Paul; 1997; Prizant, et al., 1997). Research has indicated that UVB, including echolalia, are more likely to occur in some situational and communicative contexts than in others (Prizant, 1981; 1983; Rydell & Mirenda, 1991). It is common for echolalia to be produced initially with limited communicative intent and used later with increased intention. Due to the Gestalt style of language acquisition in ASD, it is possible that these children will rely on echolalia and repetition of speech in interacting with others.

As noted earlier, some forms of echolalia are highly unconventional, and therefore may be barriers to effective communication. Other forms, however, are clearly produced with intent and will be more recognizable because of their greater conventionality and relevance to the communicative context. In order to use echolalia in
enhancing speech and language development, intervention should involve appropriate
contexts containing communicative exchanges and activities of high motivation and
interest. Furthermore, interventions or activities must be engaged in functional language
use, cognitive learning, and independence (Prizant, et al., 1997).

*Perception and Production of Music in Children with Autism Spectrum Disorder*

Music has been used for ASD treatment for decades because of the beneficial
effects of musical stimuli and the positive musical responses in children with ASD (Thaut,
1999). Since 1965, music therapy literature has reported the positive effects of musical
activities on cognitive, social and sensory development for children with ASD. A few
music therapy studies have focused on the positive effect of music on communication and
language improvement. The following section in this chapter will review research
literature regarding music perception and production in children with ASD, and the link
between music and speech production.

*Music perception*

Music is composed of many separate yet interconnected components such as pitch,
melody, rhythm, harmony, form, timbre, and dynamics. These elements are typically
arranged in patterns and perceived as “music” (Radocy & Boyle, 2003; Berger, 2002).
These musical patterns are organized in such a way so as to bring anticipation of
incoming patterns over a temporal order (Berger, 2002). Musical pattern perception is
commonly ruled by Gestalt laws of perception (Radocy & Boyles, 2003). Gestalt
psychology emphasizes the importance of figure-ground relationships in perceptual
pattern organization. Gestalt psychologists have proposed a number of organizational
principles including: proximity, similarity, common direction, simplicity, and closure (Eysenck, 2001).

The law of proximity states that objects that are close to one another tend to be grouped together. Similarity involves the grouping of objects that share common attributes. Common direction results when either visual or aural objects appear to have the same motion trajectory. Simplicity (i.e., good continuation) states that perceptual organization will always be as good as prevailing conditions allow, and closure involves the perceptual completion of an object that is physically incomplete. These laws assist in the process of recognizing the most important sensory events and abstracting them perceptually from a less significant background of activity.

Similar perceptual processes are at work when perceiving organized patterns in music (Eysenck, 2001; Lipscomb, 1996; Radocy & Boyle, 2003). Essentially, the Gestalt perceptual laws as applied to melodic perception suggest that listeners are likely to perceive musical tones close together in time and auditory space as a melodic unit (i.e., proximity). Listeners tend to perceive similar, repeated tones as a unit (i.e., similarity), and hear a melodic sequence as moving in a common direction toward completion (i.e., common direction) based on what was heard previously. In addition, listeners are likely to perceive and organize the Gestalt, the melodic contour or pattern, in its simplest form (i.e., simplicity) (Radocy & Boyle, 2003).

Musical patterns are necessary for cortical perception (Berger, 2002). As soon as information is perceived as structured and organized musical patterns, which the brain prefers to process rather than random items, the brain begins activation in the higher
channels of cognition. Berger (2002) states that children with ASD might also utilize this perceptual-pattern organization in processing musical sounds.

*Similarities between music perception and speech/language perception*

Perception of musical elements is similar in many ways to the perception of speech/language information. Trehub and Trainor (1993) indicated that pitch-contour processes in music perception and prosodic processes in speech require a common auditory perceptual strategy. Rhythm perception is similar to perception of numbers in a metric organization (i.e. counting or dividing) (Clarke, 1987; Radocy & Boyle, 2003; Sloboda, 1985). Timbre or dynamic perception is directly involved in perception of intensity, quantity or quality of a stimulus. Lastly, memory of musical elements such as memory of melody contour or rhythm pattern is based on the perceptual organization of sequential structure. Thus, perception of music is similar to learning speech or to memory of events (Radocy & Boyle, 2003). Music perception follows the same principles of general perceptual organization, such as pattern recognition and grouping information into categorical units. The mechanism of music perception is similar to the perceptual/cognitive mechanism of speech/language information (Radocy & Boyle, 2003; Thaut, 2005). The similarities between music and speech/language suggest a link between the two domains in terms of their evolution and development.

McMullen and Saffran (2004) explain that both music and language are organized temporally, with pattern structures unfolding in time. Furthermore, both speech and music reach the perceptual system as frequency spectra arrayed as pitches, which are generated from a finite set of sounds (notes or phonemes), selected from a larger possible set of sounds. These sounds are organized into discrete categories which
facilitate recalling and representation of the sounds. In other words, auditory events in both language and music domains are subject to the process of categorical pattern perception (McMullen & Saffran, 2004).

Researchers also indicate that general auditory processing mechanisms responsible for pattern analysis are involved in the perception of both speech and music. The vast stores of knowledge pertaining to these separate domains, however, may be stored in separate places in the brain (Patel, 2003). Researchers suggest that basic similarities between speech and music learning mechanisms would be expected.

McMullen and Saffran (2004) reported that brain mapping data overwhelmingly support separate cortical regions sub-serving some aspects of musical and linguistic processing in adults. The authors also indicated that young children may bring some of the same skills to bear on learning in each domain. The brains of young children are quite plastic and show a remarkable ability to reorganize auditory events, which suggests that early experience or training has a profound effect on cortical/perceptual organization of auditory stimuli (McMullen & Saffran, 2004).

A large body of research literature regarding music and language supports the idea that their roots are indistinguishable. In particular, early perception of sound seems to involve common processes across music and language (Hafteck, 1997; McMullen & Saffrain, 2004; Michel & Jones, 1992). Furthermore, researchers have found evidence of grouping processes such as repeating the same utterance in infancy for both music and language. Both musical and spoken phrases are perceptual units for infants, and their perception is based on larger units such as pitch contour and rhythmic structure (Hafteck, 1997). This perceptual phenomenon is found in egocentric speech and spontaneous
songs of infants. Furthermore, pitch-contour processing is an important perceptual organizational device for infants not only for processing musical sequences, but also for speech sequences, especially the prosodic aspects of speech (Hafteck, 1997; Trehub & Trainor, 1993).

From early perception of sounds to the emergence of singing and speech, a close relationship between music and language development is evident. During the early stages of development, when music and speech are highly integrated, the closer link between the two domains should be encouraged (Hafteck, 1997). From the perspective of young listeners, who must learn about each system of music and speech before discovering its communicative intent, the similarities between music and speech perceptions may be heightened. Due to the similarities, children might learn how to speak and sing in close tandem; furthermore, music can incorporate singing and speech in a very structurally cohesive way (Thaut, 2005).

Neuro-anatomical commonalities between music and speech perception

More recently, investigators have tested the idea that exposure to music benefits speech/language due to perceptual similarities (Schon, Magne & Besson, 2004). One way to explore the speech-music link is to examine the neuro-anatomical relationship between the two domains. Researchers have tested patients with either speech/language deficits or music perception deficits following brain damage in order to find the common cortical regions or neural pathways involved in both music and speech. The recent development of brain-imaging techniques also offers new possibilities for testing the hypothesis that music influences other cognitive domains including speech/language.
Results of these neuro-anatomical studies might provide more conclusive answers for the link between music and speech.

Patel, Peretz, Tramo and Labreque (1998) examined the relationship between the perceptual processing of melodic and rhythmic patterns in speech and music. The researchers tested the prosodic and musical discrimination abilities of two amusic patients who had music perception deficits secondary to bilateral brain damage. Results revealed parallel perceptual processes between linguistic prosody and musical structure in the two patients. The researchers found preservation of both prosodic and musical discrimination abilities in one amusic patient, and impairments of both abilities in the other amusic patient. This particular result indicates that perception of speech prosody and melodic contour share common cognitive and neural resources, as do the perception of rhythmic patterns in speech and music (Patel, etc., 1998). In both amusic patients, the level of performance was similar across domains, suggesting that they shared the same neural resources and structures for prosody and music. The authors indicate that prosody in speech and music may overlap in the processes used to maintain auditory patterns in working memory (Patel, et al., 1998).

A case study of an aphasic patient with brain damage was conducted in order to explore the perception and memory of texts and melodies in songs (Peretz, Gagnon, Hebert & Macoir, 2004). The patient had a severe speech disorder, marked by phonemic errors and stuttering, without a concomitant musical production disorder. The patient was instructed to repeat song lines in a sung version (with words), a “la, la, la” version (singing melody without words), and a spoken version. The performance of the melody version was significantly higher than performance of the spoken version (i.e., text), but
singing did not help the patient’s articulation of syllables. In addition, the characteristics of the errors were similar in singing and speaking versions. The impairment in speech, such as the disruption of speech planning procedures, affected both speaking and singing in a similar manner. Speech disruption, however, left melodic production intact. The speech route was distinct from the spared melodic route. The researchers concluded that verbal production, whether it was sung or spoken, was mediated by the same impaired language output system. The neural pathway of speech production, however, was distinct from the spared pathway of melody production (Peretz, et al., 2004).

Researchers examined activation of the cortical system using brain-imaging while amateur musicians sang repetitions of novel melodies, sang harmonization with novel melodies, or vocalized monotonically (Brown et al., 2004). Major blood flow increased in the primary and secondary auditory cortex, primary motor cortex, frontal operculum, supplementary motor area, insula, posterior cerebellum, and basal ganglia during monotonic vocalization and melody repetition. Melody repetition and harmonization produced patterns of cortical activation very similar to each other; however, harmonization showed more intense activations in the same areas.

The results also revealed a strong overlap in cortical activation during all tasks. This overlapping brain activation suggests that participants perceived monotonic vocalization as a musical task, despite the use of a single pitch (Brown, et al., 2004). Vocalization with a steady beat might be perceived as a simple form of music. The researchers stated that this antiphonal, monotonic vocalization task should be seen as a model of some of the most important and cardinal features of music. The results also revealed that the melody repetition and harmonization tasks activated a part of the tertiary
auditory cortex which is specialized for higher-level pitch processing and the storage of pitch information in working memory.

The researchers proposed the existence of a cortical system that is specialized for imitative auditory behavior, including antiphonal imitation, and that such a system would be a foundation for audio-vocal matching functions, such as song and speech (Brown et al., 2004). Both music and speech development in children are based on a process of imitating adult role models during critical periods in brain development. The frontal operculum in Broca’s area could be the system of antiphonal imitation and the antiphonal production of song, since the system involves audio-vocal matching for both pitch and rhythm. Therefore, Broca’s area might process pitch and rhythm information in music, and transfer the information into motor aspects of vocalization. The authors conclude that the imitative function of Broca’s area is directly involved in the imitative audio-vocal processes underlying music and speech. They further suggest that singing is mediated by a vocalization system based on antiphonal imitation in the frontal operculum of Broca’s area (Brown et al., 2004).

Most recently, in order to examine a direct neuro-anatomical link between music and language, parallel generational tasks for melody and speech were compared using positron emission tomography (PET) (Brown, Martinez & Parsons, 2006). Ten right-handed amateur musicians were asked to vocally improvise melodies or sentences, while PET scans were performed to measure cortical activations. Results of the study suggested that common, parallel, and distinctive neural mechanisms for music and language might exist. The results indicate that a striking overlap exists in the brain activations between music and speech conditions. The overlapping activations were
found in the primary auditory cortex, sub-cortical auditory system and motor cortex underlying vocalization. This indicates that both music and language tasks require common perceptual processing of auditory stimuli and share neural resources for the control of phonation and articulation during singing and speaking.

The parallel activations for music and speech were found in the secondary auditory cortex that is involved in the perceptual processing of complex sound patterns, and in Broca’s area where vocal production and syntax processing are generated (Brown, Martinez & Parsons, 2006). The secondary auditory cortex is an interface area for organizing patterns of sounds and interpreting the phonology of complex sounds and then transmitting it to areas supporting the semantics of words and phrases. A parallel activation in the secondary auditory cortex was found while processing pitch, interval and scale structure in a melody generation task. In addition, Broca’s area was activated by the vocal production of melody and the sequential ordering of music which is processed in a similar manner to language syntax (Brown, Martinez & Parsons, 2006). The regions of parallel activation might share the common decoding or encoding system for music and language, or the neural system might have an “adaptive coding” capacity to process different types of information (i.e., music or language).

The distinct, non-overlapping, domain-specific activations in extra temporal areas were observed for the melody and sentence generation tasks. The authors indicate that these distinctive cortical activations might result from task operation-related differences, particularly different informational content (i.e., semantics) (Brown, Martinez & Parsons, 2006). Collectively, melody and sentence generating tasks share neural processing in the primary auditory cortex and primary motor cortex. Parallel
representations for phonological processing and sound sequences in speech and music were observed in the secondary auditory cortex and in Broca’s area. Finally, domain-specific and non-overlapping representations for distinctive functions in music and language were also found (Brown, Martinez & Parsons, 2006).

In summary, researchers have found neuro-anatomical structures that are involved in both music and speech perception and production. Broca’s area, the primary and secondary auditory cortices (i.e., bilateral temporal and right inferior frontal lobe), and the primary motor cortex are the regions involved in processing both music and speech, particularly prosody in speech (Brown, Martinez & Parsons, 2006; Brown, et. al., 2004; Patel, et. al., 1998). Even though the neural pathways are not identical between the production of music and speech, similar patterns of cortical activation have been observed in the perception and production of music and speech.

Perception of music in children with ASD

As discussed earlier, children with ASD may have abnormal auditory-cortical activation that causes a dysfunction of specific temporal regions which specialize in the processing of spoken words and the integration of complex sounds (Boddaert et al., 2004). Nevertheless, the majority of children with ASD show an intact ability to perceive and produce complex sounds, including musical sounds (Schuler, 1995; Tager-Flusberg, 1997). A distinct sensitivity and attention to music has been frequently mentioned in the literature on children with ASD (Thaut, 1999). Researchers have reported that some ASD children respond favorably and appropriately to musical sounds (Thaut, 1999). Furthermore, children with ASD are able to perceive and produce well-organized musical patterns, such as melodic and rhythmic patterns. The following section will review the
music literature regarding the music behaviors of children with ASD, including their perception and production of music.

Thaut (1987) investigated the perceptual preferences of autistic children by comparing their responses to musical and visual stimuli. Each child was given the choice of watching picture-slides or listening to music. The autistic children showed a slight preference for the musical stimulus, and stayed significantly longer in the music than in the visual condition. By contrast, children matched by chronological age and developmental age preferred the visual stimulus over the music. Most of the autistic children showed strong motor reactions whenever the music was playing (Thaut, 1987). The author indicated that there is a marked difference in preference and response to music between autistic children and typical children. The results suggest that children with autism have a perceptual preference for and an intact capacity to perceive auditory musical stimuli. In spite of their developmental deficits, children with autism are able to perceive musical sounds in a meaningful way.

Researchers have reported that children with ASD who have affective and interpersonal deficits can perceive affect in music (Heaton, Hermelin & Pring, 1999). Children with and without ASD were asked to identify the affective patterns in musical melodies by matching musical mode (i.e., major or minor key) with happy and sad faces. The results showed that children with and without ASD did not differ in their ability to identify the affective patterns in major and minor musical modes. In spite of their affective and interpersonal deficits, children with ASD showed no deficit in processing affective patterns in musical stimuli. The authors concluded that children with ASD could recognize emotional expression in music at a simple level, and that they might have
intact perception of melody patterns in music, such as a key or melodic mode. Furthermore, they appear to have an intact perceptual ability to connect the musical pattern with a corresponding affective pattern.

In summary, researchers have found that the perception of musical stimuli is intact in some children with ASD, and that they often have a perceptual preference for music. Musical stimuli are commonly organized and perceived by patterns, and this musical pattern perception is observed in children with ASD (Heaton, Hermelin & Pring, 1999; Orr, Myles & Carlson, 1999). Pattern perception is also the primary mechanism for speech and language in children with ASD (Prizant, 1983). In particular, language development in children with ASD is heavily influenced by their capacity for pattern perception and production (Prizant, 1983; Prizant, et al., 1997). Thus, researchers have suggested that pattern perception is the common phenomenon in children with ASD for processing both music and speech (Berger, 2002; Thaut, 2005).

**Pattern perception of music and speech in children with ASD**

Previous research has explored the commonalities in pattern processing and organization for both language and music. Therefore, it is worthwhile to examine how the perception and production of musical elements relates to the perception and production of the corresponding patterns in language. In addition, how children with ASD perceive and produce the primary elements or patterns in both music and speech needs to be explored. As in music, in language the presence of patterns is very evident (Berger, 2002). Every word, when divided into its syllabic rhythm, displays patterns. The functions of patterns in speech/language can be detected in musical patterns.
**Rhythmic patterns.** In music, rhythm provides the temporal ordering of sounds and is perceived as a temporal figure based on the beat patterns. Rhythmic patterns are organized by grouping and perceived as a unit. In addition, rhythmic patterns have a constant and repetitive nature (Berger, 2002). The continuity and perseverative nature of rhythm provides the perceptual elements that ultimately facilitate focused cortical activation, since the brain attends to the repetitive nature of pulse interacting with rhythmic patterns (Berger, 2002; Thaut, 2005). Rhythm is also considered one of the most important elements in the learning of spoken language (Berger, 2002). The temporal pattern in language embellishes and causes anticipation of the following pattern (Barrett, 1999; Peters, 1983).

Berger (2002) suggests that non-verbal autistic children tend to be more attentive and motivated to imitate and learn word sounds that are broken down into rhythmically patterned syllables, spoken, clapped, and/or sung. A simple verbal phrase repeated to a rhythmic pattern might sustain the children’s attention and interest in the verbal input and increase their anticipation of the following phrase.

**Pitch, melodic contour and prosody.** In music, pitch is a psychoacoustic property of tones that is perceived categorically (i.e., C, D#, G, etc). Frequency, the physiological property of tone, produces certain patterns of neural excitation in the higher auditory pathway. Listeners tend to perceive these patterns and organize them into a certain category (Radocy & Boyle, 2003). Perceived patterns are stored and recalled as pitch memory. Organized patterns of different pitches constitute a melody in music, as well as prosody of speech.
Perception of melody is based on melodic contour which is the overall shape or particular pattern of pitch movement. Melodic contour strengthened by rhythm is commonly perceived as a “Gestalt” or whole figure. Therefore, melody perception follows the principles of Gestalt perceptual organization (Lipscomb, 1999; Radocy & Boyle, 2003). In the perception of melody, hierarchical orders of melodic contour are based on the temporal structure. In other words, rhythmic pattern is a necessary perceptual unit for melody perception. Melodic pattern or contours, such as wide sweeps and leaps from low to high pitches (e.g., “Twinkle, Twinkle, Little Star”), or scale-wise melodies (e.g., “Mary Had a Little Lamb”) create different physiological and psychological states of anticipation in listeners (Berger, 2002; Radocy & Boyle, 2003). Therefore, melodies can stimulate a center in the brain that analyzes different sequences of pitch and processes the melodic pattern, such as Broca’s area. Melodic contour might prompt the center to intone the prosody of letters or words in songs, and then eventually lead to proper speech production (Berger, 2002).

The ability to analyze auditory information accurately in a melodic pattern is of vital importance in learning speech prosody. Prosody refers to the variation of tones used when speaking (i.e., intonation or pitch) and vocal stress, which is the relative emphasis given to certain syllables in a word (McCann & Peppe, 2003). The functions of prosody are to provide an indication of the speech affect and speaker’s intention. Many children with ASD lack speech affect or prosody (Tager-Flusberg, 1997). They are, however, able to imitate melodic patterns in songs and produce prosody in musical speech (Edgerton, 1994; Wigram, 2000). Children with ASD might be able to express their emotions
through musical melodies, and furthermore to produce prosodic vocal self-expression through melodic patterns.

*Dynamics.* In music, perception of sound intensity occurs via patterns of energy, resonance and level of auditory stimulation through the variance of volume (Berger, 2002; Radocy & Boyle, 2003). A listener is immediately engaged by dynamic nuances, because other musical elements such as melody or rhythm are emphasized by dynamics. Patterns of musical dynamics such as loud, soft, and gradual increases and decreases of volume contribute to the emotional content of music. In particular, dynamics indicate intention and emotion of a musical passage and thus contribute to musical prosody. Musical dynamics parallel human dynamics in terms of moods, levels of excitability, and physical and psychological states (Berger, 2002; Thaut, 2002). The ability to perceive and produce dynamics in music can be used in the perception and production of the dynamics of speech which also indicate the urgency and level of the emotional state.

In individuals with ASD, it is not unusual to find either very few dynamics in their playing of musical instruments or predominantly loud mono-dynamic pounding (Berger, 2002). By contrast, some musical expressions by children with ASD display erratic dynamic changes, such as overall soft playing, except for a sudden loud attack. These musical displays are possibly due to impairments in expression of inner states and reciprocal social interaction with others.

*Form and structure.* Musical form is another type of auditory pattern. It ultimately shapes the musical pattern development and limits the size of the pattern presentation, thus elicits the anticipation of structural closure. Musical form evokes cognitive processes for planning and organizing (Berger, 2002). Perception of musical
form is based on perceiving the repetition of a combination of musical elements and anticipating the upcoming patterns within a given time frame.

Perception of musical form might be related to the semantic and pragmatic aspects of language. Such aspects as how a tune begins, where it goes next, how long it should continue, how it can conclude, and how many repetitions would be enough, are evident in both musical and speech form. According to Berger (2002), attention is actually a state of anticipation. The brain waits and attends to information by remaining in a holding pattern until some resolution allows the processing to conclude. This particular perceptual process can be applied to semantic and pragmatic language ability.

Linguistic form and sentence structure have been used in developing language teaching methods for children with language impairments and ASD (Sundberg & Partington, 1998). Once children understand a particular form or sentence structure, they tend to perceive new elements (e.g., vocabulary words) easily within this structure (Sundberg & Partington, 1998). Repetitive learning is needed to acquire a particular language pattern, and eventually leads to linguistic memory.

In conclusion, the perception of music parallels perception of language, and these processes appear to be intact in children with ASD. Both musical stimuli and language stimuli are perceived as patterns in the sensory channels of children with ASD. Research has shown that children with ASD perceive and produce speech-sounds or words in musical songs in the same way that typical children do, in spite of their dysfunction in auditory areas (Edgerton, 1994; Wimply & Nash, 1999; Wigram, 2000). Children with ASD also appear to have intact neuro-anatomical regions for processing speech sounds, as well as musical elements, including melody, rhythm, and structure in songs.
Consequently, musical patterns may stimulate and activate intact neural processes for other types of auditory information (Berger, 2002; Thaut, 1988; Heaton, et al.; 1999).

Musical behaviors in children with ASD

Previous research has indicated that children with ASD recognize and perceive patterns in music (Berger, 2002; Heaton, Hermilin & Pring, 1999; Orr, Myles & Carlson, 1999; Thaut, 1987). Analysis of the literature indicates that the mechanism of musical behaviors in children with ASD is based on pattern perception and production. If children with ASD have an intact ability to perceive musical patterns, it is worthwhile to explore their capacity to produce well-organized musical patterns, such as melodic and rhythmic patterns.

Thaut (1988) analyzed improvised musical tone sequences produced by children with and without autism and children with mental retardation. The children were asked to play a xylophone arranged in a pentatonic scale, and were allowed to continue playing until they came to a natural ending. The produced musical tone sequences were then analyzed according to the following five criteria: rhythm, restriction, complexity, rule adherence of melodic subunits, and originality. Rhythm indicated the imposition of and adherence to temporal order, and restriction indicated the use of available musical elements. Complexity indicated the generation of recurring melodic patterns, while rule adherence of melodic subunits indicated the application of melodic patterns to the total sound sequences, and originality indicated the production of original melodic patterns.

The results suggested that the autistic children scored significantly higher than the mentally retarded children in terms of rhythm, restriction, originality, and total performance score. No significant differences were found between typical children and
autistic children in terms of rhythm, restriction, total performance score, or originality. These findings indicate that the overall performance of autistic children was not significantly different than typical children. Among the three groups of children, autistic children achieved the highest mean score on the restriction scale, and the lowest mean score on the complexity scale (Thaut, 1988).

The rhythmic scores by the autistic children might reflect a tendency to adhere rigidly to temporal rules. Their high scores for restriction show that autistic children perceive and explore the available musical material just as typically developing children do. In terms of complexity and rule adherence, the tone sequences of autistic children showed some abnormal features, resembling the performance of mentally retarded children, by being relatively short and repetitive. The low performance on complexity and rule adherence of autistic children suggest that they might have difficulty organizing and retaining complex temporal sequences. In summary, the results indicate that some children with autism are able to produce musical tone sequences containing melodic and rhythmic patterns. Moreover, musical pattern production in children with autism appears to be not much different from that of typically developing children.

The communicative and interactive behaviors of autistic children are commonly presented through their musical behaviors. Eleven children with ASD were asked to improvise musical patterns on various musical instruments, including voice, piano, snare drum, and cymbal, in an interactive music therapy setting (Edgerton, 1994). The children were able to match a fast, basic beat to the experimenter’s playing, simultaneously imitate the rhythm of a melodic motif, and participate in exchanging rhythmic patterns with the experimenter during improvisational instrument playing. A number of verbal,
vocal, and instrumental behaviors were initiated by the children in an attempt to influence the experimenter’s improvisation or musical behaviors. The analysis indicates that children with ASD are able to produce various rhythmic patterns, vocalize in a steady tempo and match the tempo of their improvisation to the experimenter’s varied tempo. In addition, they all produced vocal responses to varied pitches.

Edgerton (1994) found a significant positive correlation between the musical vocal behavior and the non-musical speech production. As musical vocal behaviors increased, non-musical speech production behaviors, such as verbal communication with gestures, also increased. Edgerton explained that communication through music might bypass the speech and language barriers of children with autism. Her study suggests that music improvisation might be effective at eliciting and increasing communicative behaviors in children with ASD within a musical setting (Edgerton, 1994).

Children with ASD have certain behavioral characteristics, and some behaviors define the disorder. Autistic children also have certain characteristics in their musical behaviors. Edgerton (1994) found communicativeness and flexibility in ASD children’s musical instrument playing. By contrast, different results were found in a more recent study. Wigram (2000) described the musical behaviors and interactions of children with ASD in order to design a music therapy assessment tool for children with autism and communication disorders. According to the author’s observation, classic rigidity was the main characteristic in musical instrument playing of the autistic children. This musical behavior was observed in repetitive and unvarying rhythm pattern production, unchanging tempo, and lack of leading during the interactive playing with others. Restricted pattern production style was presented through unvaried volume, tempo and
other musical elements (e.g., pitch range). Their musical playing was also considered to be systematic and methodical (Wigram. 2000).

According to the author, autistic children produced perseverative and repetitive scale playing on the piano, xylophone and metallophone, and made monotonous rhythm sequences on the guitar and the autoharp. In addition, children with ASD demonstrated a lack of skill or intuitive ability in turn-taking, sharing, anticipating, copying, reflecting or empathic playing (Wigram, 2000). They also showed a lack of ability to respond to or share changes in tempo, rhythm, timbre, intensity, and many other elements of shared musical engagement. Children with ASD were, however, able to copy varied rhythm patterns and imitate the changes in scales on the keyboard. With repetitive experience and practice, children with ASD produced varied rhythm figures and more flexible tempi and dynamics (Wigram, 2000).

These findings suggest that children with ASD produce musical patterns, although the patterns are restrictive. Furthermore, some autistic children perceive other people’s musical patterns, and then change their own musical patterns accordingly. Restrictive and unrefined musical behaviors might be changed or corrected through imitating less restrictive and more refined patterns. Children with ASD might replace restrictive patterns with more varied and advanced musical patterns with appropriate training and practice.

Researchers indicate that children with ASD have intact auditory areas that can process various patterns in sounds, including musical sounds and speech. The children’s similar perceptual mechanisms in music and speech have been discussed. If autistic children can produce less restricted and more varied patterns by imitating other people’s
musical patterns, it is worthwhile to apply this learning process to the production of speech patterns. Children with ASD might produce more advanced speech patterns by imitating musical patterns in a song which contains such elements of the speech patterns as vocabulary words, rhythmic patterns, prosody, and form. Perception and production of musical patterns might positively influence the perception and production of speech in children with ASD.

*The Effect of Music on Speech and Language in Children with ASD*

A small number of researchers have explored the influence of music on speech and language skills in children with language impairments, including children with ASD. Hoskins (1988) investigated the relationship between sung and spoken versions of two standardized speech tests, and examined the effect of singing on language abilities. Sixteen children with developmental delay and mental retardation ranging from 2 to 5 years old took the Expressive One-Word Picture Vocabulary Test (EOWPVT; Gardner, 1979), and the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1981) in its original form and in a modified sung version. In addition, three prerecorded music subtests designed by the investigator were used to determine participants’ rhythm imitation, pitch imitation, and melodic imitation abilities. For each subtest, participants were asked to repeat the item immediately after hearing it. Participants attended 10 weeks of music activities with an emphasis on increasing expressive language skills, including group singing (e.g., action songs) and antiphonal singing (Hoskins, 1988).

The results of pre- and post-testing showed significant differences between each of the six pretests and posttests. A statistically significant improvement in responses to the sung version of the PPVT was found in the pre- to post-test analysis. The author
indicated that the beneficial effect of antiphonal singing with picture cards might exist in the improvement of the melodic version of the language test. The sung version of the language test may have been more similar to the antiphonal singing in the music treatment, and this could indicate why significant improvement was found in the melodic version but not in the spoken version. The results also indicated a strong positive correlation between the spoken and sung versions of the PPVT. Pre- and post-tests of the spoken and sung versions of the PPVT were highly related, so that they could be said to measure the same attribute, and that no advantage existed for using one version over the other (Hoskins, 1988). Children with ASD might perceive auditory patterns equally well, whether presented in a sung version or in spoken version. The positive effects for music in this study were largely attributed to an increased attention factor (Hoskins, 1988).

Hoskins’s study demonstrates how musical patterns can be integrated with speech patterns. The spoken version of a standard speech test was transformed into a musical form, a sung version of the same test. In addition, the sung version of the speech test measured speech/language ability in children with ASD. The results suggest that musical activities such as singing could be used to enhance scores on standardized language tests (Hoskins, 1988). More recently, a study was conducted to explore the use of music as a strategy to promote better short-term memory for manual signs in children with autism (Buday, 1995). The author explained that many autistic children respond positively to music, and accounted for this effect in terms of attention to a regularly repeated musical pattern which could be easily recalled. The researcher measured the number of signs and spoken words correctly imitated in a story verse context by children diagnosed with autism. In one condition, signed and spoken words were paired with music (singing). In
the other condition, signed and spoken words were paired with rhythmic speech without music. The author hypothesized that signs taught with music would result in correct imitation of more signs and more spoken words than the signed words, which were taught with rhythmic speech (Buday, 1995).

The results showed that the average number of signs and spoken words correctly imitated during the music condition was significantly higher than the average number of signs imitated during the rhythm condition (Buday, 1995). The author suggested that music enables a child with autism to focus more intently on on-task behaviors by reducing boredom. During the music condition, the investigator observed fewer stereotyped behaviors such as hand flapping and head movement, as well as less incoherent babbling. The author also reported that another explanation for the positive results is that music provides a more enjoyable learning situation for many of the children (Buday, 1995).

Buday’s findings indicate a positive effect of music on sign and speech imitation in children with autism. This effect supports the use of music as an augmentative form of didactic language intervention (Buday, 1995). Her study, however, did not enhance the understanding of how musical patterns influence language acquisition in children with autism. In addition, Buday examined children’s imitative ability only. She did not include several critical elements of speech/language development such as pragmatics, semantics, prosody or phonology, in her measurement. In order to understand the positive effect of music on speech/language skills in children with ASD, it is necessary to examine changes in the children’s communicative behaviors including use of language after engaging in musical experiences.
If music has positive effects on enhancing communicative behaviors including speech and language, music stimuli can be used as antecedents for the verbal or non-verbal communicative responses. The effects of social versus musical antecedents on communication responsiveness in children with developmental disabilities were compared (Braithwaite & Sigafoos, 1998). Children with developmental disabilities and severe communication impairments participated in two antecedent conditions: social interaction or social interaction with musical antecedent. The social interaction condition consisted of a teacher’s initiations of verbal communication, eye contact, and facial expression. The musical antecedent condition consisted of a teacher’s singing and acoustic guitar playing with the same social interaction. In each condition, the children were given opportunities to make communication responses such as greeting, naming, and requesting. During the musical antecedent condition, each opportunity was embedded within songs.

Results of the study showed moderate increases in communication responsiveness of three children during the musical antecedent condition. In contrast, the other two children had comparable levels of appropriate communicative responding across both conditions. The results suggest that musical antecedents can facilitate communication responsiveness in some children with developmental disabilities (Braithwait & Sigafoos, 1998). According to the authors, children were more responsive to the opportunities for greeting and requesting than for naming. Appropriate responding to the naming opportunities was observed for two children who had the highest overall levels of adaptive behaviors among the five children. This finding indicates that the range of
communicative intentions expressed by children and their overall functioning levels might be related to social-communicative ability.

A number of factors in the music antecedent condition may have influenced the increased communicative behaviors. The authors suggest that motivational and attentional factors in the music had generated greater communicative responsiveness in children with developmental disabilities. Providing opportunities for communicative responses in a musical form (i.e., songs) during the musical antecedent condition can increase the probability of producing previously learned communicative responses (Braithwait & Sigafoos, 1998).

The application of communicative responses learned from musical stimuli in non-musical behaviors of children with ASD needs to be examined in order to justify the clinical use of music in ASD treatment. Brownell (2002) investigated the effect of musical social stories on the behaviors of children with autism in four experimental case studies. A social story is a short story that adheres to a specific format and guidelines to objectively describe a person, skill, event, concept, or social situation. A unique social story was created for each child that addressed current target behavior goals, such as eliminating delayed echolalia of movie and television media, following directions, and using a quiet voice. After baseline data collection, the spoken and sung versions of the social stories were alternately presented to the participants using a counterbalanced treatment order (i.e., ABAC/ACAB).

Results from all four case studies indicated that both the spoken version and the sung version were significantly more effective in addressing the target behavior than the control condition. The sung version was significantly more effective than the spoken
version in only one case, in which the target behavior was using a quiet voice. Brownell (2002) reported that the frequency of the negative target behaviors, such as excessive talking about TV, displaying difficulty in following directions and using an intensely loud shouting voice, occurred least often during the presentation of the sung social story. The findings suggest that children with ASD might comprehend a message in a song, and that this comprehension might be enhanced by perception of patterns in the presented music.

Brownell suggests that the use of a musical social story is an effective and viable treatment option for modifying behaviors with children with autism. Although the study did not focus on the effect of music on speech/language skills, the findings indicate that both reading and singing a story containing a target behavior have a similar beneficial effect on the target behavior. Consequently, children with autism might perceive the musical story and spoken story in a similar manner. Furthermore, they might be able to apply what they learned (or perceived) from either version in their subsequent behaviors.

The positive effects for music in these studies were largely attributed to increased attention, enjoyment, and optimal social context (Brownell, 2002; Buday, 1995; Hoskins, 1988). A close link exists between music and language development in children, and the common perceptual mechanisms for sound might be the primary cause for the link (Hafteck, 1997). Researchers suggest that when speech patterns or elements are paired with a melodic pattern, the speech patterns are easily produced if the melodic pattern is provided as a cue (Buday, 1995). Children with autism may prefer certain kinds of music stimuli to other kinds of stimuli because of the repetitiousness and concreteness of the patterns in music (Nelson, Anderson & Gonzales, 1984).
Some researchers have suggested that the positive effect of music stimuli on speech/language improvement might be attributed to the perception and production of patterns in both music and speech (Hafteck, 1997; Nelson, Anderson & Gonzales, 1984). A great part of discussion in their studies, however comes from speculation, and is not based on careful examination of music and speech. The effect of music on speech production and specific elements of language improvement in children with ASD has not yet been examined. The postulation for the positive effect of music on speech/language skills must be investigated and verified with scientific evidence. Moreover, this evidence will be most beneficial for the ASD population who experience abnormal speech/language development.

**Summary of the Literature Review**

Children with ASD show aberrant features of speech, including impaired pragmatics (i.e., conversational use of language) and semantics, echolalia, unresponsiveness to questions, aberrant prosody, and lack of drive to communicate. The language disorders in children with ASD may be due to problems in cortical mechanisms for perception and production of sounds including a persistent abnormality in the secondary auditory cortex. In spite of this finding, more than half of all children with ASD show an intact ability to perceive and produce speech-sounds, and develop functional speech.

One of the most salient aspects of deviant speech in autism is the occurrence of echolalia. Echolalia is regarded as a speech imitation skill that may predict further speech-communication development in children with ASD. More recent research indicates a connection between echolalia and development of functional speech, thus
Echolalia has been reevaluated as a developmental behavior that might accomplish communicative purposes. Moreover, echolalia has been offered as evidence of Gestalt processing in autism. Echolalia is a Gestalt language form which is based on pattern perception and production. The Gestalt style of language acquisition such as formulaic utterances, use of unanalyzed units of speech, and prefabricated routines, might explain characteristics and deficits of communicative behavior in autism. The Gestalt style of language acquisition and echolalic behaviors can be used to develop functional speech in children with ASD.

Music perception also reflects Gestalt processing. Combinations of musical elements, such as pitch, rhythm, melodic contour, dynamics, and form, are perceived as ‘Gestalts’ or patterns of the whole figure. Therefore, music perception is commonly ruled by Gestalt laws of perceptual pattern organization. A large body of research indicates similarities between music perception and speech/language perception. Collectively, perception and production of music and speech in children with ASD follow the same principles of Gestalt pattern perceptual organization.

Common neuro-anatomical structures and similar patterns of cortical activation mediate the perception and production of both speech and music. The perception of speech prosody and melodic contour share certain cognitive and neural resources, as does the perception of rhythmic grouping in both linguistic and non-linguistic domains. Specifically, Broca’s area is directly involved in imitative audio-vocal processes underlying both music and speech.

The perception of musical elements appears to be intact in children with ASD. They are able to perceive regulated rhythm patterns in music and to process affective
patterns in melodies. In addition, children with ASD are able to perceive and produce musical tone sequences containing melodic and rhythmic patterns. They are able to imitate varied rhythm patterns and scales on the keyboard. Musical pattern perception and production appear to be intact in children with autism, and not much different from that of typically developing children.

The positive effects of music on speech and language in children with ASD have been discussed in recent literature. Some researchers suggest that music improves communicative behaviors in children with ASD, including speech and language. The positive effects of music in the previous studies, however, were largely attributed to increased attention, enjoyment, and optimal social context. The studies did not explain the inherent musical structures or the neural mechanism for the use of musical components to enhance speech production. The present study attempts to examine the use of songs as developmental speech-language training for children with ASD. The study explores how perception and production of musical patterns in singing impacts the perception and production of speech.
CHAPTER THREE

Method

Participants

Fifty-one children were recruited from various local treatment facilities for children with Autism Spectrum Disorder (ASD), and a total of fifty children were selected to participate in this study. One child was screened with her parents’ consent; however the child did not participate in the study due to having a dual diagnosis of Autism Spectrum Disorder and Down Syndrome. Recruiting sites included: The Victory School for Children with Autism, The Kendall Speech and Language Center, Children’s Resources First Step Preschool, United Cerebral Palsy Association of Miami, Dave and Mary Alper Jewish Community Center, and Creative Children Therapy. Each child had already been diagnosed with autism spectrum disorder (ASD) by his or her own health care provider, such as a pediatric psychologist or neurologist. All children met the diagnostic criteria for autism, including qualitative impairments in social functioning and communication, and repetitive behaviors and interests (American Psychiatric Association [APA], 1994). Each child displayed the four central characteristics of ASD, including: (1) disturbances of language and communication, (2) ritualistic and compulsive behaviors and insistence on sameness, (3) disturbed social relationships, and (4) onset of the disorder prior to 30 months of age (Kanner, 1946; Paul & Sutherland, 2005; Prizant & Duchan, 1981; Tager-Flusberg, 1997).

Age range. Participants in this study ranged in age from 3 to 5 years. In typically-developing children, conventional use of language begins around 12 months. At this age, children usually say their first recognizable words and, by responding
appropriately to specific words outside the context of routine games, give the first clear evidence that they comprehend single words (Lord & Paul, 1997; Maurice, Green & Luce, 1996; Tomasello & Kruger, 1992). Twelve to 18-month-old children exhibit a rapid increase in both receptive and expressive vocabulary (Lord & Paul, 1997). From ages 3 to 5, typically-developing children begin to understand their world in more symbolic ways expressed through language, and start to comprehend abstract concepts (Schopler, Van Bourgondien & Bristol, 1993). In typically-developing children, a variety of conversational skills emerges and becomes refined as they move toward age 5 (Lord & Paul, 1997).

In contrast, most children with autism begin to speak later and develop speech at a significantly slower rate than non-autistic children (Le Couteur et al., 1989; Lord & Paul, 1997; Paul & Sutherland, 2005; Rapin & Dunn, 2003; Tager-Flusberg, 2003). Their language is severely delayed by the age of 2 years, and young autistic children’s expressive skills continue to develop at a slower rate through age 5 in comparison with children who do not have autism.

Researchers have reported that achievement of useful expressive language by the age of 5 years has been perhaps the most powerful predictor of behavioral and vocational outcomes in autism (Lord & Paul, 1997; Rutter, 1970). Many ASD researchers agree with the finding that children with autism without speech by age 5 have a poor prognosis in terms of eventual independence. Venter, Lord, and Schopler (1992) defined the presence of fluent speech in autism as at least three-word utterances produced spontaneously, regularly, and communicatively. The presence of fluent speech in children with autism by the age of 5 has predicted adaptive skills and academic
achievements in adolescence equally as well as early IQ or language tests (Venter, Lord & Schopler, 1992).

In summary, the chronological age range of 3 to 5 years appears to be a critical time for language development in children with autism (Schopler, Van Bourgondien & Bristol, 1993). Furthermore, Prizant and Wetherby (1993) indicate that language acquisition should be emphasized in early intervention and preschool services for young children with ASD. The investigator of the present study wanted to capture the critical window of time during which children with autism could develop their linguistic and communication skills. Therefore, the participants in this study included the age range from 3 to 5 years. Six participants were 3 years old, 20 participants were 4 years old, and 24 participants were 5 years old. The average age for the participants was 4 years and 8 months (See Table 1).

**Level of functioning.** The participants’ functioning level was classified as high or low. The determination of level of functioning was based on scores on either the Childhood Autism Rating Scale (CARS) (Schopler, Reichler & Renner, 1988), or the Autism Diagnostic Interview Revised (ADI-R) (Rutter, LeCouteur & Lord, 1994). All of the potential recruitment sites for the present study administered either the CARS or ADI-R; therefore, all participants in this study already had a score for overall functioning.

The CARS is a brief rating scale for autism which was developed over a 15-year period with more than 1,500 cases. This instrument is commonly used to evaluate young children over the age of 2 who may have autistic spectrum disorders (Schopler, Reichler & Renner, 1988). The internal consistency of the CARS is high, with a coefficient alpha of .94, and the reliability has a correlation coefficient of .88 (p<.01). The validity of the
CARS has a coefficient alpha of .81 (Kramer & Conoley, 1992). The CARS is used to determine various levels of functioning, such as relating to people, body use, adaptation to change, listening response, and verbal communication. The CARS is designed to rate the child on a scale from 1 to 4 in each of the functioning areas. According to the CARS, children with ASD can be identified as “age-appropriate,” “mildly impaired,” “moderately impaired,” or “severely impaired.” In this study, children who were identified “age-appropriate” and “mildly impaired” on the CARS were considered to be participants of high functioning level. Children who were identified “moderately impaired,” and “severely impaired” on the CARS were considered to be at a low functioning level.

The ADI-R is a semi-structured interview for a clinician to use with the child’s parent or principal caregiver who is familiar with the development history and current behavior of the individual being evaluated. This interview provides a thorough assessment of individuals suspected of having autism or other autism spectrum disorders, with mental age from about 18 months and above (Rutter, LeCouteur & Lord, 1994). The ADI-R focuses on three functional domains including language and communication, reciprocal social interactions, and restricted, repetitive, and stereotyped behaviors and interests. The interview is designed to collect maximal information from the parent (or caregiver), and it provides categorical results for defining autism rather than scales or norms.

Elevated scores of the ADI-R indicate problematic behavior in a particular area. For each item, the child receives a score ranging from 0 to 3. A score of 0 is given when “behavior of the types specified in the coding is not present;” a score of 1 is given when
“behavior of the type specified is present in an abnormal form, but not sufficiently severe or frequent to meet the criteria;” a score of 2 indicates “definite abnormal behavior” meeting the criteria specified, and a score of 3 is reserved for “extreme severity” of the specified behavior. In this study, children who obtained an average score of 0 or 1 for the three functioning areas on the ADI-R were considered to be at a high functioning level. Children who obtained an average score of 2 or 3 for the three functioning areas on the ADI-R were considered to be at a low functioning level.

In order to maximize the potential for learning language, children at both levels of functioning were included in the present study. Furthermore, it is not known how functioning level relates to the effects of a language development intervention for autism. In this study, 25 children were considered to be at a high functioning level and 25 children were considered to be at a low functioning level (See Table 1).

Language age and characteristics. Deficits in communication skills such as limited functional language skills were demonstrated in all of the participants. Participants in the present study already had a score on a standard verbal intelligence test of expressive and receptive language skills. The investigator used previously-administered test scores of language tests from the recruiting sites in order to determine each participant’s language age. Acceptable tests for the present study included: the Preschool Language Scale (Zimmerman, Steiner & Pond, 1992), the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), and the Expressive or Receptive One Word Picture Vocabulary Test (Gardner, 2000).

The primary goal of these standardized language tests is to identify a child’s age-equivalent linguistic level, known as language age. Language age indicates the age of
typical children who speak at about the same level as the child being tested. This level is obtained by a comparison of a child’s score on the test with the average language abilities of a large group of typically developing children. The scores on the two parts are converted to obtain the child’s chronological functioning level, and the scores are either summated (e.g., language age of 2.1 years), or presented individually (e.g., receptive functioning level of 3.2 years and an expressive functioning level of 2.4 years) (Sundberg & Partington, 1998). The three identified language tests determine language age in a comparable way, and the measurement criteria for language age are constant across the language tests.

These standardized tests are typically divided into receptive and expressive sections. Children’s language skills are tested by presenting them with increasingly complex sets of receptive and expressive tasks. The tests assess several aspects of language, including vocabulary, grammar, syntax, sentence construction, and the mean length of utterances. The tests evaluate both comprehension and production of words and sentences. These standardized language tests are required by most school districts in the United States, including potential recruiting sites for the present study, in order for a child to qualify for special speech and language services (Sundberg & Partington, 1998). Speech/language pathologists administered these tests in the potential recruiting sites, and the investigator for the present study had access to the scores for each participant.

The language characteristics of the participants as measured by the standardized tests range from “no formal means of intentional communication” to language skills seen in typical 5-year-old children.” The semantic skills in the typical language development of 5-year-old children include an average expressive vocabulary size of 1,500 words.
Their comprehension strategies include supplying the most probable missing information in answer to difficult questions, and relying on word order to process sentences. The syntax of typical 5-year-old children includes consistent grammatical morphemes, mature forms of negatives, and developed questions. For phonology, a typically-developing 5-year-old child is able to make correct sound production with good prosody and nearly 100% intelligible speech. Their pragmatic skills include referring to past/future events, and stating sequences of events. These skills also include expressed communicative functions of projecting, such as the ability to explain ideas or characteristics, and referring to or imagining what other people said (Barrett, 1999; Lord & Paul, 1997).

In this study, ten participants had a language age of 4, eight participants had a language age of 3.5, five participants had a language age of 3, and seven participants had a language age of 2.5. In addition, six participants had a language age of 2, five participants had a language age of 1.5, and nine participants had a language age of 1 (See Table 1).

Echolalia. The diagnostic criteria for autism include stereotyped and repetitive use of language or idiosyncratic language (APA, 1994). Echolalia is one example of stereotyped use of language. For a thorough description of echolalia, see Chapter Two. Prizant and Duchan (1981) found that the percentage of echolalia produced with evidence of comprehension was greater than the percentage produced with no evidence of comprehension. Furthermore, the authors predicted that echolalia is characteristic of a majority of children with ASD who acquire speech (Prizant, 1987; Rydell & Prizant, 1995). The presence of echolalic behavior was obtained from reports from both speech/language pathologists and teachers at the recruiting sites. The investigator asked
speech/language pathologists and teachers for each participant a question “does this child produce echolalia?” The researcher in the present study consulted with two speech/language pathologists and one special education teacher who specialize in working with ASD children. All of the consultants confirmed that they are easily able to determine whether or not at least some of an autistic child’s verbal productions are echolalic.

Children with and without echolalia were included in this study. Thirty-two participants produced echolalia and 18 participants did not produce echolalia. In terms of speech-language training through music, it is not known how echolalia interacts with learning of words and the speech production. Therefore, the participant’s echolalic behavior was used in the present study as an independent variable for the analysis. Please see Table 1 for the demographic information of the participants’ age, level of functioning, language age, echolalia, and gender.

Materials

Target words. The target words were selected from functional vocabularies that typically developing children can use effectively in everyday interactions (Prizant, et. al., 1997; Sundberg & Partington, 1998; Tager-Flusberg, 1997). The selection of vocabulary words in this study was based on a number of criteria, including meanings and intentions that are commonly expressed in early language communication and that can be practiced often. The general criteria for vocabulary words included: names of familiar objects and persons, functional action words such as words for social control (e.g., “stop,” “help,” “more”), and emotional words (e.g., “happy,” “sad,” “mad”) (Maurice, Green & Luce, 1996; Prizant et. al., 1997).
Table 1

*Participants’ Age, Level of Functioning, Language Age, Echolalia, and Gender*

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants (N=50)</td>
<td>18</td>
<td>18</td>
<td>14</td>
</tr>
</tbody>
</table>

Age

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 years (n=6)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4 years (n=20)</td>
<td>6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>5 years (n=24)</td>
<td>10</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

Level of Functioning

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (n=25)</td>
<td>8</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Low (n=25)</td>
<td>10</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Language Age

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year (n=9)</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>1.5 years (n=5)</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2 years (n=6)</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2.5 years (n=7)</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3 years (n=5)</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3.5 years (n=8)</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4 years (n=10)</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Echolalia

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence (n=32)</td>
<td>13</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Absence (n=18)</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Gender

<table>
<thead>
<tr>
<th></th>
<th>Music</th>
<th>Speech</th>
<th>No Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl (n=6)</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Boy (n=44)</td>
<td>15</td>
<td>16</td>
<td>13</td>
</tr>
</tbody>
</table>
Learning words for social control such as protesting or rejecting, and for expression of emotion is particularly critical in ASD and pervasive developmental disorders (PDD) (Prizant, et. al., 1997; Tager-Flusberg, 1997). Using these words demonstrates pragmatic skills, and the pragmatic aspect of language is specifically impaired in children with ASD. In addition, due to impairments in social functioning and communication, children with ASD have difficulty learning vocabulary words for social interaction and emotion. The specific criteria for vocabulary selection and target words for the present study are presented in Table 2.

The criteria for vocabulary word selection were concordant with many language assessments which have been commonly used by speech/language pathologists and teachers who work with children with autism (Gardner, 2000; Paul & Sutherland, 2005; Prizant & Bailey, 1992; Sundberg & Partington, 1998; Zimmerman, Steiner & Pond, 2006). The target words consisted of vocabulary words that a typically developing 3-year-old child would produce (Prizant, Shuler, Wetherby & Rydell, 1997; Zimmerman, Steiner & Pond, 2006). A total of 36 words were used in the present study.

Testing materials. The pre- and post-test used the 36 target words as listed in Table 2. Both tests were in the same format, and were administered individually. The pre- and post-tests were a form of fill-in-the-blank, intra-verbal communication. Intra-verbal communication is a type of expressive language in which a word or phrase evokes another word or phrase by cuing or prompting, but the cue is not identical to the response (target word) (Sundberg & Partington, 1998). If the cue and the response were identical, this would be an example of echolalia. For example, a child’s tendency to say “farm” when cued by “Old MacDonald had a…” is a form of intra-verbal communication. The
first phrase, “Old MacDonald had a…” evokes, but is not identical to the response
“farm.” The fill-in the blank, intra-verbal communication should be enjoyable and
directly relevant to the child’s ongoing interests. It should also contain words that the
child hears in his daily environment (i.e., functional vocabularies).

Each pre- and post-test included 36 phrases that were structured so as to end
with one particular target word. The investigator said the first part of each phrase, and
then left off the target word. A corresponding picture was also presented with each target
word. The pictures were selected from The Picture Exchange Communication System
(PECS) that is commonly used in teaching language for children with ASD (Frost &
Bondy, 1994). Each picture was used only one time with its corresponding target word.
The objective of the procedure in the pre- and post-test was to prompt the participant to
produce a target word at the end of a phrase upon hearing the first part of the phrase.

Each participant had two practice trials before taking each pre- and post-test. The
investigator conducted the practice trials. A practice trial consisted of two examples of
verbal modeling. In the first trial, a complete phrase with a sample word was presented
with a corresponding picture for the sample word. For example, the phrase “The monkey
wants a cookie” was presented with a corresponding picture of a cookie. Next, the same
phrase was presented, but with the sample word omitted, “The monkey wants a ____.”
The corresponding picture for the sample word “cookie” was shown after the first part of
the phrase. This process was repeated with an additional phrase and sample word for the
second practice trial. The sample words for the practice trials were not selected from the
36 target words listed in Table 2, but were commensurate in terms of complexity,
difficulty, and criteria for vocabulary word selection.
Table 2

Criteria for Vocabulary Word Selection and Target Words

<table>
<thead>
<tr>
<th>Criteria for vocabulary words</th>
<th>Target words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words that express early semantic functions:</td>
<td></td>
</tr>
<tr>
<td>Nonexistence or cessation</td>
<td>“all gone”</td>
</tr>
<tr>
<td>Recurrence</td>
<td>“more”</td>
</tr>
<tr>
<td>Action</td>
<td>“stop”</td>
</tr>
<tr>
<td>Locative action</td>
<td>“down”</td>
</tr>
<tr>
<td>Words to request motivating things:</td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>“apple”</td>
</tr>
<tr>
<td>Animals</td>
<td>“brown bear”</td>
</tr>
<tr>
<td>Objects</td>
<td>“key”</td>
</tr>
<tr>
<td>Activities</td>
<td>“sing”</td>
</tr>
<tr>
<td>Words for routine independent living activities</td>
<td>“eat”</td>
</tr>
<tr>
<td>Words for expressing agreement of affirmation</td>
<td>“yes”</td>
</tr>
<tr>
<td>Names (calling) of significant people</td>
<td>“mommy”</td>
</tr>
<tr>
<td>Words to request:</td>
<td></td>
</tr>
<tr>
<td>Assistance</td>
<td>“help”</td>
</tr>
<tr>
<td>Affection or comfort</td>
<td>“kiss”</td>
</tr>
<tr>
<td>Interaction</td>
<td>“play”</td>
</tr>
<tr>
<td>Words for common environment</td>
<td>“home”</td>
</tr>
<tr>
<td>Action words of general application</td>
<td>“point”</td>
</tr>
<tr>
<td>Words of attribution</td>
<td>“pretty”</td>
</tr>
<tr>
<td>Words to express feelings or internal states</td>
<td>“happy”</td>
</tr>
</tbody>
</table>
Music stimuli. In the present study, a music therapy technique called developmental speech and language training through music (DSLM) was used to address acquisition of the target words. “Developmental speech and language training through music (DSLM) is designed to utilize musical as well as related materials to enhance and facilitate speech and language development in children with developmental speech and language delays” (Thaut, 2005, p. 173). Therapeutic music experiences in DSLM may range from simple singing exercises to connecting sounds or pictures with words to elicit proper vocal production of the target words. Structurally and functionally organized music experiences of DSLM may enhance speech production and vocabulary acquisition in children with ASD.

Six songs composed by the investigator were used for the study. The songs included the 36 target words. Each song lyric included six target words, and each lyric line ended with a target word. For example, “Hello, hello, brown bear. What do you like to eat? I’d like to eat an apple. When I eat apples, I am happy. Brown bear says I want more. Daddy bear says the apples are all gone.” The six songs in the music condition were composed in a simple song structure with a limited number of vocabulary words. Each song was composed in a distinctly different style. For example, the songs varied by key (i.e., C Major, d minor, and E Major), tempo, and meter. Three songs were in either a 4/4 or 2/4 meter, and three songs were in 3/4, 6/8 or 12/8 meter). See Appendix A for the complete notation for the songs.

The arrangement of musical elements within the songs was developmentally appropriate for 3-year-old children. Therefore, all six songs included melodies within a limited pitch range, adjacent intervals and repetitive melodic contour (Radocy & Boyle,
Syllabic production (i.e., pronunciation) of the target words was emphasized by the rhythmic and harmonic structure as needed to preserve prosody and speech rhythm. The researcher predicted that the inherent musical structures and organized musical patterns in the six songs would provide a perceptual Gestalt to facilitate processing of both musical and speech information. Music is composed of many separate yet interconnected components, such as pitch, melody (melodic contour), rhythm, harmony, form, text (lyric), sound quality (timbre), and dynamics (musical prosody). A combination of these musical elements can be perceived as a pattern. The musical patterns in songs consist of a variety of elements organized in such a way so as to facilitate the perceptual process and anticipation of information.

All of the songs were recorded on to video tape and presented to participants through a television monitor. Each song was presented with guitar accompaniment. A female music student sang all six songs, and the investigator played the guitar. Pictures from the PECS for each of 36 target words were also presented by the singer as she sang the congruent target word. Each song was presented two times consecutively in the music video.

*Speech stimuli.* The same texts for the six songs used in the music stimuli were used for the six stories in the speech stimuli. The texts included 36 target words. Each story included six target words, and each sentence of a story ended with a target word. For example, “Look at the pink pig. He is wearing shoes. Oh, where will you go? I’m going to see the baby. I’ll give him a big hug. Babies are so much fun.” The same female music student in the music condition spoke the stories. Pictures for each of 36 target words in the stories were also presented by the vocalist as she spoke the
corresponding target word. All of the stories were recorded on to video tape and presented to participants through a television monitor. Each story was presented two times consecutively in the speech video.

Procedure

Each child’s parents learned about the purpose of the study by reading a prewritten document, and then signed the informed consent form to give permission for his/her child to be part of the study. The informed consent form gave permission to the investigator to use the participant’s standardized test scores and to ask the participant’s speech/language pathologists or teachers about the child’s use of echolalic behaviors. All screening procedures were completed by the investigator. The investigator arranged to meet the participant in a room at the recruiting site for all subsequent research procedures.

A total of 50 children with ASD participated in the study. Each participant was randomly assigned to one of three training conditions: music, speech or no-training. Eighteen children with ASD participated in the music condition, and another group of 18 children participated in the speech condition. A control group of 14 children did not receive any training. Table 1 contains the demographic distribution of the 50 participants by three training condition (See Table 1).

Each participant was tested individually. The investigator administered the same pre-test to each participant. Two practice trials were given to each participant prior to the pre-test. On the next day, the investigator met the participant and showed the appropriate video. Participants in the music condition watched the music video, while participants in the speech condition watched the speech video. Each participant watched the same video presentation two times per day (i.e., morning and afternoon) for three
days. Six training sessions were completed in a minimum of three days and a maximum of seven days. Participants in the control group did not watch the music or speech video for three consecutive days.

Edible reinforcers (e.g., crackers or cookies) and tangible prizes (e.g., rubber balls, or yo-yos) were used to keep the participants engaged during both the music and speech conditions. If the child could not have an edible reinforcer due to dietary restraints, stickers were offered. Participants received an edible reinforcer after each song or story if they stayed in the room and refrained from behaviors that interfered with their perception of the information being delivered in the video (e.g., shouting, crying, and tantruming). If participants earned edibles for each song or story, then they received a tangible prize to keep at the end of each training session.

On the next day after completing the music or speech training condition, the investigator administered the post-test. The investigator administered the same post-test individually to participants in all conditions. The training condition and pre and post–tests for each participant were completed in a minimum of five days and a maximum of nine days. The flexibility of the treatment application time allowed for changes in participants’ schedules due to absence, illness, etc.

Data collection. During the pre- and post-tests, data were collected in regard to each child’s production of the target words. A correct verbal response consisted of four components, including: (1) semantics, (2) phonology, (3) pragmatics, and (4) prosody. These speech components tend to be most impaired in children with ASD (Rapin & Dunn, 2003; Tager-Flusberg, 2003). Some components such as pragmatics and prosody tend to be especially impaired in children with Autism Spectrum Disorders (Rapin & Dunn,
A verbal production evaluation scale designed by the investigator measured participants’ productions of the target words according to the four speech components. Please see Appendix B.

The term “semantics” refers to the meanings which are encoded in language (Barrett, 1999). In the present study, semantics refers to the meaning of each spoken target word, and indicates whether the child produced the correct target word. The semantic component was determined by answering ‘yes’ or ‘no’ to the question: Does this verbal production include the correct target word? If a child produced the correct target word, he received one point.

Phonology refers to correct pronunciation of the target word. For correct pronunciation, the word produced has no articulation errors or phonetic errors in vowels or consonants (Barrett, 1999). If a child produced the word free of errors, he received one point. Pragmatics refers to the relationship between language and particular behavior, social, or linguistic contexts (Barrett, 1999). Pragmatics in speech includes communicative functions, conversation and discourse, which require individuals to take successive turns in adopting the roles of speaker and listener. Therefore, pragmatics is directly involved in speaking with coherence, relevance, sequential organization, and a correct time frame within the given context. In this study, pragmatics were measured by the temporal and sequential aspect of the verbal production. The latency of the child’s verbal response ranged from an immediate response to 10 seconds after seeing the corresponding picture (Braithwaite & Sigafoos, 1998). If a target word was produced within the designated period of time, the child received one point.
In addition, the prosody of each target word was measured. Prosody refers to the variation of tone used when speaking (i.e., intonation) and vocal stress, which is the relative emphasis given to certain syllables in a word (McCann & Peppe, 2003). The functions of prosody are to provide an indication of the identity of words. Prosody can indicate syntax, turn-taking in conversational interactions, types of utterance such as question versus statement, and the speaker’s attitudes, or intention and feelings. The elements of prosody are derived from the acoustics of speech; they include the pitch accent on stressed syllables, loudness (i.e., dynamic accent), and full vowel length (i.e., duration) of sound production (McCann & Peppe, 2003; Peppe & McCann, 2003).

All prosodic elements are present in varying qualities or quantities in every spoken utterance. The verbal production evaluation scale designed by the investigator measured how each target word was produced in terms of the prosodic elements of pitch accent, length of vowel sounds, and intensity (i.e., volume). Pitch accent was defined by whether the stressed syllables in each target word had a high or low pitch given the context of the cue phrase. For instance, the last syllable in a yes/no question has a relatively high pitch, but the last syllable in a declarative statement has a low pitch relative to the rest of the sentence. Length of vowel sounds was defined by whether the stressed syllables in each target word had fuller and longer vowel sounds than unstressed syllables and proper duration given the context of the cue phrase. For example of a sentence “I think so,” either the first syllable “I” or “i” in the word “think” has fuller and longer vowel sounds depending on the given context: emphatically (i.e., “I think so.”) or doubtfully (i.e., “I think so.”).
Intensity was defined by whether the spoken utterance for each target word had proper loudness or volume given the context of cue phrase. Typically, stressed syllables are louder than unstressed syllables (Peppe & McCann, 2003). For example, in the word “Sunday,” the stressed syllable “Sun” has fuller and louder vowel sounds than unstressed syllable “day.” Each of the three prosodic elements was worth one point; therefore, the maximum prosodic score for one target word was three points. The highest overall score for the verbal production of each of the 36 target words was 6 points, and the lowest score was 0 points. The total score possible for the pre-test or post-test was 216 points. Please see Appendix B for the verbal production evaluation scale and scoring system.

All of the pre- and post-tests were administered by the investigator and video-taped. Each test took approximately 10 minutes. Two speech/language pathologists who specialize in treating young children with language impairment and who were blind to the purpose of the study were trained for data coding. The two speech/language pathologists watched all video tapes one time and completed the coding of the four speech components on the verbal production evaluation scale simultaneously. The level of inter-rater reliability was .999, and the composite scores of the two raters were used for analysis of the collected data.

Statistical analysis. One Analyses of Covariance (ANCOVA), a two-way Analysis of Variance (ANOVA) and an Independent Samples t-test were used to analyze the data. The independent variables were: (1) type of training condition: music, speech or no training, (2) each participant’s level of functioning: high versus low, and (3) echolalia: presence versus absence of echolalia. Participant’s language age and pre-test scores were used as covariates. The dependent variables were the total score from the Verbal
Production Evaluation Scale (i.e., post-test), and the four components of verbal production: semantics, phonology, pragmatics, and prosody. A 3x2x2 ANCOVA was conducted to evaluate the effects of treatment condition, level of functioning, and ecolalia on the post-test scores of the Verbal Production Evaluation Scale after controlling for pre-test scores on the same measure and language age. A two-way ANOVA was conducted to analyze the effects of variables which were found to be significant in the initial ANCOVA. An Independent Samples $t$-Test was conducted to determine whether or not the post-test scores on the four components of speech production (i.e., semantics, phonology, pragmatics, and prosody) were significantly different between participants in the music and speech training conditions.
CHAPTER FOUR

Results

This chapter will report the results obtained from data collection and the statistical analysis of the data. The chapter will discuss the descriptive and the inferential results from the statistical analysis according to the four research questions.

Research Question # 1: Does level of speech production in children with ASD differ by training condition: music, speech, or no-training?

Descriptive analysis. To address the first research question, a descriptive analysis was completed by using the Statistical Package for the Social Sciences (SPSS) v.14.0 Software. The descriptive results, shown in Table 3, include means (M) and standard deviations (SD) of the pre-test, post-test, and the pre to post-test change scores for the total scores on the Verbal Production Evaluation Scale (VPES) of 50 participants by training condition. The total score was the cumulative score of the four speech components on the VPES which could range from 0 to 216 points.

Table 3

Means and Standard Deviations of the Pre-test, Post-test, and Change Scores for the Total Scores on the Verbal Production Evaluation Scale by Training Condition

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Pre-test Score</th>
<th>Post-test Score</th>
<th>Change Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Music</td>
<td>55.64</td>
<td>58.06</td>
<td>133.11</td>
</tr>
<tr>
<td>Speech</td>
<td>60.50</td>
<td>67.53</td>
<td>125.83</td>
</tr>
<tr>
<td>No-Training</td>
<td>42.14</td>
<td>60.89</td>
<td>43.07</td>
</tr>
</tbody>
</table>
The 18 participants in the music training condition achieved the highest mean score ($M = 133.11$) on the post-test among three training conditions, and they produced the highest change score (77.47 points). The 18 participants in the speech training condition achieved the second highest mean score ($M = 125.83$) on the post-test among the three training conditions, and they produced the second highest change score (65.33 points). The 14 participants in the no-training condition achieved the lowest mean score ($M = 43.07$) on the post-test among the three training conditions, as well as the lowest change score (0.93 points).

*Figure 1*. Mean change scores by treatment condition
Inferential analysis. A 3x2x2 analysis of covariance (ANCOVA) was conducted to evaluate the effect of training condition, level of functioning, and echolalia on the post-test scores of the Verbal Production Evaluation Scale (VPES). The ANCOVA controlled for the influence of the pre-test scores on the same measure and the participants’ language age. In this analysis, a composite of the pre-test scores for the two raters on the VPES was used as a covariate, and a composite of the post-test scores for the two raters on the VPES was used as the dependent variable. This particular analysis ensured that the effects of training condition were interpreted above and beyond the participants’ initial speech production and performance level on the VPES.

The analysis indicated a significant effect of treatment condition on the participants’ verbal productions after controlling for the pre-test score and language age, \( F (2, 37) = 21.54, p < .001, \) partial \( \eta^2 = .54 \). The effect size indicated that 54% of the variance in the participants’ verbal production was explained by the training condition, which is considered to be a large effect size (Keppel & Wickens, 2004). Therefore, the researcher failed to prove the null hypothesis: There is no difference between the training conditions; that is, type of training condition has no effect on speech production in the children with ASD.

An additional analysis was conducted to examine the effect of each training condition on the participant’s verbal productions. Dunnett’s \( t \)-test was conducted using the pre- to post-test change scores as the dependent variable, since a post-hoc test can not be used when the researcher includes a covariate. Dunnett’s \( t \)-test treats one group as a control, and compares all other groups against the control group (i.e., no-training).
The analysis indicated a significant difference between the music training condition and the no-training condition. The results also indicated a significant difference between the speech training condition and the no-training condition. According to the analysis, there was no significant difference between the music training condition and the speech training condition. Please see Table 4.

Table 4

*Analysis of Training Conditions on Change Scores: Dunnett’s t-Test (2-sided)*

<table>
<thead>
<tr>
<th>Training Conditions</th>
<th>Mean Difference</th>
<th>SE</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music vs. No-Training</td>
<td>76.544 *</td>
<td>12.973</td>
<td>.00</td>
</tr>
<tr>
<td>Speech vs. No-Training</td>
<td>64.405 *</td>
<td>12.973</td>
<td>.00</td>
</tr>
</tbody>
</table>

*Note: *The mean difference is significant at the .05 level.

The large effect size (partial $\eta^2 = .54$) and the significant $p$ value of the mean differences led the researcher to examine the individual effect size for each training condition. The effect size, Cohen’s $d$, was calculated by hand to compare the difference between training conditions. Partial Eta squared computed by SPSS showed the effect of the independent variable (training condition) on the dependent variable (verbal production). However, in order to answer the research question about which training condition has a higher effect, the researcher computed Cohen’s $d$ by hand. The results, as shown in Table 5, indicated that for the children with Autism Spectrum Disorders, music training had a large effect ($d=1.275$) on the speech production compared to the no-training condition. The speech training also had a large effect ($d=1.141$) on the speech
production compared to the no training condition. An effect size $d$ above 0.8 is considered large and $d$ from 0.2 to 0.5 is considered small (Cohen, 1977).

Table 5

*Cohen’s d Effect Size between Training Conditions*

<table>
<thead>
<tr>
<th>Training Conditions</th>
<th>Cohen’s $d$</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music vs. No-Training</td>
<td>$d = 1.275$</td>
<td>Large effect size</td>
</tr>
<tr>
<td>Speech vs. No-Training</td>
<td>$d = 1.141$</td>
<td>Large effect size</td>
</tr>
<tr>
<td>Music vs. Speech</td>
<td>$d = 0.092$</td>
<td>Negligible effect size</td>
</tr>
</tbody>
</table>

*Research Question #2: Does level of speech production in children with ASD differ by level of functioning, language age, or presence of echolalia?*

*Descriptive analysis.* In this study, 25 participants were considered to have a high level of functioning and 25 participants were considered to have a low level of functioning. Participants’ language age, as shown in Table 1, ranged from 1 year to 4 years. In this study, 32 participants produced echolalia and 18 participants did not produce echolalia. The descriptive results, as shown in Table 6, include means ($M$) and standard deviation ($SD$) for the post-test total composite score and change score on the Verbal Production Evaluation Scale (VPES), by level of functioning and echolalia. Language age was included as a covariate to further control for initial level of language ability. Thus, results of the descriptive analysis focus on the mean differences in VPES performance based on the independent variables, such as level of functioning and echolalia. The range of the total score of the VPES was from 0 to 216 points.
Table 6

*Means and Standard Deviations of the Post-test Score and the Pre to Post-test Change Score by Level of Functioning and Echolalia*

<table>
<thead>
<tr>
<th></th>
<th>Post-test Score</th>
<th>Change Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Low level of functioning</td>
<td>38.24</td>
<td>47.92</td>
</tr>
<tr>
<td>N=25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High level of functioning</td>
<td>172.32</td>
<td>50.88</td>
</tr>
<tr>
<td>N=25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Echolalia</td>
<td>135.23</td>
<td>70.65</td>
</tr>
<tr>
<td>N=32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No echolalia</td>
<td>52.03</td>
<td>79.50</td>
</tr>
<tr>
<td>N=18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The highest possible score in the post-test was 216 points.

The analysis indicated that participants with a high level of functioning achieved higher scores on the post-test ($M=172.32$) than did participants with a low level of functioning ($M=38.24$). Participants with a high level of functioning also produced the highest change score ($M$ of changes = 76.08 points). Participants with a low level of functioning produced a smaller change score ($M$ of changes = 27.26 points). The analysis further indicated that participants with echolalia achieved a higher score on the post-test ($M=135.23$) than participants without echolalia ($M=52.03$). Participants with echolalia
also produced a higher change score ($M$ of change = 64.33 points) than participants without echolalia ($M$ of change = 29.17 points).

**Inferential analysis.** The same 3x2x2 analysis of covariance (ANCOVA) was used to evaluate the effects of level of functioning and echolalia on the post-test scores of the VPES after controlling for the pre-test scores on the same measure. Language age was used as a covariate for this analysis. Results indicated a significant main effect for level of functioning, $F (1, 37) = 25.61, p < .001$, partial $\eta^2 = .41$.

The effect size (Cohen’s $d$) was calculated by hand to compare differences in level of functioning on verbal production. The effect of high level of functioning versus low level of functioning had a large effect size ($d = 1.605$). Since there was no control group for the level of functioning, it was not possible to get a separate effect size for high functioning and low functioning. The results indicated a significant difference between high and low level of functioning on the speech production in children with ASD. Therefore, the researcher failed to prove the null hypothesis: There were no differences in speech production in children with ASD due to different levels of functioning or different language age. As shown in Figure 2, the participants with a high level of functioning improved their verbal production to a greater degree than participants with a low level of functioning. Level of functioning in children with ASD might be a significant factor to indicate the level of performance and improvement in their speech production regardless of the training condition.
Figure 2. Mean change scores by level of functioning

The results of the ANCOVA indicated that the presence of echolalia did not have a significant effect after controlling for pre-test score and language age, $F(1, 37) = 0.064, p = .802$, partial $\eta^2 = <.01$. Although the participants with echolalia and without echolalia produced different scores on the pre-test and different mean changes from the pre to post-test, the difference was not statistically significant. Therefore, the researcher failed to reject the null hypothesis: There is no difference between presence of echolalia and absence of echolalia on speech production in children with ASD. However, the results of a one-way ANOVA evaluating only two variables of training condition and echolalia on the post-test scores after controlling for the pre-test score indicated that a main effect of echolalia might exist, $F(1,43) = 3.907, p = .055$. The statistical results of
this main effect of echolalia on speech production in children with ASD approached
significance.

Correlation coefficients for the independent variables and covariates were
conducted in order to examine the reliability of the present statistical design and the
relationships between various variables. Table 7 contains the results for the correlations.
This particular result provided the rationale and reliability for using language age and the
pre-test score as a covariate for the present statistical analysis.

Table 7

*Pearson Product-Moment Correlation Coefficient 2-tailed for Type of Training, Level of
Functioning, Echolalia, Language Age, and Pre-test Total Score*

<table>
<thead>
<tr>
<th></th>
<th>Training condition</th>
<th>Level of functioning</th>
<th>Echolalia</th>
<th>Language age</th>
<th>Pre-test total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training condition</td>
<td>1</td>
<td>-.05</td>
<td>-.18</td>
<td>-.121</td>
<td>-.081</td>
</tr>
<tr>
<td>Level of functioning</td>
<td>-.05</td>
<td>1</td>
<td>.417**</td>
<td>.797**</td>
<td>.699**</td>
</tr>
<tr>
<td>Echolalia</td>
<td>-.180</td>
<td>.417**</td>
<td>1</td>
<td>.596**</td>
<td>.378**</td>
</tr>
<tr>
<td>Language age</td>
<td>-.121</td>
<td>.797**</td>
<td>.596**</td>
<td>1</td>
<td>.722**</td>
</tr>
<tr>
<td>Pre-test total</td>
<td>-.081</td>
<td>.699**</td>
<td>.378**</td>
<td>.722**</td>
<td>1</td>
</tr>
</tbody>
</table>

N=50.
** Correlation is significant at the 0.01 level (2-tailed).

The analysis indicated that level of functioning, echolalia, language age and the
pre-test total score composite were positively correlated to each other and that the level of
significance was considerably high. The most highly correlated variables were language
age and level of functioning. The analysis showed that two covariates of language age
and the pre-test total score were also strongly correlated. Language age is also strongly
correlated to echolalia. In addition, there was a strong correlation between the pre-test score and level of functioning, and between the pre-test score and echolalia (See Table 6).

According to the correlation analysis, the training condition did not appear to relate to any of the other independent variables or covariates. In fact, training condition was an external variable, whereas other independent variables and covariates were internal variables acquired from the participants. The analysis indicated that the participants’ language age was positively correlated to their level of functioning, which means that the high language age and high level of functioning were strongly related in the children with ASD \( (r = .797, p = .000) \). The analysis of correlation also indicated that the participants’ language age was positively correlated to echolalia. High language age and the presence of echolalia were strongly related in the children with ASD \( (r = .596, p = .000) \). In addition, the result indicates that a high level of functioning and the presence of echolalia are strongly related in children with ASD \( (r = .417, p = .003) \).

Furthermore, the correlation analysis indicated that the participants’ pre-test score was positively related to their language age \( (r = .722, p = .000) \), level of functioning \( (r = .699, p = .000) \), and the presence of echolalia \( (r = .378, p = .007) \). This particular result indicates that a participant’s high score on the pre-test of the Verbal Production Evaluation Scale was strongly related to his/her high language age, high level of functioning, and the presence of echolalia.

Research Question # 3: Does any interaction exist between type of training condition, level of functioning, language age, presence of echolalia, and overall speech production in children with ASD?
Descriptive analysis. To address the research question, the descriptive analysis was completed by using SPSS v.14.0 Software. Table 8 contains the results for interaction of the three variables including the music training condition, echolalia, and level of functioning for the post-test total score composite.

Table 8

Descriptive Results for the Post-test by Music Training Condition, Echolalia, and Level of Functioning

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Echolalia</th>
<th>Level of Functioning</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music</td>
<td>No echolalia</td>
<td>Low</td>
<td>33.50</td>
<td>41.6133</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>215.00</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>69.80</td>
<td>88.8099</td>
<td>5</td>
</tr>
<tr>
<td>Echolalia</td>
<td>Low</td>
<td>99.92</td>
<td>31.3088</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>206.79</td>
<td>6.9121</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>157.46</td>
<td>59.2222</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Low</td>
<td>73.35</td>
<td>47.9386</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>207.81</td>
<td>7.0353</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>133.11</td>
<td>77.2261</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

The descriptive results indicated that 8 high functioning participants who received the music training achieved the highest mean score of all participants’ groups \(M = 207.813\) on the post-test. Among the 8 participants, 7 participants produced echolalia, and 1 participant did not produce echolalia. Among 10 low functioning
participants who received the music training, 6 participants produced echolalia and 4 participants did not produce echolalia, and they achieved the lowest score in the music training condition \((M=33.50)\).

Table 9

*Descriptive Results for the Post-test by Speech Training Condition, Echolalia, and Level of Functioning*

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Echolalia</th>
<th>Level of Functioning</th>
<th>(M)</th>
<th>(SD)</th>
<th>(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech</td>
<td>No echolalia</td>
<td>Low</td>
<td>.00</td>
<td>.00</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>177.667</td>
<td>36.0636</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>88.833</td>
<td>99.9493</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Echolalia</td>
<td>Low</td>
<td>56.00</td>
<td>49.00</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>173.778</td>
<td>40.8412</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>144.333</td>
<td>66.9852</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Low</td>
<td>28.00</td>
<td>43.6028</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>174.750</td>
<td>38.1138</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>125.833</td>
<td>81.0329</td>
<td>18</td>
</tr>
</tbody>
</table>

As shown in Table 9, among 12 high functioning participants who received the speech training, 9 participants produced echolalia. Three high functioning participants did not produce echolalia, and they achieved the highest score of all participants who received the speech training \((M = 177.667)\). Among 6 low functioning participants in the
speech training condition, 3 participant produced echolalia and 3 participants did not produce echolalia. The children were low functioning and did not produce echolalia achieved the lowest mean score of all participants in this study ($M = 0.00$).

Table 10

*Descriptive Results for the Post-test by No-Training Condition, Echolalia, and Level of Functioning*

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Echolalia</th>
<th>Level of Functioning</th>
<th>$M$</th>
<th>$SD$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Training</td>
<td>No echolalia</td>
<td>Low</td>
<td>7.786</td>
<td>20.5991</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>7.786</td>
<td>20.5991</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Echolalia</td>
<td>Low</td>
<td>.00</td>
<td>.00</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>109.70</td>
<td>63.3528</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>78.357</td>
<td>74.4377</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Low</td>
<td>6.056</td>
<td>18.1667</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>109.70</td>
<td>63.3528</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>43.071</td>
<td>63.9849</td>
<td>14</td>
</tr>
</tbody>
</table>

From the descriptive analysis as shown in Table 10, all of the 5 high functioning participants in the no-training condition produced echolalia. Among the 9 low functioning participants in the no-training condition, 2 participants produced echolalia and they achieved the lowest mean score of all participants ($M = 0.00$), and 7 participants did not produce echolalia.
The descriptive analysis indicated that among all of the low functioning participants in this study \((n=25)\), 6 participants who produced echolalia and received the music training achieved the highest mean score \((M=99.917)\). The analysis also indicated that 3 low functioning participants in the speech training who did not produce echolalia and 2 low functioning participants in the no-training who produced echolalia achieved the lowest mean score \((M=.00)\). The descriptive results indicated that among all of the high functioning participants in this study \((n=25)\), 84% produced echolalia \((n=21)\). In contrast, among all of the low functioning participants \((n=25)\), only 44% of the participants produced echolalia \((n=11)\). The other 56% of the low functioning participants did not produce echolalia \((n=14)\). The presence of echolalia did not appear to interact with the training condition; however, it did appear to interact with level of functioning.

Inferential analysis. The same 3x2x2 analysis of covariance (ANCOVA) was also used to examine the interaction of training condition, level of functioning, and echolalia on verbal production in children with ASD. In this analysis, a composite of the pre-test scores for the two raters on the Verbal Production Evaluation Scale (VPES) was used as a covariate, and a composite of the post-test scores for the two raters on the VPES was used as the dependent variable. This approach ensured that the effects of training condition, level of functioning, and presence of echolalia were interpreted above and beyond the participants’ initial level on the VPES. Language age was also included as a covariate to further control for initial levels of language abilities.

The analysis indicated no significant interaction between training condition and the presence of echolalia after controlling for the pre-test score and language age, \(F(2, \)
The analysis also indicated no significant interaction between training condition and level of functioning after controlling for pre-test score and language age, \( F(2, 37) = 0.65, p = .528 \). Lastly, the results of the ANCOVA indicated no significant three-way interaction between training condition, echolalia, or level of functioning after controlling for the pre-test score and language age, \( F(1, 37) = .41, p = .524 \). The researcher, therefore, failed to reject the null hypothesis: There is no interaction between type of training condition, levels of functioning, language age, presence of echolalia, and speech production in children with ASD. All mean differences between the speech production in children with ASD are therefore explained by the main effects of training condition and/or level of functioning.

**Analysis of Training Condition and Level of Functioning**

As discussed earlier for research questions 1 and 2, the results of the ANCOVA indicated a significant main effect of training condition (music, speech, and no-training), and another significant main effect of level of functioning on speech production in children with ASD. After the researcher determined the significant main effects of training condition and level of functioning, an additional analysis was completed to examine the effects of training condition and level of functioning, and the relationship between the two independent variables.

**Descriptive analysis.** The descriptive results, shown in Table 11, include means (\( M \)) and standard errors (\( SE \)) for the pre- to post-test differences on the VPES according to training condition and level of functioning. Please see Table 11 and Figure 3.
Table 11

*Mean Change Scores and Standard Error by Training Condition and Level of Functioning*

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Level of Functioning</th>
<th>Mean</th>
<th>SE</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music</td>
<td>Low</td>
<td>60.550</td>
<td>11.539</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>98.625</td>
<td>12.901</td>
<td>8</td>
</tr>
<tr>
<td>Speech</td>
<td>Low</td>
<td>15.167</td>
<td>14.897</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>90.417</td>
<td>10.534</td>
<td>12</td>
</tr>
<tr>
<td>No-Training</td>
<td>Low</td>
<td>-1.667</td>
<td>12.163</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.600</td>
<td>16.318</td>
<td>5</td>
</tr>
</tbody>
</table>

Analysis of the mean scores revealed that participants with a low level of functioning who received music training gained an average of 45.38 more points from the pre-test to the post-test than participants with a low level of functioning in the speech training condition. In contrast, participants with a high level of functioning who received music training gained an average of 8.208 more points between the pre-test and post-test compared to the participants with a high level of functioning in the speech training condition. These results indicate that an interaction may exist between the two training conditions and level of functioning in this population.

The analysis showed that the low functioning participants produced greater changes on the verbal production scores in the music training than in the speech training (See Table 11). In contrast, high functioning participants produced positive changes in
verbal production in response to both music and speech training conditions. The mean difference in verbal production between the low functioning and high functioning participants in the music training was 38.08 points. The mean difference in verbal production between the low functioning and high functioning participants in the speech training was 75.25 points. The mean difference in verbal production that the low and high functioning participants produced after the speech training was two times higher than the mean difference that the low and high functioning participants produced after the music training (See Figure 3). In other words, low functioning participants produced greater changes on the VPES after the music training than after the speech training.

As shown in Figure 3, for participants with a low level of functioning, music training elicited the greatest changes in verbal production compared to the other two training conditions: speech or no-training. For participants with a high level of functioning, music training also elicited the greatest changes in verbal production in comparison with speech training or no-training.
Figure 3. The effects of training condition and level of functioning on mean changes in verbal production.

Inferential analysis. A two-way ANOVA was conducted to investigate the effects of training condition and level of functioning on change scores for the Verbal Production Evaluation Scale (VPES). This particular analysis was conducted in order to further analyze and interpret the main effects which were found to be significant in the initial 3x2x2 ANCOVA. The results of this analysis indicated that an interaction between training condition and level of functioning may in fact exist in the population, since the statistical results of this interaction approached significance ($p = .053$).
Table 12

ANOVA for Between-Subjects Effect of Training Condition and Level of Functioning on Pre to Post-Test Changes

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sig</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Condition</td>
<td>45304.558</td>
<td>2</td>
<td>22652.279</td>
<td>17.013</td>
<td>.000</td>
<td>.436</td>
</tr>
<tr>
<td>Level of Functioning</td>
<td>18499.103</td>
<td>1</td>
<td>18499.103</td>
<td>13.894</td>
<td>.001</td>
<td>.240</td>
</tr>
<tr>
<td>TC x LF*</td>
<td>8367.135</td>
<td>2</td>
<td>4183.567</td>
<td>3.142</td>
<td>.053*</td>
<td>.125</td>
</tr>
<tr>
<td>Total</td>
<td>272726.750</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>139237.305</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Research Question #4: Does any relationship exist between training condition: music versus speech, and various aspects of speech production: semantics, phonology, pragmatics, and prosody, in children with ASD?

Descriptive analysis. To address the research question, the researcher used the participants’ separate post-test scores for each of the four speech components: semantics, phonology, pragmatics, and prosody, as a dependent variable. The descriptive results, shown in Table 13, include the means and standard deviations of the post-test score composites for each speech component within the music and speech training conditions. The maximum score for semantics, phonology, and pragmatics was 36 points each. The maximum score for prosody was 108 points.
Table 13

Means and Standard Deviations for the Post-test Composite Scores for Semantics, Phonology, Pragmatics, and Prosody by Training Condition

<table>
<thead>
<tr>
<th></th>
<th>Music Training</th>
<th>Speech Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Semantics</td>
<td>23.28</td>
<td>12.43</td>
</tr>
<tr>
<td>Phonology</td>
<td>20.94</td>
<td>13.87</td>
</tr>
<tr>
<td>Pragmatics</td>
<td>23.28</td>
<td>12.34</td>
</tr>
<tr>
<td>Prosody</td>
<td>65.56</td>
<td>39.40</td>
</tr>
</tbody>
</table>

Music and speech training conditions did not differ greatly in the mean post-test scores for each speech component. Music training produced a higher score for semantics, pragmatics, and prosody than did speech training. The post-test scores for phonology after the music and the speech training were almost the same.

**Inferential analysis.** An independent samples t-test was conducted to determine whether or not the post-test scores on the individual components of the speech production test were significantly different between participants in the music and speech training conditions. The mean performance of participants who received the music training condition versus participants who received the speech training condition on the four components of the Verbal Production Evaluation Scale (i.e., semantics, phonology, pragmatics, and prosody) were analyzed for differences. Participants who received no-training were not included in this analysis since the research question only concerned the
music and speech training conditions. In addition, an independent samples $t$-test compares only two groups at a time.

The results indicated that the difference between participants in the music and speech conditions on the semantics post-test scores was not statistically significant, $t (34) = 0.38, p = .709$. The difference between participants in the music and speech conditions on the phonology post-test scores was not statistically significant, $t (34) = -.01, p = .995$. The difference between participants in the music and speech condition on the pragmatics post-test scores was not statistically significant, $t (34) = 0.38, p = .709$. The difference between participants in the music and speech condition on the prosody post-test scores were not statistically significant, $t (34) = 0.30, p = .768$. The researcher failed to reject the null-hypothesis: There is no relationship between two training conditions: music versus speech, and various aspects of speech production, including: semantics, phonology, pragmatics, and prosody, in children with ASD.

**Target Word Analysis**

A descriptive analysis for the target word production in the pre-test and post-test scores of the Verbal Production Evaluation Scale was completed by the researcher. Target word analysis was conducted to examine what kind (i.e., category) of target words children already knew before training, and which target word production was enhanced after training. A total of 36 target words were analyzed by training condition and the number of participants who produced the target words. Tables 14, 15, 16, 17, 18, and 19 contain the 36 target words listed by song - story and the number of participants who produced each target word in the pre- and post-tests by training condition.
Table 14

*Target Words for Song-Story # 1 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test*

<table>
<thead>
<tr>
<th>Song-Story # 1</th>
<th>Music Training</th>
<th>Speech Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Words</td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
</tr>
<tr>
<td>Bear</td>
<td>7 13 8 11</td>
<td>8 11</td>
</tr>
<tr>
<td>Eat</td>
<td>6 14 5 12</td>
<td>5 12</td>
</tr>
<tr>
<td>Apple</td>
<td>11 13 11 14</td>
<td>11 14</td>
</tr>
<tr>
<td>Happy</td>
<td>5 9 6 12</td>
<td>12</td>
</tr>
<tr>
<td>More</td>
<td>5 11 3 11</td>
<td>11</td>
</tr>
<tr>
<td>All gone</td>
<td>2 12 3 14</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 15

*Target Words for Song-Story # 2 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test*

<table>
<thead>
<tr>
<th>Song-Story # 2</th>
<th>Music Training</th>
<th>Speech Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Words</td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
</tr>
<tr>
<td>Play</td>
<td>3 12 3 8</td>
<td>3 8</td>
</tr>
<tr>
<td>Sing</td>
<td>4 12 5 10</td>
<td>10</td>
</tr>
<tr>
<td>Pretty</td>
<td>3 9 2 12</td>
<td>12</td>
</tr>
<tr>
<td>Stop</td>
<td>8 13 7 11</td>
<td>11</td>
</tr>
<tr>
<td>Home</td>
<td>2 14 7 12</td>
<td>12</td>
</tr>
<tr>
<td>Mommy</td>
<td>7 12 5 13</td>
<td>13</td>
</tr>
</tbody>
</table>
Table 16

*Target Words for Song-Story #3 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test*

<table>
<thead>
<tr>
<th>Song-Story #3</th>
<th>Music Training</th>
<th>Speech Training</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
</tr>
<tr>
<td>Pig</td>
<td>10</td>
<td>14</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Shoes</td>
<td>10</td>
<td>15</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Go</td>
<td>4</td>
<td>14</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Baby</td>
<td>11</td>
<td>16</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Hug</td>
<td>7</td>
<td>13</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Fun</td>
<td>3</td>
<td>10</td>
<td>5</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 17

*Target Words for Song-Story #4 Produced by Participants in Music and Speech Training Conditions on Pre-test and Post-test*

<table>
<thead>
<tr>
<th>Song-Story #4</th>
<th>Music Training</th>
<th>Speech Training</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
<td>Pre-test (n=18)</td>
<td>Post-test (n=18)</td>
</tr>
<tr>
<td>Down</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Look</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Point</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Kiss</td>
<td>5</td>
<td>12</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Sleep</td>
<td>8</td>
<td>13</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Again</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 18

Target Words for Song-Story # 5 Produced by Participants in Music and Speech
Training Conditions on Pre-test and Post-test

| Song-Story # 5 | Music Training | | | Speech Training | | |
|---------------|---------------|---------------|---------------|---------------|---------------|
| Target Words  | Pre-test (n=18) | Post-test (n=18) | Pre-test (n=18) | Post-test (n=18) | |
| Story         | 3             | 9             | 3             | 9             | |
| School        | 7             | 13            | 5             | 10            | |
| Key           | 7             | 14            | 9             | 13            | |
| Come on       | 2             | 9             | 3             | 8             | |
| Here          | 2             | 6             | 2             | 6             | |
| All Done      | 3             | 14            | 4             | 12            | |

Table 19

Target Words for Song-Story # 6 Produced by Participants in Music and Speech
Training Conditions on Pre-test and Post-test

| Song-Story # 6 | Music Training | | | Speech Training | | |
|---------------|---------------|---------------|---------------|---------------|---------------|
| Target Words  | Pre-test (n=18) | Post-test (n=18) | Pre-test (n=18) | Post-test (n=18) | |
| Out           | 5             | 9             | 2             | 8             | |
| Yes           | 3             | 9             | 2             | 7             | |
| Banana        | 12            | 13            | 11            | 14            | |
| Dirty         | 6             | 11            | 4             | 11            | |
| OK            | 2             | 8             | 2             | 10            | |
| Help          | 2             | 13            | 5             | 7             | |
Pre-test results. Target word analysis indicated that the most produced target words in the pre-test before music training were “apple” (n=11), “baby” (n=11), and “banana” (n=12); and before speech training were also “apple” (n=11), “baby” (n=10), and “banana” (n=11). The analysis showed that approximately 60% of all participants had already known and correctly produced “apple,” “baby,” and “banana” before any type of training. According to the criteria for vocabulary word selection which were used in the method for the present study, both “apple” and “banana” are the words to request motivating food (See Table2). The target word “baby” falls under the category of names (calling) of significant people. However, the target word “mommy,” which is in the same category as “baby,” was produced by only 30% of the participants at the pre-test.

The least produced target words in the pre-test before music training were: “all gone,” “home,” “look,” “point,” “again,” “come on,” “here,” “OK,” and “help.” The least produced target words in the pre-test before speech training were “again,” “pretty,” “look,” “point,” “here,” “out,” and “dirty.” Only 2 participants out of 18 in each training condition correctly produced these target words. Three words: “look,” “point” and “here,” were the common target words produced least often before the both music and speech training. According to the criteria for vocabulary word selection, both “look,” and “point” are action words of general application. The word “here” is in the category of words to request assistance, and the target word “help” under the same category as “here,” was also produced in the pre-test by a small number of the participants.

Post-test results. The most produced target words after music training were “baby,” “shoes,” “eat,” “home,” “pig,” “key,” and “all done.” A total of 25 target words were produced by more than 50% of the participants (more than 9 participants)
after the music training condition. The least produced target words after music training were “again,” and “here.” A total of five words were produced by less than 50% of the participants after the music training condition. The most produced target words after speech training were “apple,” “all gone,” “pig,” “shoes,” “kiss,” and “banana.” A total of 26 target words were produced by more than 50% of the participants (more than 9 participants) after the speech training. The least produced target words after speech training were “look,” “point,” and “here.” A total of nine words were produced by less than 50% of the participants after the speech training.

Among the *Criteria for Vocabulary Word Selection* (See Table 2), action words of general application (i.e., “point,” and “look”), words that express early semantic functions of recurrence (i.e., “again”), and words to request assistance (i.e., “here”) were produced by the least number of children with ASD before any training. The action words of general application (i.e., “point,” and “look”), words to request interaction (i.e., “play,” and “come on”), and words to request assistance (i.e., “help,” and “here”), and words for expressing agreement of affirmation (i.e., “yes,” and “OK”) were least produced by the children with ASD after the speech and/or music training conditions.
The purpose of the present study was to explore how the perception and production of musical patterns in singing would impact the perception and production of speech and language in children with Autism Spectrum Disorders (ASD). The study examined the effect of developmental speech-language training through music on the speech production of children with ASD. A total of 50 children with ASD completed the pre-test, six sessions of training, and the post-test within 5 to 9 days. The pre-and post-tests measured each participant’s verbal production including semantics, phonology, pragmatics, and prosody, of the 36 target words.

In this chapter, the results of the statistical analysis will be interpreted. Findings from the results will be discussed in depth and will be compared with findings in the previously reviewed literature. The clinical application of the results will be also discussed in this chapter, as will the limitations of the study and recommendations for future research.

Discussion of the Research Questions

The effect of music and speech training on speech production in children with ASD. The results of the present study showed that participants in both music and speech training increased their scores on the Verbal Production Evaluation Scale from the pre-test to the post-test. This finding suggests that both music and speech training are effective for enhancing speech production, including semantics, phonology, pragmatics, and prosody in children with ASD. Participants who received music training gained a higher mean difference than participants who received the speech training; however, the
difference was not statistically significant. In summary, music training is as effective as speech training for improving speech production.

The results of the present study suggest that children with ASD perceive and produce the linguistic information from music as they do from speech. This finding is consistent with conclusions from previous research indicating that the mechanism of music perception is similar to the perceptual/cognitive mechanisms of speech/language information (Radocy & Boyle, 2003; Thaut, 2005). McMullen and Saffran (2004) found that both music and language follow the same principles of perceptual organization, such as pattern perception and recognition. Both musical and speech pattern perception are ruled by Gestalt laws of perception and organized temporally, with pattern structures unfolding in time (Mcmullen & Saffran, 2004; Radocy & Boyle, 2003).

In the present study, the musical patterns and the speech patterns were organized temporally in six songs and six stories, and all of the phrases in the six songs and stories were identical. In addition, the verbal information (or content) in both music and speech conditions utilized a common Gestalt law of perception; the law of good continuation, which involves perceptual completion. The target word in both the music and speech stimuli was placed at the end of each phrase, which allowed the children to anticipate the location and time of the target words to be produced. This research suggests that children with ASD appear to have an intact ability to perceive musical and speech patterns, and they utilize the same principle of Gestalt pattern perceptual organization to produce speech.

The findings from this study also support a close relationship between music and language in early childhood development (Hafteck, 1997). Hafteck indicated that
during the early stages of development, music and speech are highly integrated; therefore the close link between the two domains can be utilized for a child’s language development. The present study found that 3 to 5-year-old children with ASD perceived the organized musical patterns in songs and subsequently produced functional speech and language. They perceived the speech information that was attached to the musical patterns in songs, and produced only speech patterns without producing musical patterns. This finding indicates similarities or associations between music and speech perception for children, and the effects of music on their speech production. Furthermore, the findings of the present study suggest the merit of using music with children who have developmental disorders, including ASD.

The results of the study suggest that children with ASD are able to perceive linguistic information, including semantics, phonology, pragmatics and prosody, that is organized by musical patterns. The results also suggest that children with ASD are able to transform the information perceived in musical patterns into speech patterns. In the present study, children with ASD who received music training perceived musical stimuli (i.e., six songs); however, they produced speech as the outcome, which is a non-musical, functional behavior. This particular result suggests that the children who received music training had a certain degree of perceptual association between music and speech. Consequently, the children with ASD were able to produce the linguistic information in speech patterns, which was transformed from the musical stimuli.

The effects of training can be found not only in the speech production of children with ASD, but also in other functional behaviors. According to the researcher’s observations, music and speech training affected various aspects of the participants’
behaviors during training sessions and during the post-test. The procedure of watching the same video six times for either music or speech training seemed to increase participants’ attention to the stimulus. This attention increase was evidenced by changes in observed behaviors during training sessions and post-testing sessions for both training conditions.

For example, participants showed more necessary behaviors for learning during training such as independent or spontaneous sitting and maintaining an upright body posture. The children also made more eye contact with the investigator while she was giving instructions, and looked more at the TV screen to watch the video during training. The participants were gradually able to follow the investigator’s directions without verbal or physical prompts. Over the six training sessions, participants demonstrated more behaviors indicating that they were paying full attention to the stimuli, and decreased stereotypical behaviors, such as rocking motions, repetitive movements and hand flapping.

These behavior changes became more evident as the number of training sessions increased. The duration of time spent looking at the pictures also increased during the post-test, compared to the pre-test for participants who watched either the music or speech video. In addition, participants who spontaneously repeated the target words after hearing them while watching the video (either music or speech), tended to achieve a higher score on the post-test compared to participants who did not verbally repeat the target words.

The procedure of the present study is considered to be valuable for eliciting positive changes in participants’ behaviors. The consistent repetition of the procedure
appeared to elicit behaviors indicating increased attention of the participants, and
improvements in their speech production following either training condition. Although
training lasted only 3 days, the same stimuli consistently presented in the same way (i.e.,
a video to watch) resulted in increased attention and gains in functional behaviors in
children with ASD. Participants also demonstrated these behavior changes and practice
effects during the post-test. Participants showed increased attention and familiarity with
the post-testing stimuli (i.e., target words and pictures) which had already been presented
in either the music or speech video during the training sessions. In summary, familiarity
with the stimuli and repetition of stimuli presentation might be a critical factor for
producing positive changes in the functional behaviors of children with ASD.

The effect of music training on speech production in children with ASD. The
researcher’s observations suggest that participants who watched the music video
responded to the target words and pictures more favorably than participants who watched
the speech video. Participants in the music training maintained a seated position longer
than participants in the speech training, and they showed fewer disruptive behaviors (e.g.,
shouting, crying, and tantruming) while watching the music video, than participants who
watched the speech video. In addition, participants in the music training spontaneously
produced a higher number of target words during and after watching the music video than
participants in the speech training. Some of the participants in the music training sang
some phrases in the songs and shouted target words during training sessions.

In addition, participants’ anticipation for the beginnings and endings of the six
songs during the music training was greater than the anticipation for the stories in the
speech training. This anticipation was evidenced by children who sang phrases or target
words ahead of the singer in the video. For example, during the first song, some children in the music training condition sang or shouted “brown bear” while the singer in the video was still singing the previous phrase “hello, hello.” Some participants in the music training condition demonstrated a certain association for the songs by recalling the first phrase of each song. For example, just before the last song played, some children recalled the last song by shouting “monkey song.” Even though the first target word of the last song was “out,” they recalled the first phrase of the song “Hello, Mr. Monkey” without seeing the picture of the monkey. In summary, the music training seemed to garner more attention from the children with ASD than the speech training did. The children in the music training demonstrated more advanced cognitive skills from the increased attention such as anticipation, association, and recall of the stimuli during the training session and the post-test compared to children in the speech training.

This particular observation of increased attention in the music condition is consistent with findings of previous studies conducted by Hoskins (1988) and Buday (1995). Hoskins (1988) suggested that the positive effects of music on vocabulary learning were largely attributed to an increased attention factor. Buday (1995) found that the average number of signs and spoken words correctly imitated during a music condition was significantly higher than the average number of signs imitated during the speech condition. The author suggested that music helps children with ASD to produce more on-task behaviors by reducing boredom and providing a more enjoyable learning situation (Buday, 1995).

The characteristics of the musical stimuli could also help explain the superior performance in speech production as well as observed attention and anticipation in
participants who received music training. Music is temporally organized by the combination of various musical elements such as pitch, rhythm, melodic contour, harmony, dynamics, and form. The particular combination of musical elements is usually organized in a repetitive and predictable manner that is perceived as a pattern (Radocy & Boyle, 2003; Sloboda, 1985). Repetition of the predictable temporal patterns in music may increase attention and elicit anticipation for the subsequent musical patterns within the structure. Furthermore, anticipation for the temporal musical pattern may enhances the listener’s further cognitive processes such as association or recall of the musical stimuli (Radocy & Boyle, 2003).

In the present study, musical stimuli were organized by the Gestalt laws of perception which are the most common principles of pattern perception (Berger, 2002; McMullen & Saffran, 2004; Radocy & Boyle, 2003; Thaut, 2005). Each phrase or segment of a phrase of the six songs was composed with a melodic contour combined with a simple rhythmic configuration, and the phrase was perceived as a musical pattern (i.e., unit). The melodic contour was organized in its simplest form; thus, the Gestalt perceptual law of simplicity was utilized. This particular melodic contour with a rhythmic configuration served as a Gestalt or a pattern which was repeated in each song. The organization of the repeated melodies with similar rhythmic patterns in the music stimuli exemplified the Gestalt perceptual law of similarity. As a result of the repetition and similarity, the listeners might develop familiarity with the music stimuli.

In addition, the pitches of the melodies in each song were relatively close together. The step-wise pitch movement was perceived as a pattern; thus, the Gestalt perceptual law of proximity was utilized, and the pattern organization facilitated the
perception of target words. For example, two adjacent pitches associated with target words (i.e., E and D for “brown bear”) helped grouping of the words. The pitches and rhythmic figures were also organized by the law of proximity and similarity, and were matched to the syllables of each target word. The combination of musical elements and target words was then perceived as a whole unit.

All of the melodic sequences in each song moved in a typical direction of western music (in term of harmonic progression) toward completion. Therefore, the Gestalt perceptual law of common direction (i.e., good continuation) and/or completion was utilized in the organization of musical stimuli. In particular, all of the target words in the six songs were placed at the end of each phrase with a proper cadence indicating perceptual completion of the each phrase. Furthermore, the predictable harmonic progression of each cadence and the relatively longer note duration or subtle changes in timing (i.e., rubato) on the last note of each phrase emphasized the target word. The described organization of musical patterns, therefore, enhanced participants’ musical anticipation and prediction, and supported their perceptual process of the target words presented with the musical patterns.

In summary, the musical stimuli in the present study were organized by Gestalt laws of perception including simplicity, similarity, proximity, common direction (i.e., good continuation), and completion. The superior performance on speech production as well as enhanced attention and anticipation in the participants who received music training may be explained by the participants’ response to the music stimuli and their perception of well organized musical patterns. Participants who received the music
training, therefore, produced more target words that were conveyed in the musical patterns than participants who received the speech training.

Some of the unique structural characteristics in the music stimuli were not found in the speech stimuli. More principles of Gestalt pattern perception were evident in the music stimuli than in the speech stimuli. For example, the law of proximity and the law of similarity were not utilized in the speech stimuli as much as in the music stimuli. Participants who received the speech training, therefore, might not have experienced the same level of anticipation and familiarity for the speech patterns in the stories as in the musical pattern in the songs. This finding suggests that music provides more predictable temporal patterns than speech does, making it easier to perceive by children with ASD. Collectively, children with ASD respond to music stimuli attentively and perceive a good amount of information conveyed in musical stimuli which are organized by the Gestalt laws of pattern perception.

*The main effect of level of functioning on speech production in children with ASD.* Results of the present study showed that previously determined level of functioning in the participants was a significant factor in target word production. The results revealed a significant difference between participants with a low level of functioning and a high level of functioning on speech production. High functioning children achieved higher scores on the post-test and a greater change from pre to post-test scores on speech production than did low functioning children. The results also indicated that participants with a high level of functioning already knew many of the target words before the training, and improved target word production after the music and speech training. Participants with a low level of functioning, however, did not score well on the
pre-test, and scored much lower on the post-test than participants with a high level of functioning after the music and speech training. This finding suggests that a strong relationship exists between level of functioning and speech production in children with ASD. Speech production in children with ASD might be a critical factor to indicate their level of functioning.

These findings support the reliability and validity of the two tests used to evaluate the level of functioning in children with ASD for the present study: The Childhood Autism Rating Scale (CARS) (Schopler, Reichler & Renner, 1988) and the Autism Diagnostic Interview Revised (ADI-R) (Rutter, LeCouteur & Lord, 1994). Both tests evaluate children’s verbal communication skills as a critical domain indicating their overall level of functioning. The findings of the present study suggest that the level of verbal production (i.e., speech-language skills) in children with ASD can indicate their current level of functioning. Furthermore, level of functioning in children with ASD can indicate the ability to acquire new vocabulary words as well as the level of speech production including semantics, phonology, pragmatics, and prosody. The level of functioning in children with ASD can also indicate the degree of improvement on speech production after speech-language training.

*The interactive effect of training condition and level of functioning.* Results of the present study indicate a relationship between training conditions and participants’ level of functioning for subsequent speech production. According to the researcher’s observation, participants with a low level of functioning tended to lose their concentration while watching the speech video during the speech training. This observation suggests that children with a low level of functioning had difficulty focusing
on the speech video and were easily distracted by other environmental stimuli, such as other people’s presence, toys in the room, and random objects in the room. The duration of time that the low functioning children spent watching the speech video tended to be shorter than the duration of their watching the music video. When the low functioning children were watching the speech video, they tended to move their hands or legs and demonstrated behaviors indicating that they were not paying attention to the speech video. They also made more noises and required more verbal/physical prompts from the investigator to redirect their attention than the low functioning children in the music training.

When the low functioning children were watching the music video, they tended to maintain a seated position throughout the first three songs without the investigator’s verbal or physical prompts, and they did not produce any other noises. After the third or fourth song, some low functioning children also tended to lose their concentration; however, they were easier to redirect to the music video than low functioning children in the speech training. These findings suggest that the music video appeared to capture the low functioning participants’ attention and maintain it for a longer period of time than the speech video.

For example, one participant with a low level of functioning cried before every training session. As soon as the investigator played the music video, however, the child stopped crying and gave full attention to watching the video. In contrast, another participant with a low level of functioning cried before every training session and he continued crying while the speech video played. Another participant with a low level of functioning and with a high language age (language age = 3.5 years old) did not pay
attention to the speech video at all for the six training sessions. He repeatedly tried to stand up and leave his seat. When the investigator gave him verbal and physical prompts to stay seated and direct his attention, he refused to attend to the speech video by demonstrating such behaviors as screaming and kicking.

According to the researcher’s observation, participants with a high level of functioning tended to give full attention to either the music or speech video, and they also sustained their level of attention during the entire training session. Since the high functioning children had a higher level of receptive language, they appeared to understand the investigator’s verbal directions much better than low functioning children. As a result, high functioning children in both training conditions completed the training procedures without any noticeable problems. Most of the high functioning children in both training conditions did not demonstrate disruptive behaviors during the training sessions, which might be explained by their high level of comprehension of the procedures.

High functioning children in the music training also produced more expressive language while watching the music video. For example, they called out the investigator’s name, mentioned “it’s music time,” or “I like this song” and asked questions such as “what’s next?” or “is that you singing in the video?” They also tended to recall or repeat target words while watching the music video more than high functioning children in the speech training condition. After two sessions of music training, some high functioning children remembered all of the target words, and they shouted each target word while watching the music video. In summary, high functioning children demonstrated more
expressive and active communication during the music training compared to high functioning children in the speech training.

In addition, a noticeable difference emerged between children with high versus low functioning on their verbal production after watching the speech video. High functioning children produced a positive outcome (i.e., training effect) after the speech training. In contrast, low functioning children did not produce much of a change in their verbal production scores after the speech training, possibly due to poor comprehension of the procedures and a lack of concentration during the training. Interestingly, the results of the study suggest that after the music training, both high and low functioning participants produced a positive outcome (i.e., training effect), such as increased number of target words produced and improved prosody scores.

Results of the study showed that the interaction between training condition and level of functioning on speech production in children with ASD approached significance. The analysis indicated that participants with a high level of functioning improved speech production after either music or speech training. Progress on speech production achieved by high functioning participants who received the music training was slightly greater than progress achieved by high functioning participants who received the speech training. This finding indicates that the difference between the music and speech training on improvement of speech production for high functioning children with ASD was small.

The analysis further indicated that low functioning participants also made progress on speech production after either music or speech training. Progress on speech production achieved by low functioning participants who received the music training, however, was much greater than progress on speech production achieved by low
functioning participants who received the speech training. In other words, low functioning participants showed a greater improvement on their verbal production scores after the music training than the speech training. This finding suggests that music training is more effective than speech training for speech improvement in children with a low level of functioning than the speech training.

In summary, speech training might be an effective training tool to improve verbal production for children with a high level of functioning, but not for children with a low level of functioning. Music training might be an effective training tool to improve verbal production for children with either a high or a low level of functioning. In particular, children with a low level of functioning seem more likely to improve their speech production in response to music training.

*The effect of echolalia and language age on speech production.* In the present study, 64% of all participants showed echolalic behaviors. According to the results, the main effect of echolalia on participants’ speech production was not statistically significant. The correlation coefficient analysis indicated that there was a significant relationship between echolalia and level of functioning in children with ASD, and that the two variables were positively related to each other. This finding suggests that children with ASD who produce echolalia are likely to also have a high level of functioning. Conversely, children with ASD who do not produce echolalia may have a low level of functioning. Logically, if echolalia is positively related to high level of functioning, which is a critical factor for improved speech production, echolalia might be related to children’s speech production as well. The results indicate that participants with echolalia achieved a much higher score on the post-test of speech production than participants who
did not produce echolalia. In addition, participants with echolalia improved their speech production regardless of training condition or level of functioning to a much greater degree than participants who did not produce echolalia.

This finding agrees with previous studies which indicate that children with echolalia are more likely to have higher functional language development than children without echolalia (Lord & Paul, 1997; Paul & Sutherland, 2005; Prizant, et al., 2005; Prizant & Wetherby, 1993; Sundberg & Partington, 1998). Findings from previous studies and the present study suggest that echolalia serves a communicative function, and that echolalic behavior can evolve into functional speech. The findings of the present study in particular suggest that echolalic speech is related to the production of speech patterns and to acquisition of further functional language in children with ASD.

In the present study, language age and pre-test scores were used to control initial levels of language abilities in the participants. The correlation coefficient analysis indicated that language age, which ranged from 1 to 4 years, had a significant positive relationship with the pre-test score, level of functioning, and the presence of echolalia. In summary, level of functioning, language age and presence of echolalia appear to be strongly related to each other in a child with ASD. Since level of functioning was a significant factor for speech production in children with ASD, the other two factors, language age and echolalia, might also influence children’s speech production. This particular speculation leads to a discussion of the relationship between training condition, level of functioning, language age and echolalia on speech production.

According to the previous discussion, a relationship exists between training condition and high/low level of functioning. The present study found that both music and
speech training enhanced speech production for high functioning children with ASD. The present study also found that music training was more effective than speech training for low functioning children with ASD. Considering the significant relationships between level of functioning, language age, and echolalia, the collective results of the present study suggest that: (1) children with ASD who have a high level of functioning, a high language age, and echolalia could improve speech production after music or speech training; (2) children with ASD who have a low level of functioning, a low language age, and no echolalia might improve their speech production after receiving music training or speech training; however, (3) children with ASD who have a low level of functioning, a low language age, and no echolalia could improve speech production to a much greater degree after music training than after speech training.

The effect of training on individual components of speech production. The findings of the present study indicate a significant effect of both music and speech training on speech production including semantics, phonology, pragmatics, and prosody in children with ASD. The results, however, revealed no significant difference between music and speech training on the four components of speech production: semantics, phonology, pragmatics, and prosody. In the present study, both music and speech training improved all four components of speech production in just three days of training (i.e., six sessions).

This finding indicates that both music and speech training were highly effective for speech production in children with ASD. Watching either a music or speech video twice a day for three days was an intensive training procedure for children with ASD. The findings suggest that an intensive, short-term developmental speech-language
training was helpful for their speech production. Collectively, the findings indicate that children with ASD are able to learn speech-language information, including semantics, phonology, pragmatics, and prosody in just three days of training. Furthermore, through their own speech production, children with ASD are able to demonstrate what they have learned from the training. In contrast, children with ASD in no-training group did not make any noticeable changes from the pre-test to post-test.

According to the descriptive analysis, participants who received the music training improved three speech components; semantics, pragmatics, and prosody, to a greater degree than participants who received the speech training. Participants who received music training achieved a slightly higher score on the semantic component of the verbal production evaluation scale than participants who received speech training. This finding is congruent with other results indicating that a larger number of target words were produced by participants who received music training than speech training. Analysis of the raw data further revealed that two aspects of prosody, the volume (i.e., intensity), and the pitch accent, were increased to a greater degree from the pre-test to the post-test in participants who received music training compared to participants who received speech training. The other aspect of prosody, the length of vowel sounds, increased from the pre-test to the post-test to a similar degree by participants in both music and speech training conditions. For the other speech component, phonology, participants in both music and speech training achieved the same degree of improvement.

Discussion of Target Words in Songs and Stories

The researcher of the present study analyzed the 36 target words to identify types or categories of target words that children with ASD produced and generally
improved. Results of the analysis might augment the understanding of target words that children with ASD tend to produce and training effects on the target word production.

**Target word production before training.** The results of the target word analysis showed that before training, more than half of the participants already knew the target words in the categories of requesting food (i.e., “apple” and “banana”), animals (i.e., “bear” and “pig”) and objects (i.e., “key” and “shoe”). Target words in the category of calling significant people (i.e., “baby”) and words indicating familiar foods (i.e., “apple” and “banana”) were most often produced by participants before either music or speech training conditions.

Certain target words were more difficult for all participants prior to training including action words of general application (i.e., “point” and “look”), requesting another individual’s assistance (i.e., “help” and “here”), expressing recurrence (i.e., “again”), and words of attribution (i.e., “pretty”). Before training, target words of expressing agreement of affirmation (i.e., “OK” and “yes”), and requesting interaction (i.e., “come on” and “play”) were also produced by the least number of children with ASD. In particular, the target words “look,” “point,” “here,” and “again” were the common target words that were least produced by participants before music and speech training.

In summary, the findings of the present study reveal some trends in the target words produced by children with ASD. The children already knew and produced a few common functional vocabulary words prior to training. Results of the study indicate that the categories of vocabulary words produced by the children before completing the two
training conditions were similar in both training conditions. The following section will discuss the effects of training on target word production in children with ASD.

*Target word production after training.* The present study found that both music and speech training conditions increased the number of target words produced from the pre-test to the post-test in children with ASD. After music training, target words for action (i.e., “go”), names of significant people (i.e., “baby”), requesting objects and animals (i.e., “shoes,” “key,” and “pig”) were most often produced by participants. Target words indicating routine independent living activities (i.e., “eat”), common environment (i.e., “home”), and expression of cessation (i.e., “all done”) were also the most often produced after music training. Target words, “all gone,” “home,” “go,” “all done,” and “help” showed the most improvement in response to music training.

After speech training, target words for requesting food (i.e., “apple,” and “banana”), animals (i.e., “pig”), objects (i.e., “shoes”), expression of early semantic functions of non existence (i.e., “all gone”), and requesting affection (i.e., “kiss”) were the most often produced by participants. Target words “all done,” and “pretty” showed the most improvement in response to speech training.

Collectively, the analysis of the post-test results indicated that after both music and speech training, vocabulary words for requesting food, animals, familiar objects, affection, and routine independent living activities were produced most often compared to the other categories of words. In addition, target words for expressing early semantic functions of non existence or cessation (i.e., “all done” and “all gone”), or action (i.e., “go”), and words for common environment (i.e., “home”) also were also most often
produced by children with ASD as the training outcome after music and speech training conditions.

Target words “look,” “point,” “again,” and “here” were the least produced by participants after either music or speech training. This finding suggests that the categories of action words of general application (i.e., words to obtain another individual’s attention), words to request another individual’s assistance, and words expressing recurrence, were difficult to learn and produce for children with ASD even after receiving music or speech training.

According to the results of the target word analysis, vocabulary words in the category of general application (obtaining another individual’s attention), expressing early semantic functions of recurrence, and requesting assistance were produced by the least number of participants before and after either training. In addition, target words in the category of requesting interaction (i.e., “play” and “come on”), and expressing agreement of affirmation (i.e., “yes” and “OK”) were also produced by a small number of participants after music or speech training. These findings agree with previous literature indicating deficits in social behavior and lack of communicative interaction with others or environments in children with ASD (American Psychiatric Association [APA], 1994; Lord & Paul, 1997; Prizant & Wetherby, 2005). Deficits in communicative and/or interactive use of language is a defining features of ASD, and the participants in the present study showed this particular difficulty.

The present study found that the music training was effective for acquiring and producing some words to express agreement or affirmation such as “yes,” words to request interaction such as “play,” and words to request assistance such as “help.” Prizant
and Wetherby (2005) indicated that social communication requires the acquisition and production of conventional and socially appropriate language. The authors suggest that providing effective intervention to improve communication skills and interactive language is a high priority in efforts to treat children with ASD, since social communication and interaction abilities are the keys for success in daily activities for children (Prizant & Wetherby, 2005). The findings of the present study suggest that developmental speech-language training through music (DSLM) can provide an effective intervention for improving such communication skills and speech/language for children with ASD.

In summary, music training resulted in a greater number of target words being produced by the majority of participants compared to the speech training. Findings of the present study suggest that in children with ASD, music training might be more effective for learning and producing functional vocabulary words than speech training.

Discussion of Music Stimuli and Speech Stimuli

As discussed earlier, all of the stimuli used in the present study were selected or composed by the researcher considering the perceptual mechanisms in children with ASD. In particular, both music and speech stimuli were organized by the Gestalt laws of perception in order to maximize participants’ perception of the 36 target words. However, due to the intrinsic nature of the two different stimuli, the participants’ responses to music and speech stimuli were different. Thus, it is worthwhile to examine the structural characteristic of music and speech stimuli and to discuss participants’ responses to the stimuli.
Music stimuli: Songs. Some of the most often produced target words after music training were located in song # 3 (“pig,” “shoes,” “go,” and “baby”). The other two target words in song # 3 were also produced by more than 50% of the participants in the music training condition. This song was composed in A Major with an 8/12 meter, and played at a moderate tempo. Six target words were placed at the end of six phrases. The song was composed via a parallel period form, in which musical phrases were symmetrically organized. The same melodic contour was repeated more than two times, and a similar rhythmic figure was used for all six phrases in this song. The pitches in the melodic contour were close together, and the same pitches were used repeatedly. In other words, most of the Gestalt perceptual laws including simplicity, similarity, proximity, good continuation, and completion were utilized in this song.

These findings of the present study suggest that when target words are embedded in simple and repetitive combinations of musical patterns that are symmetrical and parallel in form, children with ASD may be better able to perceive and produce the words. The organization of musical patterns via the Gestalt laws might facilitate the perception and subsequent production of target words conveyed in the musical patterns.

According to the researcher’s observation, participants with a low level of functioning tended to lose their concentration while watching the music video, although it appeared to capture their attention longer and better than the speech video. In addition, the music video (9 minutes) was longer than the speech video (5 minutes and 40 seconds). During the first three songs, low functioning children tended to focus on the music video and demonstrated behaviors indicating they were paying attention. After the third or
during the fourth song in the music video, however, low functioning children tended to be distracted and lose their concentration.

The fourth song was composed in d minor and with a 6/8 meter. The song was played in a slow tempo with a soft dynamic, thus played like a sad lullaby. The form of the song was not symmetrical or parallel, but the same rhythmic figure was repeated throughout the song. The melodic contour and pitch movement were repeated infrequently, and therefore the temporal patterns were not predictable. The song included a chord change from minor to Major in the third phrase. The six target words were located at the end of each phrase; however, the length of the phrases was varied. The structural characteristics of this song might suggest that the Gestalt laws of perception were not actively utilized. As a result, the three least produced target words out of five after the music training, “look,” “point,” and “again,” came from the fourth song.

After the fourth song, some of the low functioning children appeared to recover their attention. The fifth song was composed in D Major with a 6/8 meter. This song was played in an up-beat (i.e., fast) tempo and included one rhythmic figure and a close pitch range. The melodic contour (i.e., pitch movement) for the beginning measures was repeated for each phrase. The overall structure of the song was symmetrical and parallel, and the rhythmic patterns of the song were matched to the syllables of words including the six target words. As a result, the target words “key,” and “all done” from the fifth song also were some of the most often produced target words after the music training.

This particular finding suggests that children with ASD may not pay full attention to a song composed in a minor key and played at a slow tempo. In addition, they might not perceive and produce target words that are embedded in an unpredictable
combination of musical patterns in an unsymmetrical structure. Collectively, the findings suggest that children with ASD tend to sustain their attention and to respond more favorably to up-beat songs in a Major key, than to slow songs in a minor key. Furthermore, children might perceive and produce more linguistic information such as target words conveyed through the musical patterns of a song in a Major key and a fast tempo compared to a song in a minor key and a slow tempo. The finding also suggests that the repetition of simple and predictable melodic structures and the placement of target words at the end of lyric line (i.e., phrase) facilitate the children’s perception and production of target words.

*Speech stimuli: Stories.* The six most often produced target words after the speech training were “apple,” and “all gone” (first story), “pig,” and “shoes” (third story), “kiss” (fourth story), and “banana” (sixth story). The least produced target words after the speech training were “look,” and “point” (fourth story), and “here” (fifth story). The researcher’s analysis indicated that the distribution of the produced target words after speech training was more random than after music training. Furthermore, children with a low level of functioning and a low language age tended not to focus on the speech video and were easily distracted by other stimuli in the room, regardless of the order of the stories or the contents of the stories.

Children with a high level of functioning tended to verbally repeat the target words while watching the speech video. Since the target words were located at the end of each phrase of the six stories, high functioning children tended to easily recall and produce the target words while watching the speech video. This observation suggests that the organization of the speech patterns with the Gestalt law of good continuation and
completion might facilitate the perception and prediction of target words in children with a high level of functioning.

**Limitations of the Present Study**

Several limitations of the present study need to be discussed including sample size, duration of the training video, training environment, and number of target words. First, a larger sample size may have revealed statistically significant interactions among training conditions, level of functioning, presence of echolalia, and language age. A high correlation coefficient between each independent variable and covariate suggests the possible interaction effects among the variables. In addition, increased sample size might have revealed a significant difference between responses to music and speech training. In summary, a larger number of participants might have positively influenced the statistical outcome of the present study.

Second, the varied durations (i.e., length) of the music video and the speech video may have influenced the findings of the present study. In this study, the music video was 9 minutes long, and the speech video was 5 minutes and 40 seconds long. Participant attention was identified as a critical factor for the training effects in the present study. Some participants could sustain their attention for the 5 minute-long presentation, but might have difficulty sustaining attention for 9 minutes. The length of stimuli presentation might have influenced participants’ attention and subsequent performance on target word production. Unequal length of stimuli presentation might also have hindered the accurate evaluation of participants’ responses to the stimuli.

In addition, low functioning children made greater progress after watching the music video than after the speech video, even though the music video was longer than the
speech video. If the music video was made shorter or equal in length to the speech video, the attention of the low functioning children might have been sustained for the entire duration of the music video. Consequently, the children might have made greater improvement on target word production.

The third limitation pertained to different environments in the recruiting sites that might have influenced the findings of the present study. The training sessions took place in six different facilities including two pre-schools and four learning centers for children with ASD. The pre-and post-tests and six training sessions were administered in different locations at the six different recruiting sites. On a few occasions, the researcher and participants had to use different rooms on different days in one facility. The change of environment and different arrangement in the location might have negatively influenced the children’s level of concentration during the training, and their subsequent performance on speech production. According to Jordan (1999), children with ASD tend to be abnormally upset by changes of established routine or unfamiliar surroundings. A different route to a class room or a rearrangement of furniture or other items could cause a tantrum, and the child may not readily be calmed until the familiar order or environment is restored (Jordan, 1999).

In addition, the participants’ familiarity with the researcher needs to be discussed. When the researcher took each participant to an arranged room for the pre-test, a few participants were upset and did not want to comply with an unfamiliar person. These participants needed to complete the pre-test with their teacher’s assistance. When working with children who have ASD, the researcher needs to consider familiarity with
the people (i.e., therapist, educator, or trainer) who conduct the speech-language training, and also the consistency of the environment.

Finally, the number of target words used in the Verbal Production Evaluation Scale (VPES) needs to be discussed. In the present study, 36 target words from 18 criteria for functional vocabulary words were selected from a previous study conducted by Prizant and Bailey (1992). Two target words fell under each criterion for target word selection in the present study. This particular number of vocabulary words, however, may not have been adequate for analyzing and generalizing participants’ target word production. In other words, having only two target words per each category was not enough to assess what kind of words children with ASD know and learn. In order to assess how children with ASD conceptualize the categories of words and how they might improve each category of the functioning vocabulary words over time, more target words should be put under each category of the vocabulary words. Furthermore, more target words under each category would provide a larger number of opportunities for the children to learn.

Theoretical Implications

Results of the present study indicate that children with ASD can improve their speech production after receiving a short-term (i.e., three days) speech-language intervention. The study shows that some children with ASD have an intact ability to perceive speech patterns and to produce the speech patterns that contain important linguistic information, including semantics, phonology, pragmatics, and prosody. This particular finding agrees with previous research indicating that children with ASD clearly have language impairments; however, they may also have an intact ability to perceive and
produce speech sounds, and may be able to develop some level of functional speech (Paul & Sutherland, 2005; Schuler, 1995; Tager-Flusberg, 1997).

The results also suggest that music is an effective method in developmental speech-language training for children with ASD. The present study shows that the children have an intact ability to perceive musical patterns combined with speech patterns, and to produce the speech patterns that contain the same linguistic information including semantics, phonology, pragmatics, and prosody, in a form of communicational speech language. The demonstrated intact ability of children with ASD to perceive has been reported in previous research (Heaton, Hermilton & Pring, 1999; Shuler, 1995; Tager-Flusberg, 1997; Thaut, 1999). According to these previous studies, in spite of dysfunction in auditory-cortical areas in the brain, children with ASD show an intact ability to perceive musical sounds. The present study indicates that organized musical stimuli are perceived as patterns, and which enhances perception and production of speech patterns in children with ASD. Musical pattern perception appears to be spared in children with ASD, as indicated by improvement in their speech production.

The findings of this study may also reflect changes in neural processing experienced by the participants. Previous research has shown that the perception of musical stimuli involves cortical activation of many different areas in the child’s brain including the primary and supplementary motor cortex, lateral pre-motor cortex, anterior insula, frontal operculum, primary and secondary auditory cortex, right and superior temporal pole, and basal ganglia (Brown, Martinez & Parsons, 2006). The activation of multiple brain areas might allow for integrated processing of the perceived stimuli, thus contributing to enhanced speech production.
The present study demonstrates that children with speech-language impairment due to ASD can acquire and produce functional vocabulary words by developmental speech-language training through music. These findings can be explained by the common perceptual principle and the mechanism of music and speech in children with ASD. The common perceptual principle is temporal pattern perception through auditory stimuli. In addition, the findings of the study suggest a significant integration of the two domains, music and speech, in the language acquisition of children with ASD. The perception of musical stimuli might influence speech production by activating the common mechanisms involved in both music and speech, even though music and speech stimuli have different characteristics: music is more aesthetic and abstract, while speech is more functional and concrete. Collectively, children with ASD perceive music, transform the perceived musical patterns into speech, and then produce functional speech.

Clinical Implications

The present research explores the use of music in treating children with ASD and augments the understanding of empirical mechanisms of music perception and speech production in children with ASD. The results indicate that low functioning children with ASD who experience severe difficulties in the production of speech-language can improve their speech production by participating in developmental speech-language training through music. In comparison, high functioning children with ASD may have already acquired a higher level of speech-language skills (i.e., high language age), and they may make noticeable progress in their speech production by either music or speech-language training. The study demonstrates, however, that low functioning children can make greater progress in their speech production through music training than
through speech training. A carefully designed speech-language training method using music stimuli might help to facilitate speech production in both high and low functioning children with ASD.

Furthermore, more target words were produced after music training than after speech training. In addition, the achieved improvement in speech production by the children who received the music training was greater than the improvement by the children who received the speech training. This particular finding is supported by previous research in which more spoken words paired with music (i.e., singing) were correctly imitated than the spoken words paired with rhythmic speech in children diagnosed with autism (Buday, 1995). The present study did not examine the imitation of the spoken words; however, both studies indicate the merit of the use of music in the perception and production of spoken words for children with ASD.

The present study demonstrates the effects of various strategies in developmental speech-language training through music including: (1) listening to songs that contain the functional target words, (2) viewing simplified pictures that correspond to each target word, (3) watching a singer’s mouth movements while she produces each word, and (4) verbally producing the target words (i.e., post-testing). Additional effects from music training include participants’ spontaneous singing (or speaking) while watching the music video. These interventions in the music training may have contributed to a greater improvement of speech production in children with ASD compared to the speech training. The music training consists of perception of speech stimuli embedded in musical patterns followed by production of speech. Speech training consists of perception of the speech patterns and followed by production of speech. In summary, music training is more
effective for speech production in children with ASD than speech training; although both trainings produced positive outcomes.

The present study provides an example of utilizing musical elements of songs to teach functional vocabulary words and to train adequate speech production of words for children with ASD. For music therapy application, the findings support the idea that listening to songs may be more effective than listening to stories for the acquisition and production of speech components, including semantics, phonology, pragmatics, and prosody of the target words. Six songs composed by the researcher conveyed the functional vocabulary words. In addition, the researcher utilized Gestalt laws of perception (i.e., laws of simplicity, similarity, proximity, common direction, and completion) as the compositional principles for the six songs. This particular use of songs demonstrates the technical aspects of song composition for teaching new vocabulary words and improving overall aspects of speech production in children with ASD including semantics, pragmatics, phonology, and prosody.

The previously discussed clinical application might be applied in music therapy sessions with a child with ASD who has deficits in communicative use of speech. As the initial step, the music therapist should assess the current language level of the child (i.e., language age, presence of echolalia, and level of functioning), and then select the appropriate category (and number) of functional vocabulary words and target words according to the assessment. Next, the music therapist should compose songs in Major keys and an up-beat tempo with a repeated melodic contour and rhythmic figure. The form of the song should be simple and predictable (i.e., symmetrical), and the target words should be located at the end of phrases. The live presentation of musical stimuli is
recommended, since the therapist can easily adapt the musical performance to the children’s response to the music stimuli and modify the music to facilitate the best outcome.

The present study also suggests that children with ASD are able to make associations between auditory stimuli and visual stimuli. During music and speech training, target words were presented with corresponding pictures. As a result, the children recognized the corresponding pictures, and verbally produced the correct target words during the post-test. This particular finding suggests that music therapists might use visual symbols (i.e., pictures or signs) in speech-language training for children with ASD. Speech-language training through music with consistent use of visual symbols might also promote literacy in children with ASD including awareness of phonemes and reading comprehension. In order for children to read and comprehend symbols, they have to establish associations between auditory and visual stimuli. Auditory and visual associations acquired from music or speech-language training could transfer to other functional stimuli such as books, picture schedules or daily environments, and enhance the children’s overall language skills.

Regardless of the type of training, the therapists should provide consistent and intensive speech-language training. Training sessions should take place on consecutive days, at the same time of the day (ideally more than once a day), and in the same room. A complete assessment of each child’s verbal production should be used to determine the number of days and length of the training. In addition, when the music therapist verbally introduces the target words, she or he should present the corresponding visual symbols (i.e., picture or sign). Children with ASD need to practice the association between
auditory stimuli (i.e., listening to songs) and visual stimuli (i.e., looking at pictures), in order to further enhance their language skills.

The music therapist should also encourage the children to sing along while they participate in the speech-language training through music. The music therapist could pause right before each target word to be sung, and give children opportunities to complete the phrases by singing the target word. Consistent and daily participation in developmental speech-language training through music, and repetition of the target word production could enhance speech production and language skills in children with ASD.

In summary, the present study provides a scientific foundation for the use of music in speech-language training for children with ASD. Both music and speech are aural forms of communication. In addition, music and speech share the same acoustical and auditory parameters such as frequency, rhythm, contour, intensity, waveform, timbre and cadential factor (i.e., principle of good direction and completion) (Thaut, 2005). Therefore, a carefully designed program of speech-language training through music might utilize a child’s unimpaired ability to perceive music stimuli in order to facilitate speech production.

Recommendations for Future Research

The findings of the present study suggest several ideas for future research. Replication of the study could be improved in various ways. Future studies could expand the duration of the developmental speech-language training. In the present study, children with ASD improved all four components of speech production including semantics, phonology, pragmatics, and prosody, in just three days. The children made a significant degree of progress in their speech production after receiving six sessions of either music
or speech training. Therefore, a larger number of training sessions over a longer period of
time could result in a greater outcome such as acquisition and production of more target
words for this population.

Future studies could increase the number of target words in each category of
functional vocabulary words. Evaluating verbal production with more target words in
each category will help researchers to generalize findings regarding the kinds of
vocabulary words children with ASD would easily acquire and produce. Results from
further categorical analysis of children’s target word production might facilitate the
design of effective speech-language training programs for the population.

Replications of the study could divide participants’ level of functioning into
more levels or categories such as superior, average, mildly impaired, and severely
impaired, etc. Future studies could then explore relationships between each level of
functioning and training condition, and could indicate more specific results regarding the
effects of training and level of functioning on speech production in children with ASD.

In addition, the presence of echolalia could be evaluated on a scale with more
than two levels for future studies. Echolalic behavior is considered to be transitional;
meaning the child could be transitioning from a meaningless imitation of utterances to a
communicative speech. Thus, purely echolalic speech can be developed into functional
language in children with ASD (Lord & Paul, 1997; Prizant, et al., 2005; Prizant
&Wetherby, 1993). Speech production in some children could be entirely echolalic, while
other children in the transitional period could produce both echolalic speech and
independent functional speech, and speech-language training might facilitate their
functional speech. Future studies could explain whether speech-language training
influences speech production for children with ASD in the transitional period of echolalia and facilitates acquisition of functional language. Thus, future research should include more levels of echolalic behavior, in order to thoroughly examine relationships between echolalia, training condition, and speech production.

Future research could also examine whether or not the speech-language skills acquired from training (i.e., semantics, phonology, pragmatics, and prosody) could transfer to a functional setting such as daily conversation. In the present study, children with ASD acquired new functional vocabulary words through music and speech training and verbally produced the words during the post-test in a laboratory setting. However, the ultimate purpose of speech-language training for children with ASD is the functional use of acquired speech-language in daily environments. Therefore, future studies should evaluate the children’s production of acquired target words including semantics, phonology, pragmatics and prosody, after the post-test in functional settings.

Replications of the present study could explore various combinations of musical components. The analysis of musical components in music stimuli could provide knowledge regarding what kind of combination or which way of organization of musical components (e.g., pitch movements, rhythm patterns, dynamics, or form) might influence speech production in children with ASD. Future research could also compare the length of the music training intervention. A shorter music video might positively influence children’s attention and the outcome of the study. Therefore, a thorough examination of various organizations of music stimuli and shorter lengths of intervention is recommended.
Finally, future research should encourage participants to sing along with the music video. Singing along with the presented songs represents children’s natural and spontaneous response to musical stimuli that could enhance subsequent recall and production of linguistic information. The findings of future research might provide a rationale to use music in the developmental speech-language training for children who need to improve fundamental communication skills.

**Conclusion**

The present research suggests that music training is as effective as speech training for improving speech production in children with ASD. Music and speech stimuli were perceived as temporal patterns, and thus facilitated verbal production of target words, including semantics, phonology, pragmatics, and prosody, in children with ASD through the repetition of training procedures. Music training resulted in greater improvement in speech production compared to speech training, in particular for low functioning children with ASD. The superior performance on speech production as well as enhanced attention in participants who received music training can be explained by the characteristics of the music stimuli. In the present study, music stimuli were composed using the Gestalt laws of pattern perception including the laws of simplicity, similarity, proximity, common direction, and completion. These structural characteristics in the music stimuli, however, were not found in the speech stimuli. The present study suggests that children with ASD respond to music stimuli attentively and perceive important linguistic information (i.e., target words) embedded in the music stimuli. As a result, they can acquire and produce functional vocabulary words after receiving music training. Furthermore, the present study essentially provides scientific evidence to support further
development and testing of protocols for developmental speech-language training through music. In conclusion, music is an effective tool for improving acquisition of functional vocabulary word and speech production in children with ASD.
References


Appendix A

Texts and Songs for Speech and Music Stimuli

1. Hello, hello brown bear

   What do you like to eat?
   I like to eat apples
   When I eat apples, I am happy.
   Brown bear says “I want more.”
   Daddy bear says, “the apples are all gone.”

2. In the morning, a yellow duck likes to play.

   A red bird likes to sing.
   Oh, the bird sounds so pretty.
   A green frog says “stop.”
   It’s time to go home.
   Everybody says, “I want to see mommy.”

3. Look at the pink pig.

   He is wearing shoes.
   Oh, where will you go?
   I’m going to see the baby.
   I’ll give him a big hug.
   Babies are so much fun.
4. On a starry night, a little boy was running down.

The little boy said “hey, look!

I see a butterfly. Can you point?

The butterfly will give you a kiss,

But only when you sleep.

The butterfly will come again.”

5. Hello, everyone. Let me tell you my story.

One day, I went to school.

But, I couldn’t find my key.

My teacher said “come on!

Your key is here.”

I said “Thank you, teacher.” My story is all done.

6. Hey, Mr. Monkey, can you come out?

“Hi, my friends, I want to say yes.

But, I’m busy now. I am peeling my yellow banana.”

Do you know, your hands are dirty?

Mr. Monkey says, “I know. That’s not ok.

I think I need a little help.”
Song #1

Music and Words by
Hayoung Audrey Lim

Hello, hello Brown Bear What do you like to eat?

I like to eat apples. When I eat apples, I am happy.

Brown Bear says, "I want more." Daddy bear says, "The apples are all gone."
Song #2

Music and Words by
Hayoung Audrey Lim

In the morning, a yellow duck likes to play.

A red bird likes to sing. Oh, the bird sounds so pretty.

A green frog says "Stop." It's time to go home.

Everybody says, "I want to see mom - my."
Song #3

Music and Words by
Hayoung Audrey Lim

Look at the pink pig. He's wearing shoes.  

Oh, where will you go?  I'm going to see the baby.  

I'll give him a big hug. Babies are so much fun.
Song #4

Music and Word by
Hayoung Audrey Lim

On a Star-ry night, a li-ttle boy is run-ning down.

The li-ttle boy says, "Hey, look! I see a but-ter-fly.

Can you point? The but-ter-fly will give you a kiss,

but on ly when you sleep. The but-ter-fly will come a-gain."
Hello, everyone. Let me tell you my story.

Oneday, I went to the school. But I couldn't find my key.

My teacher said, "Come on! Your key is here."

I said, "Thank you, teacher." My story is all done.
Song #6

Music and Words by
Hayoung Audrey Lim

Hey, Mister Monkey, can you come out? Hi, my friends. I want to say "Yes."

But I'm busy now. I'm peeling my yellow banana. Do you know, your hands are dirty?

Mister Monkey says, "I know. That's not okay. I think I need a little help."
Appendix B

*Verbal Production Evaluation Scale*

Instructions: Please answer each question by circling ‘Yes’ or ‘No.’ Each item answered ‘Yes’ is scored as 1 point.

Participant Code: _______________ Date: ________

Target word: __________________________________________________________________________

**Semantics**

1. Does this verbal production include the correct target word?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
</table>

**Phonology**

2. Is the target word pronounced free of any articulation or phonetic errors?

<table>
<thead>
<tr>
<th>Yes (free of error)</th>
<th>No (with error)</th>
</tr>
</thead>
</table>

**Pragmatics**

3. Is the target word produced at the right time, defined as ranging from immediate to within 10 seconds?

<table>
<thead>
<tr>
<th>Yes (produced in time)</th>
<th>No (not produced in time)</th>
</tr>
</thead>
</table>

**Prosody**

4. Is the target word produced with proper pitch accent, given the context of the cue phrase?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
</table>

5. Is the target word produced with proper length of vowel sounds given the context of the cue phrase?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
</table>

6. Is the target word produced with proper intensity (volume) given the context of the cue phrase?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
</table>

Total score: _________ points
May 3, 2007

Dear Parent,

Your child is invited to participate in a dissertation research project for speech production in children with Autism Spectrum Disorders (ASD). The project is designed to examine the effect of developmental speech-language training through music on speech production in children with ASD.

As part of this project, your child will be randomly assigned to one of three training conditions: music, speech or no training. Before and after the training, your child’s verbal production of 36 target words will be videotaped. The videotapes will be watched by two speech-language pathologists who have been trained to measure specific verbal production. Ultimately, the results should help determine whether music is an effective training method for the proper verbal production in children with ASD.

If you would like for your child to participate in this project, please sign the attached consent form and return it to your child’s teacher. If you would like to discuss this project with me in greater detail, please do not hesitate to contact me. Thank you.

Sincerely,

Hayoung Audrey Lim, MM, MT-BC
INFORMED PARENTAL CONSENT FORM

PURPOSE OF STUDY:

Your child is being asked to participate in a research study on music and speech production in children with Autism Spectrum Disorders (ASD). The purpose of this study is to explore how the perception and production of musical patterns in singing will impact the perception and production of speech and language in children with ASD. The study will examine the effect of music as part of developmental speech-language training in the speech production of children with ASD.

PROCEDURES:

The investigator will visit the school where your child is currently attending. At your child’s convenience and daily schedule of the school program, the investigator will arrange to meet your child at a room at the school for all subsequent research procedures. Your completed consent form will give permission to the investigator, Hayoung Audrey Lim, (1) to use your child’s standardized language test score and (2) to access your child’s score on autism rating scale and (3) to ask your child’s speech/language pathologists or teachers about his/her use of echolalic behaviors. All screening procedures for your child will be completed by the investigator.

Your child will be randomly assigned, like the flip of a coin, to one of three training conditions: music, speech or no training. He/she will be tested individually. On the first day of the experiment, the investigator will administer a pre-test to your child, and two practice trials will be given to the child prior to the pre-test. The pre-test will use 36 target words that are functional vocabularies. The pre-test will be a form of fill-in-the-blank, intra-verbal communication. The pre-test will include 36 phrases that are structured so as to end with one particular target word. The investigator will say the first part of each phrase, and then leave off the target word. A corresponding picture will also be presented with each target word. The objective of the procedure in the pre-test will be to prompt your child to produce a target word at the end of a phrase upon hearing the first part of the phrase. The pre-test will take approximately 10 minutes.

On the next day, the investigator will meet the participant and show the appropriate video. If your child is assigned to the music condition, he/she will watch the music video. If your child is assigned to the speech condition, he/she will watch the speech video. Your child will watch the same video presentation two times per day (i.e., morning and afternoon) for three consecutive days. Video watching will take approximately 15 minutes. If your child is assigned to the control group, he/she will not watch the music nor speech video for three consecutive days.
Food (e.g., crackers or cookies) and tangible prizes (e.g., rubber balls, or yo-yos) will be used to keep your child engaged during both the music and speech conditions. If the child cannot have any food due to dietary restraints, stickers will be offered. Your child will receive food after each song or story if he/she stays in the room and refrain from behaviors that interfere with his/her perception of the information delivered in the video (e.g., shouting, crying, and tantruming). If your child earns a reinforcer for each song or story, then he/she will receive a tangible prize to keep at the end of each training session.

On the next day after watching the training video two times per day for three days, the investigator will administer the post-test. The post-test is identical to the pre-test. Each condition will be completed in a minimum of 5 days and a maximum of 9 days in order to allow for changes in your child’s schedule due to absence, illness, etc.

**Your child’s verbal production during the pre-and post-tests will be video-taped.** Two speech/language pathologists who specialize in treating young children with language impairment and who are blind to the purpose of the study will be trained for data coding. The speech/language pathologists will watch each video tape one time and will evaluate your child’s verbal production. The results of the study for your child will be shared with you.

**RISKS:**
No foreseeable risks or discomforts are anticipated for your child by participating in this study.

**BENEFITS:**
No benefit can be promised to your child by participating in this study.

**ALTERNATIVES:**
You have the alternative for your child to not participate in this study. During the study, you may stop your child’s participation at any time. You and your child have the right to refuse to participate in the study and nothing will happen to you and your child as a result. Your child’s care at their school will not be affected.

**COSTS:**
No costs are anticipated for you as a result of participating in this study.

**PAYMENT TO PARTICIPANT:**
No monetary payment will be given for participation in this study.

**CONFIDENTIALITY:**
The researcher will consider your child’s records confidential to the extent permitted by law. The U.S. Department of Health and Human Services (DHHS) may request to review and obtain copies of these records. Your child’s records may also be reviewed for audit purposes by authorized University of Miami employees or other agents who will be bound by the same provisions of confidentiality.
The investigator, Hayoung Lim has authorized access to your child’s records of language age and level of functioning by the agreement with your child’s school. These records will be kept in electronic format (e.g., data base, PDAs, etc) and paper format. The collected data will not hold any information that will identify your child. Data will be closed by special number rather than by your child’s name. Hard copies or discs containing your child’s records will be secured in a locked file cabinet to which the investigator, Hayoung Lim, will have the only access.

Your child’s speech during the pre-and post-tests will be video-taped. The video tape will be kept by the investigator, Hayoung Lim until the research project is completed.

RIGHT TO WITHDRAW:
Your agreement for your child’s participation is voluntary; you and your child have the right to withdraw from this study. If you do not want for your child to participate or do not follows the procedures, the researcher can also remove you from the study without your consent.

OTHER PERTINENT INFORMATION:
Dr. Shannon de l’Etoile (305-284-3943) will gladly answer any questions you may have concerning the purpose, procedures, and outcome of this project. If you have any question concerning the research study or your child’s participation in this study, please contact Hayoung Audrey Lim, MM, MT-BC at (305) 479-1351. If you have any question about your child’s right as a research participant in this research, you can contact University of Miami, Human Subjects Research Office at (305) 243-3195.

PARTICIPANT AGREEMENT:
I have read the information in this consent form and agree to allow my child to participate in this study. I have had the chance to ask any questions I have about this study, and they have been answered for me. I am entitled to a copy of this form after it has been read and signed.

Name of Your Child

Name of Parent(s) or Caregiver

Signature of Parent(s) or Primary Caregiver Date

Signature of person obtaining consent Date
Videotaping

By signing this section you give consent for your child to be video taped during this study.

______________________________________                       ______________________
Signature of Parent or Primary Caregiver                                                      Date

_______________________________________                     ______________________
Signature of person obtaining informed consent                                             Date