Clinical Focus

A Case Study on Vocal Loudness With a Young Adult With Autism Spectrum Disorder and Developmental Delay

Amanda Pack, a Paola Colozzo, b B. May Bernhardt, b Bosko Radanov, b Riley Rosebush, b and Stefka H. Marinova-Todd b

Purpose: This clinical focus article describes an exploratory case study addressing reduction of vocal loudness in a young adult with a history of autism spectrum disorder and developmental delay. The need for a short-term pullout individual intervention arose from his participation in the Advancing Language and Literacy group, a program that provides support for enhancement of speech, language, and literacy skills for young adults with developmental disabilities.

Method: The participant attended individual treatment sessions weekly for 9 sessions, all of which were digitally audio-recorded. Client awareness of different loudness levels was first established using nonspeech sounds and speech samples. Loudness in spontaneous speech was monitored with support of verbal and visual feedback in the individual sessions and tracked in individual sessions as well as sessions of the Advancing Language and Literacy group, which served as a generalization context.

Results: In the individual sessions, the participant’s spontaneous production of acceptable volume improved from 42% to 92% of utterances. Observation of group participation indicated generalization, with the majority of utterances produced posttreatment at an acceptable volume.

Conclusions: This clinical focus article emphasizes suprasegmental aspects as a dimension of communicative competence and highlights the need for continued intervention research and services for young adults with developmental disabilities.

Young adults with a history of developmental delay generally receive minimal speech, language, and/or literacy services (Freeman, 2003; Rydzewska, 2012; Wright, Brooks, D’Astous, & Grandin, 2013). A group program conducted in the speech-language pathology department at the University of British Columbia offers such services, while simultaneously presenting opportunities for student training, case studies, and research. The Advancing Language and Literacy (ALL) group has existed since 2004 and has resided at the University of British Columbia since 2009. Participants are 10 to 12 young adults with a history of developmental delay and associated speech, language, and literacy difficulties of various types and levels. This clinical focus article describes a short-term pullout intervention for one of the group participants. Like several others in the group, this particular individual showed social-communication challenges characteristic of autism spectrum disorder (ASD) and developmental delay. Following the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association [APA], 2013), ASD is defined as a lifelong neurodevelopmental condition characterized by persistent deficits in social communication and social interaction and restricted, repetitive patterns of behavior, interests, or activities. Although ASD is typically first observed and diagnosed in childhood, many adults with ASD continue to face significant challenges in socialization, communication, and emotional regulation (Wright et al., 2013). One suprasegmental aspect of the participant’s communication was negatively affecting his communication and participation in the group—excessive vocal loudness. A short-term block of individual therapy was provided as a supplement to the ongoing ALL group program with the goal of helping him learn to moderate his vocal levels. The group also served as a generalization context.

The objectives of this clinical focus article are twofold: (a) to emphasize suprasegmental aspects as a dimension of communicative competence and (b) to set the stage for further treatment research in this area and with similar...
clients. The next section provides background on social-communicative competence pertaining to suprasegmental factors such as vocal loudness.

**Social-Communicative Competence and Suprasegmental Aspects of Speech**

Social-communicative competence is a multidimensional construct that indicates the degree of success (competence) of a person’s verbal and nonverbal communication for social purposes (Adams, 2005; APA, 2013). Suprasegmental aspects of speech production make up one dimension that contributes to social-communicative competence (Scharfstein & Beidel, 2015). Aspects of prosody (i.e., phrasing, stress, rate) and voice (i.e., loudness, pitch, voice quality, resonance quality) serve grammatical, pragmatic, and affective functions (Shriberg et al., 2001). These speech characteristics also influence listeners’ perceptions of a speaker’s competence, personality, and level of ability, which may in turn either facilitate or hinder social and vocational integration (Paul et al., 2005; Shriberg et al., 2001; Shriberg & Widder, 1990). Shriberg et al. (2001) proposed that even infrequent differences in prosody or voice “can affect listeners’ perceptions of a speaker’s emotional status and attractiveness” (p. 1110). Furthermore, there is evidence that listeners make consistent—but not necessarily accurate—judgments regarding the personal characteristics of a speaker on the basis of nonverbal indices including pitch, loudness, and speaking rate (Mohammadi, Vinciarelli, & Mortillaro, 2010).

Vocal loudness was the suprasegmental focus of interest for this clinical focus article. Literature is sparse on this topic for both children and adults. Nonetheless, some studies suggest that listeners dislike or attribute negative traits to loud voices. In a study with typical adults, Page and Balloun (1978) examined reactions of 63 college students to a female interviewee answering questions at low, moderate, or high vocal volume. When louder volume was used, listeners rated the speaker’s desirability as a communication partner lower and her aggression and dominance higher than at the moderate volume. In another study, Hollien, Gelfer, and Carlson (1991) asked 80 typical and healthy adults in four age groups (20 to 90 years) to listen to recordings of speakers reading a standard passage under low, medium, and high intensity conditions. All four participant groups exhibited a strong and consistent preference for medium intensity, and in fact, the intensity variable accounted for the largest variance in observer ratings of voice volume did not significantly differ between the groups. Ratings and digital analyses were nonetheless in agreement for another group of participants (children with social phobia) who showed not only lower vocal intensity but also higher and more variable pitch.

In another relevant study, Shriberg and Widder (1990) documented the prosody and voice characteristics from speech samples of 20 men with intellectual disability aged between 20 and 55 years. The percentages of utterances with acceptable loudness were 92% and 83% for the groups with mild and moderate intellectual disability, respectively, although the ranges reflected high levels of individual variability.

Most studies have used noninstrumental auditory-perceptual procedures to obtain listener judgments of loudness. The use of such listener ratings is supported by research indicating that listeners are sensitive to changes in intensity and effort in their judgments of loudness (Brandt, Ruder, & Shipp, 1969). Taken together, the data from these studies suggest that individuals with ASD with accompanying intellectual disability may use excessively quiet or excessively loud levels of voice in conversation, which can then further interfere with their ability to communicate successfully with others.

The intervention research focusing on loudness is sparse. One study used a Social Story (http://www.thegraycenter.org) protocol to successfully decrease the frequency of excessively loud voice in the classroom for a 9-year-old boy with autism (Ozdemir, 2008). To our knowledge, no previous study has targeted vocal loudness in adults with ASD. The current case study begins to address this gap. The treatment program aimed to increase the participant’s awareness and self-regulation of acceptable conversational loudness and thereby increase the proportion of utterances produced at acceptable loudness levels.

**Method**

**Participant**

The participant was a young adult in his mid-20s with ASD, normal hearing, and challenges in adaptive functioning.
The participant and his family provided consent to participate in the project in accordance with procedures approved by the University of British Columbia’s Behavioural Research Ethics Board. He used speech augmented with written language (i.e., spelling words aloud or writing messages) to communicate. His utterances were generally short and simple but not always completely intelligible. Overall, he presented with moderately severe communication difficulties and associated anxiety, evidenced by occasional sudden outbursts with loud voice and agitated expression.

**General Procedure**

Nine individual treatment sessions (30 min each) were conducted at the university by a student clinician (the first author) and faculty supervisor (the third author) as pullout sessions from the ALL group. The participant was audio-recorded during the individual weekly treatment sessions using the recorder’s built-in microphone. On three occasions (Weeks 3, 4, and 5), another ALL group member participated in the individual sessions in order to provide peer modeling and promote generalization once the participant had shown that he understood the purpose and structure of the individual sessions. The regular ALL group sessions served primarily as a context in which to document whether any generalization occurred from the individual sessions to the group context. Only minimal attention was drawn to the participant’s vocal loudness levels in the group context (see later). However, two of the authors were present in the ALL group itself at all times, maintaining potential stimulus control outside of the individual sessions.

The student clinician took precautions to reduce the participant’s anxiety during the individual therapy sessions. First, she ensured that the participant notified his caregivers of his arrival by phone prior to starting the sessions. Second, she and the participant completed a written Plan of the Day for the session activities, a familiar routine in the ALL group.

**Vocal Loudness Assessment and Treatment Goals**

A 15-min language sample recorded 7 months prior to the study was used to derive a prebaseline score for vocal loudness. The participant wore a lapel microphone while the student clinician used an interview format to encourage him to talk about favorite, recent, and upcoming activities (e.g., watching a hockey game, going to a birthday party, walking around downtown, holiday decorations and events, etc.). Because the sample data were not current, they could not be directly compared with the treatment-session data. However, in that earlier sample, only 14% of the participant’s utterances were rated as having acceptable conversational loudness (see rating procedure later), and thus the sample provides support for the concern regarding vocal loudness as a problematic area for him. Furthermore, the ALL group leaders agreed that the participant’s vocal use remained unchanged in the group setting over the intervening period. His excessively loud voice had been a long-standing concern of theirs, and it continued to disrupt group activities and irritate some of the other group members.

The participant’s goals were the following:

1. Identify sounds and speech produced as quiet, medium, or loud.
2. Use an acceptable range of conversational loudness without prompting in the majority of his utterances.

**Individual Treatment Sessions**

The individual sessions included activities and strategies that were meant to help the participant develop awareness regarding the relative loudness levels of auditory stimuli and also support self-recognition and adjustment of his vocal volume. Figure 1 presents an overview of the structure of the treatment sessions.

**Identification Trials**

During the identification trials, the participant listened to live sound stimuli or audio recordings and identified them as quiet, medium, or loud. The stimuli included nonspeech sounds (e.g., clapping, knocking, chords played on a keyboard) presented at different intensities (e.g., quiet, medium, and loud knocking) and previously recorded speech stimuli produced by the student clinician (words and short sentences) and the participant (utterances produced during prior sessions; see later). Identification trials began with a block of nonspeech sounds followed by a block of speech stimuli, because the latter were anticipated to be more challenging for the participant to identify.

The nonspeech sounds were produced live at intensities that were meant to clearly contrast quiet, medium, and loud. The prerecorded speech stimuli were also selected to be sufficiently far apart for it to be obvious which stimuli were of acceptable conversational volume versus too soft or too loud. This evaluation is supported by the reliability data for ratings of production trials (see later).

The student clinician first modeled the activity using three pictograms (clip art depicting the head and torso of men and women), representing the loudness levels, as a visual support. Quiet was represented by a young person whispering something into a woman’s ear; medium/acceptable was represented by a woman speaking (with lines indicating sound coming out of her open mouth); loud was represented by a man grimacing and covering his ears. The participant also used the visual schematic in his responses by pointing to the corresponding pictogram to identify the loudness level. Finally, the student clinician used the visual schematic to provide feedback regarding the accuracy of the participant’s loudness ratings. In cases of misidentification, she indicated the misidentified volume (as selected) using the visual schematic, replayed the stimulus, and then asked the participant to select a new response on the schematic and verbally label it as quiet, medium, or loud. If the second selection was still misidentified, the student clinician provided the answer.
The first session focused on introducing the sound identification task (with both nonspeech sounds and speech stimuli) and the visual schematic. During the first session, the participant also produced sounds (five trials; e.g., clapping, knocking) and words (four trials) at different volumes. This task was discontinued at the participant’s request and was not included in any of the subsequent sessions, although he had completed it successfully.

Production Trials

The production trials took place in the context of question-and-answer conversation activities. The number of trials increased in frequency as the treatment progressed, and achieved a relatively stable level for Sessions 7–9 (see Table 1). This was possible because of the time freed up by the reduction of identification trials and because the participant had achieved a greater level of comfort with the task. The production activities were audio-recorded to provide loudness identification stimuli for use in later sessions and for analysis.

During these activities, the student clinician regularly highlighted instances of acceptable conversational loudness and also provided prompts to reduce loudness. The frequencies of positive feedback and corrective prompts per session were approximately equivalent. For the first three sessions, (positive and corrective) feedback was provided using the same visual schematic as already described for the identification trials. In cases where the participant produced utterances that were excessively loud, the student clinician prompted for a self-correction; she said, “That was a loud voice,” indicated loud on the visual schematic, and asked him to try again with a medium voice while pointing to medium on the schematic (or to the Bla bla bla app). For Sessions 4–9, a cell-phone app (Bla bla bla; Bravi, 2011) served as an online indicator of conversational loudness and was used as feedback instead of the visual schematic. This sound-reactive application shows a face that changes (e.g., mouth and eyes become larger) when sound volume increases. It was included as a means of providing clear feedback for the participant regarding vocal volume—both his and that of the student clinician, for comparison.

Use of the static pictograms had worked well at the beginning of the intervention, but the student clinician decided to introduce a more active form of feedback in order to keep the participant engaged in later sessions. During the transition/familiarization phase, the participant chose which face he preferred for the application. The student clinician then demonstrated how the face changed in response to different volumes and pointed to the corresponding picture on the visual schematic (e.g., large mouth and eyes on the application in response to a loud voice associated to loud pictogram). The participant was then provided with the...
opportunity to practice with the application prior to using it in the identification trials. The student clinician cued him to monitor it and asked him to indicate if changes occurred to the face when either of them spoke. The participant held the cell phone at a comfortable distance to see the visual feedback. The application was then used as feedback during production trials for the rest of the sessions, the participant appearing to enjoy this form of monitoring and feedback.

During the first few sessions in particular, the student clinician did not prompt the participant to self-correct all utterances that he produced with excessive loudness but rather focused on developing his awareness and understanding of the different volumes. Furthermore, given that a high proportion of the participant’s utterances were excessively loud at the beginning of treatment (58% and 24% for the first two sessions, respectively; see Table 1) and that he was prone to anxiety upon correction, systematic prompting would likely have increased his level of anxiety. By Session 7, once the participant had gained confidence, was producing more utterances with acceptable loudness (≥91%), and was more tolerant of prompting, the student clinician prompted him to reduce his loudness for all utterances produced with excessively loud voice (see Table 2). The frequency of prompting varied from six to 17 instances per session, which represented 3% to 9% of total utterances produced by the participant. The proportion of excessively loud utterances that were prompted was below 30% for Sessions 1, 2, and 6, but reached 100% for Sessions 7–9.

### Group-Session Observations

As mentioned previously, the regular ALL group sessions served primarily as a context in which to document generalization rather than as additional therapy time. The group meets 10 months of the year, once weekly for 2 hr. Learning goals are broad, and one of the main objectives of the ALL program is to help participants improve in social-communicative competence. Two speech-language pathologists lead the group of 10–12 adult participants with the assistance of volunteers. Sessions begin with information sharing, conversational exchanges, and planning. They end with a review and wrap-up of information gained within small-group breakout sessions.

During group sessions, the participant’s spontaneous use of moderate conversational loudness was acknowledged on occasion either publicly or privately by one of the speech-language pathologists or authors. Prompting for reduction of vocal loudness to an acceptable level was not generally provided. The first author kept written notes on the number of utterances produced by the participant and rated each utterance as it occurred as quiet, medium (acceptable), or loud (unacceptable; see later).

### Analysis

In order to evaluate outcomes, a 3-point listener judgment scale was used: quiet, medium/acceptable, and loud (in terms of conversational acceptability). The judgment task was in accordance with the primary treatment goal of increasing social-communicative competence in regular conversational settings and followed previous research methodology in which noninstrumental auditory-perceptual procedures have been used (Hollien et al., 1991; Koegel & Frea, 1993; Ozdemir, 2008; Page & Balloun, 1978; Shriberg et al., 2001; Shriberg & Widder, 1990). Judgment tasks are also supported by research indicating that listeners can make even finer grained loudness judgments than the scale used in the current study (Brandt et al., 1969).

Using headphones, the first author listened to the digitally audio-recorded utterances from an individual session and rated them as loud (unacceptable) or medium (acceptable) conversational loudness. (Quiet voice did not occur.) The files were played over a computer with volume set at medium output gain (no sound waves were portrayed). Another final-year master’s student in speech-language pathology who was unaware of the purposes of the study rated 10% of randomly selected utterances using the same procedure, with 93% agreement. For the group sessions, only online ratings for vocal loudness were obtained.

Percentage of spontaneous utterances with acceptable conversational loudness over total utterances was calculated for the prebaseline sample and for six of the nine individual sessions: Sessions 1, 2, and 6–9. We also considered responses to prompts (to diminish excessive loudness), but separately from spontaneous utterances. Data from the first 2 weeks were chosen in order to provide a baseline at the beginning of treatment. Data from the last 4 weeks were selected because improvement was deemed most likely to have appeared by that point. For the group sessions, percentage of spontaneous utterances with acceptable conversational loudness was calculated only for Weeks 1 and 6–9.

### Table 1. Percentage of unprompted utterances with acceptable conversational loudness in the individual and group sessions.

<table>
<thead>
<tr>
<th>Session type</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 6</th>
<th>Week 7</th>
<th>Week 8</th>
<th>Week 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>42% (69/163)</td>
<td>76% (70/92)</td>
<td>79% (104/132)</td>
<td>96% (163/169)</td>
<td>97% (156/161)</td>
<td>91% (172/189)</td>
</tr>
<tr>
<td>Group</td>
<td>25% (3/12)</td>
<td>—</td>
<td>74% (14/19)</td>
<td>83% (19/23)</td>
<td>83% (10/12)</td>
<td>88% (7/8)</td>
</tr>
</tbody>
</table>

Note. Em dash indicates data not obtained.

---

3For further information on the ALL group, please contact the second or third author.
Results and Discussion

The first three sessions included identification trials with both live nonspeech sounds (e.g., clapping, knocking, sounds played on a keyboard) and prerecorded speech stimuli (words and sentences). The participant was able to identify the stimuli as quiet, medium, or loud in 89% (8/9), 83% (10/12), and 100% (6/6) of nonspeech trials and 80% (8/10), 83% (5/6), and 100% (6/6) of speech trials for Sessions 1–3, respectively. He was also able to self-correct the occasional misidentifications. For Sessions 4–9, only speech trials were included (generally nine per session), all of which were accurately identified. The participant thus demonstrated awareness regarding the relative loudness levels of auditory stimuli, in addition to off-line recognition of his vocal volume and that of the student clinician.

Turning to production trials, as Table 1 shows, during the first individual treatment session 42% of the participant’s utterances were rated as having acceptable conversational loudness. The percentage of spontaneous utterances with acceptable loudness increased to 76% in Session 2, then 79% in Session 6, and remained above 90% for Sessions 7–9. The participant also reduced loudness successfully after the majority of prompts for Sessions 7–9 (94%–100%, see Table 2). The group setting also showed increases in percentage of utterances with acceptable conversational loudness, although at a lower level than in the individual sessions (and with many fewer tokens): 25% in Week 1, 74% in Week 6, 83% in Weeks 7 and 8, and 88% in Week 9 (see Table 1). Thus, the participant appeared to benefit from the focus on vocal loudness, even in the group setting, where it was not generally addressed.

Clinical Implications, Limitations, and Future Research

This clinical focus article highlights the importance of intervention and research for adults with developmental delays and, in this case, ASD with sudden vocal outbursts and excessively loud conversational speech—an area that has limited research evidence to date. Such a clinical case study can remain only exploratory and has several limitations. Both for the individual participant and for the greater population, generalizability remains a question because of the short-term nature of the treatment and the flexible approach taken. Treatment techniques were adjusted in relation to the participant’s needs in each session. Using a more stringent design would have allowed stronger conclusions to be made. However, given the participant’s anxiety, the clinicians placed maximal importance on adaptability.

Another limitation concerned evaluation of vocal loudness in the group context; naturally occurring utterances within the group setting were less frequent than in the individual sessions and were judged online by a single rater. Further research should use additional controls, including a longer baseline period closer to the start of treatment and recordings of participant utterances during both the individual and group sessions.

In spite of these limitations, the data indicate that the participant’s use of excessive vocal loudness declined, both in the individual and the ALL group settings after only nine weekly sessions. He also generally responded to requests to reduce his vocal loudness by using acceptable conversational loudness. These changes can be interpreted in competing yet nonexclusive ways, including a change in internal calibration of acceptable vocal loudness, an increase in awareness of others’ negative reactions to loud outbursts, and an increase in self-regulation. Regardless of which (if any) of these interpretations is correct, there is reason to think that the changes observed in conversational loudness and responses to feedback may lead to more positive perceptions and reactions from potential communication partners and increased conversational interactions in various contexts.

Informal observation in the ALL group in the months following this intervention revealed consistent use of acceptable loudness most of the time and continued self-corrections in response to feedback when required, which are positive outcomes for all involved. The individual treatment sessions provided a direct, relatively private opportunity to address a communication issue with negative social consequences, and the group provided a context for immediate generalization. This preliminary report, albeit with some limitations, adds to a very sparse intervention literature on suprasegmentals in the speech of individuals with ASD and intervention with young adults with developmental disabilities. Communicative partners tend to react negatively to unaccustomedly loud voices, and such vocal loudness may be more frequent in at least some speakers with developmental challenges. Further exploration of the effects and effectiveness of treatment for vocal loudness appears very much warranted in children and adults.

Communication challenges for people with developmental delays often persist into adulthood. This clinical focus article draws both clinical and research attention to this underserved group with the intent to stimulate further intervention research, particularly as it concerns aspects of social-communicative competence, including suprasegmental factors that may impinge on successful social interaction.

Table 2. Frequency of prompting for vocal loudness and of adjustment to loudness following prompting in the individual sessions.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 6</th>
<th>Session 7</th>
<th>Session 8</th>
<th>Session 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompted utterances/total utterances</td>
<td>16% (15/94)</td>
<td>14% (3/22)</td>
<td>29% (8/28)</td>
<td>100% (6/6)</td>
<td>100% (5/5)</td>
<td>100% (17/17)</td>
</tr>
<tr>
<td>Prompted utterances/total utterances</td>
<td>9% (15/163)</td>
<td>3% (3/92)</td>
<td>6% (8/132)</td>
<td>4% (8/169)</td>
<td>3% (5/161)</td>
<td>9% (17/189)</td>
</tr>
<tr>
<td>Postprompt adjusted utterances</td>
<td>40% (8/15)</td>
<td>100% (3/3)</td>
<td>50% (4/8)</td>
<td>100% (6/6)</td>
<td>100% (5/5)</td>
<td>94% (16/17)</td>
</tr>
</tbody>
</table>
Acknowledgments

The authors wish to express their sincere appreciation to the participant and his family, ALL group members, and the School of Audiology and Speech Sciences at the University of British Columbia for space and support.

References


