Assessing Prosody in Minimally to Nonverbal Children with Autism

Jill Thorson[1], Steven Meyer[2], Daniela Plesa-Skwerer[2], Rupal Patel[1] and Helen Tager-Flusberg[2]

[1] Northeastern University, USA
[2] Center for Autism Research Excellence, Boston University, USA
j.thorson@neu.edu

Abstract

A procedure for assessing basic prosodic perception and production abilities of minimally to nonverbal children and adolescents with autism spectrum disorder is described (AP: Assessment of Prosody). The procedure consists of three sections: an optional primer phase, a learning phase, and an assessment phase. It includes the assessment of both the perception of basic pitch accent structure distinctions (low versus high) as well as elicits expressive productions of these contrasts. The goal of the procedure is to evaluate the extent to which this population can perceive and produce prosodic distinctions. The overarching aim is to create a pre and post assessment to quantify prosodic competence and performance of minimally to nonverbal children and adolescents who are eligible for music-motor based intervention therapies (i.e. AMMT: Auditory Motor Mapping Therapy). Current and future versions of the assessment are discussed.

Index Terms: prosody, assessment, autism spectrum disorder

1. Introduction

Expressing and understanding prosodic variation is essential for successful human interaction. Individuals with autism spectrum disorder (ASD) show deficits in communication and pragmatic use of language ([1]), with mixed results for how stress, intonation, and phrasing distinctions are employed ([2], [3], [4]). Two characteristic differences of autism are atypical development in reciprocal social interaction as well as atypical communication ([5]), both likely impacted by differences in their prosodic usage. There are few prosodic assessments currently available, and none specifically designed to be used with minimally to nonverbal children with ASD. The goal of this research is to develop and implement a new prosodic assessment in order to evaluate the perception and production of basic prosodic distinctions in minimally to nonverbal children with ASD.

The primary motivation for this work is to create an assessment that can be administered before and after music-motor intervention therapies, which aim to increase verbal ability in minimally to nonverbal individuals with ASD (AMMT: Auditory Motor Mapping Therapy, [6]). It is vital to have a way to capture the prosodic abilities of this population and to design a procedure that can assess both expressive and receptive capacities. Currently, there exists a gap in the prosodic assessments available for this population, with no current type of controlled evaluation.

The number of prosodic assessments to date is limited. The primary procedure in use by clinicians is the Profiling Elements of Prosodic Systems, PEPS, and the PEPS-C (Child version) ([7], [8]). The procedure was designed provide a comprehensive assessment of prosody and intonation and to assess prosody in four communication areas: interaction, affect, boundary (chunking), and focus. The PEPS-C was originally normed on 120 southern British English typically-developing children aged 5 to 14 years old ([9]) as well as 18 children with autism, speech impairment, pragmatic language impairment, SLI, hearing impairment, and stammering ([10], [11]). Previous procedures include the Prosody Voice Screening Profile ([12]) and PROP ([13]), but neither includes normative data like the PEPS and PEPS-C. Critically, while the PEPS-C has been implemented for verbal children with ASD, it is not an appropriate assessment for children and adolescents who are minimally or nonverbal.

One of the primary challenges regarding a minimally or nonverbal population with ASD is that there are often intellectual and behavioral differences present as well ([14]). Therefore, the same types of assessments that are used with typically developing children are not appropriate. Our procedure is inspired by Prosodic Marionette, a visuo-spatial task that allows users to express prosodic knowledge via the physical manipulation of word-icon blocks on a tablet screen ([15]). For this population, we created a simplified visuo-spatial manipulation task and moved it into the physical domain (versus a digital one).

The current procedure focuses on both the perception and production of basic prosodic distinctions. It is primarily concerned with 'form' over 'function' level processing. That is, the task assesses the auditory discrimination of prosody, the ability to represent and communicate this discrimination visually, and the voice skills required to create these differences in speech. Future versions of the task will include a 'function' level and assess the communicative and interactive intent of prosody as well. Also, while the current version is created for minimally to nonverbal children and adolescents with ASD, the task has the potential to be adapted for adults and other populations that would benefit from a simplified task design.

2. Method

The current version of the Assessment of Prosody (AP) is in its sixth iteration, with improvements made after piloting each version of the materials and procedure. This section describes the materials required to run the protocol, how the stimuli were created and selected, information about the participants who piloted the protocol, and a general description of the procedure (described in detail in Section 3).

2.1. Materials

Due to complexities of using a digital interface with this population, the protocol is administered in the physical

modality with hands-on objects. Examples of difficulties that emerge from using a digital interface include the inability to disengage with a tablet- or computer-based assessment and decreased attention to the researcher during administration. Attention was also a consideration when determining the length of the procedure, with an effort made to keep the sections effective and brief.

The materials used in this procedure are:

- Computer with speakers, equipped with stimuli presentation software (e.g., PowerPoint).
- White magnetic board (11x8.5 inches).
- Two square magnets (2x2 inches). Selected due to maneuverability (over Velcro or no attachment).

2.2. Stimuli

Careful consideration was taken to select the stimuli for the assessment (see Table 1 for a full list of all stimuli used during the procedure). Target patterns included *low-high* and *high-low* constructions, where a low was equivalent to a deaccented syllable or word, and the high was a focused element carrying an H* pitch accent. A female speaker who was trained in prosodic analysis and who was not an Administrator in the task recorded the stimuli. A trained Administrator uttered all live versions of the stimuli in the same manner. Criteria for the stimuli included:

- Familiar and appropriate syllables/words for participants of all ages and with varying ASD diagnoses. Selected from a list of commonly known/produced words by this population.
- Use of primarily sonorant sounds as to provide a more stable track of the fundamental frequency during analysis of vocal productions.
- Open and closed monosyllables with no more than two-consonant onsets or codas: CV, VC, CVC, CCVC, and CVCC.

Table 1. *Full list of individual stimuli used in the procedure.*

Sound/Syllable/Word	Syllable Structure
<cow moo>	*<sound>*
<bird chirp>	*<sound>*
ma	CV
di	CV
red	CVC
car	CVC
big	CVC
dog	CVC
lions	CVCC
sleep	CCVC
dogs	CVCC
eat	VC

2.3. Participants

The protocol has been piloted on ten participants ranging in age from 6 to 13 years old. All participants were minimally verbal or nonverbal and had a diagnosis of autism spectrum disorder. The participants are part of a larger study looking at language, cognitive, and neural abilities and differences in this population.

2.4. Procedure

The procedure consists of three primary sections. The first two sections orient the participant with the task and the structure of the assessment. The final section is the assessment portion (Figure 1). The procedure is administered in a quiet room with two Administrators (1 and 2). The primary task of the assessment asks the participant to physically manipulate magnets on the magnetic board to visually recreate auditory stimuli of pitch accent structures. Figure 2 shows an example magnet configuration demonstrating a *low-high* sequence. Auditory stimuli is either said orally by Administrator 1 or played from a pre-recorded presentation by Administrator 2. Administrator 1 interacts with the participant and aids them in the task. Task orientation and learning was based on a Task Analysis design where there is decreasing support from the Administrator during learning ([16]). Details of the administrator guidance levels are outlined in Section 3.
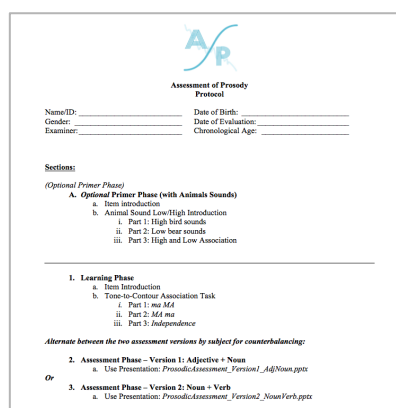

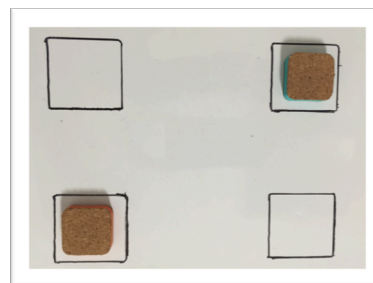Figure 1: *Prosodic Assessment Protocol, Version 6.*

Figure 2: *Magnetic board with magnets in low-high configuration pattern.*

3. Procedure

There are three primary sections to the AP protocol. The first is an Optional Primer Phase, the second is the Learning Phase, and the third is the Assessment Phase. Each phase and their components are described in more detail in the following sections. See Figure 3 for a detailed list of trials, stimuli, and administrator guidance level. Finally, scoring of the protocol is discussed in the final section.

3.1. Primer Phase

The Primer Phase was added after the first five rounds of piloting the protocol. It was included after experiencing difficulty engaging the task with younger participants. This phase introduces the concept of low and high via animal

sounds first before human speech and is optionally available as a starting point for the procedure. The high location is represented by a bird with corresponding bird chirp sounds. The low location is represented by a cow with a corresponding cow moo sound. Cartoon birds are present in each of the top two squares on the magnetic board, and cartoon cows are present in the two bottom squares. The items are first introduced and then the animal sound low/high association task begins with a series of eight trials.

Primer Phase			
Trial	Stimuli	Pattern	Guidance/Criterion
1	1 bird chirp	high	Tap/Repeat 2x
2	2 bird chirps	high-high	Tap Hand/Repeat 2x
3	2 bird chirps	high-high	Fade away/Repeat 2x
4	1 cow moo	low	Tap/Repeat 2x
5	2 cow moos	low-low	Tap Hand/Repeat 2x
6	2 cow moos	low-low	Fade away/Repeat 2x
7	chirp-moo	high-low	Fade away/Repeat 2x
8	moo-chirp	low-high	Fade away/Repeat 2x
Learning Phase			
Part 1 – Low-High			
1	ma MA	low-high	Tap/Repeat 2x
2	ma MA	low-high	Tap Hand/Repeat 2x
3	ma MA	low-high	Guide Hand/Repeat 2x
4	ma MA	low-high	Fade away/2 independent
Part 2 – High-Low			
5	MA ma	high-low	Tap/Repeat 2x
6	MA ma	high-low	Tap Hand/Repeat 2x
7	MA ma	high-low	Guide Hand/Repeat 2x
8	MA ma	high-low	Fade away/2 independent
Part 3 – Independence			
9	ma MA	low-high	Fade away/Repeat 2x
10	MA ma	high-low	Fade away/Repeat 2x
11	ma MA	low-high	Fade away/Repeat 2x
12	MA ma	high-low	2 in row independently
13	DI di	high-low	2 in row independently
14	di DI	low-high	2 in row independently
Assessment Phase			
Syllable Phase			
1	MA ma	high-low	Tap/Repeat 2x
2	di DI	low-high	Tap Hand/Repeat 2x
3	MA di	high-low	Guide Hand/Repeat 2x
4	ma DI	low-high	Fade away/2 independent
Utterance Phase			
5	red CAR	low-high	2 in row independently
6	BIG dog	high-low	2 in row independently
7	RED car	high-low	2 in row independently
8	big DOG	low-high	2 in row independently
9	red CAR	low-high	2 in row independently
10	big DOG	low-high	2 in row independently
11	RED car	high-low	2 in row independently
12	BIG dog	high-low	2 in row independently

Figure 3: *Full list of stimuli, configuration pattern, and administrator guidance level for each trial in each phase. Assessment phase is Version 1 with Adjective plus Noun stimuli. Version 2 is not shown and uses Noun plus Verb stimuli.*

Trials 1 to 3 introduce the *high-bird chirp* association. For Trial 1, one magnetic piece is put into the high position before playing the sound, and Administrator 1 guides the participant's hand to tap the magnet as the sound plays (repeated two times). For Trial 2, both magnets are placed into the high positions while playing two repetitions of the bird chirps. The participant's hand is guided to tap each magnet while the sound plays (repeated two times). Trial 3 starts with the two magnets in a neutral position (halfway between the high and low squares) and Administrator guidance begins to fade away as to let the participant move the pieces independently (repeat two times). Trials 4 to 6 introduce the *low-cow* sound association, and proceed in the same manner as the first three trials. Trials 7 and 8 combine the two sounds together with first a *bird* then *cow* sound for a high-low sequence (and target configuration). Trial 8 plays a *cow* then *bird* sound for a low-high sequence. The magnets begin in a neutral position for both of these trials and Administrator 1 is instructed to fade away from aiding the participant, and allow them move the pieces independently if possible. For each trial, Administrator 1 only says, "*Look. Listen.*" to the participant.

3.2. Learning Phase

The goal of the learning phase is to orient the participant with the task (if they did not complete the Optional Primer Phase) and demonstrate the associations for how the physical objects relate to the melodic contours. The Learning Phase first identifies *low-high* and *high-low* contrasts using open CV syllables (*ma* and *di*). There are three parts to the Learning Phase: *Part 1 Low-High* (*ma MA*), *Part 2 High-Low* (*MA ma*), and *Part 3 Independence* (mixing *ma-ma* and *di-di sequences* in varying *high-low* and *low-high* patterns).

The first two parts have four trials each and consist of a decreasing amount of Administrator guidance. Guidance begins by tapping the pieces in their target configuration while the corresponding stimuli plays. The magnets are then moved back to a neutral position, the sound is played and Administrator 1 says, "*Listen*" and guides the participant's hand as the pieces move into place, saying only "*Look*". For the final trial in each of the first two parts, the Administrator fades away from guiding the participant and just says "*Look, Listen*". Each trial is repeated two times until the final trial, where the criterion for advancing is the completion of two iterations performed in a row independently by the participant. The stimuli for the first trial in Parts 1 and 2 are said aloud by the trained Administrator 1. For the rest of the trials, Administrator 2 plays the stimuli from a recording on the presentation software in a wizard-of-oz type setting, where the participant hears the stimuli over speakers as they perform the task.

The third part *Independence* consists of six trials, alternating between *low-high* and *high-low* target trials. The first four trials consist of *ma-ma* stimuli, while the final two trials transition to a new syllable combination (*di-di*, with the same *low-high* and *high-low* pitch accent structures). The first three trials are repeated two times each, while the final four trials must meet a criterion of two repetitions performed independently by the participant. After each trial, Administrator 1 moves the pieces back to a neutral position and says, "*Can you move the pieces? Listen. Can you say the words too?*" The target sound is played from a pre-recorded stimuli presentation. The participant is asked to move the pieces into either a *low-high* or *high-low* sequence and to reproduce the sequence verbally if possible (verbal imitation is not always feasible with all participants).

3.3. Assessment Phase

The goal of the Assessment Phase is to evaluate if a participant can reconstruct *low-high* and *high-low* contrasts with limited or no guidance by the Administrator to express receptive knowledge of pitch accent/focus structure. This phase begins with a Syllable Phase then extends into a two-word combination Utterance Phase using lexical items. The Administrator 1 says "*Listen*", Administrator 2 plays the corresponding pre-recorded stimuli, Administrator 1 continues with "*Can you move the pieces now?*", and then prompts the repetition of the utterance by saying "*Can you repeat the words? Can you say it too?*"

There are two versions of the assessment phase. A future protocol may collapse the two versions, but currently they are being piloted separately to ensure the assessment is not overly complex and does not last more than thirty minutes. Version 1 consists of lexical items that create adjective plus noun combinations (Adj+N). The two stimuli pairs are "*big dog*" and "*red car*", with utterance focus/prominence randomized between the adjective and the noun by trial. Version 2 consists of lexical items that create noun plus verb combinations (N+V). The stimuli pairs are "*lions sleep*" and "*dogs eat*", with focus randomized between the noun and the verb by trial.

Both versions of the assessment begin with the Syllable Phase consisting of four trials that use *ma-ma*, *di-di*, and *ma-di* sequences. As with Part 3 of the Learning Phase, there is limited to no guidance by the Administrator. The Utterance Phase has eight trials, where each stimulus item is tested two times each with no direct Administrator guidance beyond the verbal instructions (Adj+N Version 1: *RED car, red CAR, BIG dog, big DOG;* N+V Version 2*: LIONS sleep, lions SLEEP, DOGS eat, dogs EAT*). The two versions are counter balanced and all stimuli are pre-recorded sound files. Administrator 1 is directed to say only "*Can you move the pieces? Listen. Can you say the words too?*" for each trial. The criterion for advancing to the next trial is for the participant to perform two iterations in a row independently.

3.4. Scoring

Scoring of the participant's performance is completed off-line from a video recording after the protocol is administered. Trials are analyzed for both receptive and expressive abilities. For receptive skills, the level of guidance required, ability to perform the task independently, and correctness of the target sound and target configuration pattern are scored. Level of guidance is noted for each trial in the Learning and Assessment Phases, but scoring is only completed for the Assessment Phase.

The scoring for the receptive abilities in the Assessment Phase is done by providing each trial with 1, 2 or 3 points. A 1 is given if the participant completes the trial correctly and with no administrator support (i.e. correctly recreates the target configuration). A 2 is given if the participant completes the trial correctly and with a very limited amount of administrator prompting or guidance (e.g., guiding participant's hand back to the magnetic board, but not moving the pieces into the target positions). A 3 is given if the participant cannot complete the trial or does not recreate the correct target configuration. If there are three trials in a row during the Assessment Phase where the participant cannot complete the trial after two attempts, then the protocol ends at this point. Thus, a score of 12 total points (1 point for each trial) is given when the participant successfully recreates the targets for each trial. A score of 36 is given when the participant completes the task, but is unable to successfully recreate the target configurations. If the Assessment Phase is ended before the final trial, scoring is adjusted to reflect the number of trials completed. For expressive skills, all speech is recorded and analyzed for its prosodic structure in relation to the target sound. The scoring for expressive abilities is still being developed and adjusted.

4. Discussion and Future Directions

The Assessment of Prosody protocol was designed to assess the basic prosodic abilities of minimally to nonverbal children and adolescents with ASD. Both receptive and expressive abilities were captured, with a focus on receptive skills due to the limited expressive language abilities of this population. An effort was made to create a protocol that successfully tests receptive prosodic abilities while only requiring a limited amount of learning to complete the task due to differences in attention, language, and cognitive abilities.

There are several possible modifications to the assessment that are currently being considered. First, a future version may collapse the two current versions (Adj+N and N+V) in order to capture greater syntactic variability in one protocol. Second, we are still experimenting with creating a more streamlined learning phase in an effort to balance over exposure to the task and the time necessary to effectively learn the task before the assessment phase. Finally, the Optional Primer Phase may be eliminated if it proves not to help participants become more engaged in the basic task design (to be determined after additional piloting). Overall, this assessment makes an important first step in designing a task targeted for minimally to nonverbal children and adolescents with ASD. This procedure tests the prosodic skills that are necessary for music-motor based intervention therapies and provides a baseline for comparing how these types of therapies impact prosodic components of the language system.

5. Acknowledgements

6. References

[1] Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior* (Vol. I, pp. 335-364). Hoboken, NJ: Wiley.

[2] Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35,* 205-220.

[3] McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language and Communication Disorders, 42,* 682-702.

[4] Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research, 44,* 1097-1115.

[5] Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*(1), 11-29.

[6] Wan, C. Y., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., Norton, A., & Schlaug, G. (2011). Auditory-Motor Mapping Training as an Intervention to Facilitate Speech Output in Non-Verbal Children with Autism: A Proof of Concept Study. PLoS ONE, 6(9).

[7] Peppé, S. J. E. (1998). *Investigating linguistic prosodic ability in adult speakers of English.* Unpublished PhD thesis, University College London.

[8] Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics and Phonetics, 7*(4-5), 345-354.

[9] Wells, B., Peppé, S., & Goulandris, A. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*(4), 749-778.

[10] McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders, 38*(4), 325-350.

[11] Wells, B., & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech, Language, and Hearing Research, 46*(1), 5-20.

[12] Shriberg, L., Kwiatkowski, J., & Rasmussen, C. (1990). *Prosody-Voice Screening Profile.* Tucson, AZL Communication Skill Builders.

[13] Crystal, D. (1982). *Profiling Linguistic Disability* (pp. 114-138). London: Edward Arnold.

[14] Boucher, J., Bigham, S., Mayes, A., & Muskett, T. (2008). Recognition and language in low functioning autism. *Journal of Autism and Developmental Disorders, 38*, 1259-1269.

[15] Patel, R., Brumberg, J., Shattuck-Hufnagel, S., Veilleux, N., and Usher, N. (2014). The Developmental Trajectory of Linguistic Prosody. Presented at *Motor Speech Conference.* Sarasota, FL.

[16] Szidon, K., & Franzone, E. (2009). *Task Analysis.* Madison, WI: National Professional Development Center on Autism Spectrum Disorders, Waisman Center, University of Wisconsin.