PROSODIC EXPRESSION OF CONTRAST IN WILLIAMS SYNDROME

Kiwako Ito & Marilee Martens

Ohio State University
Ito.19@osu.edu, martens.22@osu.edu

ABSTRACT

Williams syndrome (WS) is a unique genetic developmental disorder characterized by excessive social behaviour accompanied by exaggerated prosody. While past studies with off-line assessment of prosodic skills report significant delays in prosodic skills in individuals with WS as compared to typically developing controls, reports typically lack speech analysis. The current study, using a speech elicitation task, tests whether individuals with WS can express contrast with prosody and what phonetic cues they produce to signal contrast.

The data from 10 WS and 10 chronological-age (CA) matched control participants demonstrate striking similarities in their use of duration and intensity to express contrast. However, F0 did not systematically mark contrast in the WS group. The present data confirm the sensitivity to words' discourse status and the ability to mark contrast in WS, and suggest that the past low ratings of prosodic skills in WS may reflect the lack of reliable F0 cues.

Keywords: prosody, contrast, production, Williams syndrome.

1. INTRODUCTION

Williams syndrome is a rare genetic developmental disorder caused by a microdeletion of up to 28 genes on chromosome 7q11.23 [5, 6]. Their excessive social and verbal behaviour has been repeatedly reported over the last three decades [1], and their 'relative strength' in expressive language, despite their general cognitive delay, has evoked a debate over the independence of linguistic function from the development of general cognition [2, 4]. Many recent studies provide evidence against the superiority of linguistic skills in WS, and claim that the linguistic performance of individuals with WS is predictable from their verbal mental age [2, 6].

Prosodic skills in individuals with WS are also reported as 'atypical' in recent studies. Most of the studies that have shown a difference between the WS group and the control groups, either CA-matched or verbal mental age-matched (VMA-matched), utilize an assessment battery called the Profiling Elements of Prosody in Speech-Child (PEPS-C) [7, 8, 9]. PEPS-C consists of assessment tasks for four sub-domains: Chunking, Affect, Interaction, and Focus; with input (comprehension) and output (production) tasks for each domain. The battery has enabled comparisons of prosodic skills across different age groups and populations with different developmental disorders, and has allowed for the comparisons of results across studies [3, 8]. However, the assessment of prosodic skills with PEPS-C relies on rating scores for each task and does not provide quantitative speech analysis. Till today, studies that report phonetic characteristics of speech in WS are sparse, despite their importance in determining the exact phonetic factors that may lead to difficulty in oral communication with individuals with WS.

This study investigates the ability of individuals with WS to express contrast (which is also often termed as 'narrow focus') with prosody. Past studies report that individuals with WS are clearly delayed in all of the four output domains as compared to CA-matched groups [3, 9]. Another recent study also showed a delayed developmental onset in expressing focus with intonational prominence in the WS group as compared to the CA group [8]. While these studies seem to converge on the inability to appropriately produce focus prosody in WS, a question remains as to how exactly individuals with WS deviate from typically developing individuals in their use of prosodic prominence. The present study devises a simple sequential picture-naming task to elicit spontaneous speech production in individuals with WS. The sequences of target pictures were manipulated such that the presence or absence of contrast and the locus of contrast altered across trials. If individuals with WS are in fact incapable of changing the rhythm and intonation according to the discourse status of the target words, none of the primary phonetic correlates of prosodic emphasis (e.g, F0, duration and intensity) should vary systematically according to the sequence of visual prompts. Alternatively, individuals with WS may fail to produce particular types of prosodic cues while maintaining their ability to change other cues in a comparable manner, as is found in typically developing children and adults.

2. EXPERIMENT

2.1. Participants

Ten individuals with WS and ten CA-matched control participants (total n= 20; Age 10-35yrs, average 18;06) were recorded while they performed the sequential picture-naming task as a part of a larger study. The overall IQ scores, and the estimated verbal and non-verbal mental age (according to their scores on Kaufmann Brief Intelligence Test, 2nd ed.) of the two groups are summarized in Table 1. All participants received a monetary compensation for their participation in the study.

Table 1: IQ-scores and mental age of participants.

Groups	WS (M=4; F=6)	CA (M=7; F=3)
Overall IQ	40-95 (Avg: 67)	99-126 (Avg: 107)
Verbal Age	6;03-16;00 (Avg: 10;01)	10;04-18;06 (Avg: 16;08)
Non-verbal Age	4;04-18;06 (Avg: 8;05)	9;03-18;06 (Avg:12;10)

2.2. Materials and Design

A total of 28 slides were prepared as visual prompts using Adobe Photoshop CS2. Each slide was divided into nine cells, and each of the eight cells surrounding the centre cell contained a drawing of an animal holding an object. Each slide contained four types of animals and four types of objects. On each slide, one cell was highlighted with a red background. A total of 14 pairs of slides were created such that the highlighted cell shifts while the layout remains unchanged between the paired slides. The four target animal-object combinations (gorilla with a basket, koala with an umbrella, racoon with a racket, rabbit with a balloon) were highlighted in the second slide of a pair and appeared once in each of the three types of sequence:

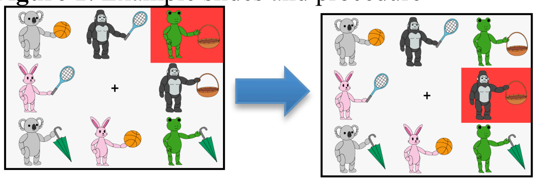
(1) Contrast on the animal (AC):
frog with a basket → gorilla with a basket

(2) Contrast on the object (OC):
gorilla with a ball → gorilla with a basket

(3) No contrast: (NO):
duck with an umbrella → gorilla with a basket

The locations of targets (highlighted cells) were roughly counterbalanced such that participants could not predict their locations in any trial. To control the overall frequency of mention, each of the four target combinations appeared exactly three times across the 12 trials.

Figure 1: Example slides and procedure



In addition to the slides with animal drawings, a total of 13 photos of familiar objects (e.g., muffin, piano, lamp, etc.) were prepared to intervene the sequential presentation of slides after each pair.

2.4. Procedure

Participants were seated in front of a computer monitor and wore a headset XLR microphone (Shure SM10A). Each participant was first familiarized with the names of animals (duck, frog, fox, gorilla, koala, panda, rabbit, racoon) and the objects (balloon, bat, basket, ball, racket, umbrella), and asked to name each target with the frame [XXX with XXX] (e.g., fox with a bat). Participants were told to name the combination highlighted with a red background on each slide, and also to name the objects presented in isolation.

The experiment began with two practice slide pairs (frog with a racket → panda with a bat; panda with a ball →duck with an umbrella) to check the participant's understanding of the task and the recording level. The following 12 trials elicited the production of the four target phrases in three conditions in a random order. Each trial began with the first slide with a beep. Once the participant finished naming the first target, the experimenter clicked on the cell with a computer mouse and it switched to the second slide (i.e., something else was highlighted). When the participant finished naming the second target and the experimenter clicked on the cell, an object photo (e.g., pie) appeared and the participant named it. This photo naming was inserted to prevent the previous slide layout and naming sequence from affecting the discourse status of the following targets.

Each participant was recorded on a PC with a sampling rate of 44.1KHz using Praat (via. TUBE MP preamplifier). Each recording lasted for 3 to 4 minutes.

2.4. Acoustic Analysis

For each target phrase, duration, mean f0, and mean intensity (RMS) of the animal noun and the object noun were measured using Praat. To compare the changes in these acoustic dimensions of the critical words across the utterances that varied in speech rate, overall pitch range and intensity within and across participants, log of the ratio between the two values was calculated for each measurement for each phrase. For example, the relative duration of two words within a phrase of a particular trial was captured by:

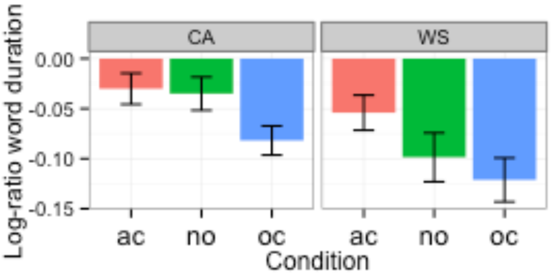
(4) log (animal duration / object duration)

Thus, if the animal noun is produced longer than the object noun in a given phrase, log ratios of the measures should have a positive value. Likewise, if the object noun has a higher value than the animal noun for any of the three measurements, the log ratio should become a negative value. If the two words had similar values for a given measurement, the log ratio should approach 0. Therefore, these dependent variables roughly indicated which of the two words in a given phrase had relative prosodic prominence.

3. RESULTS

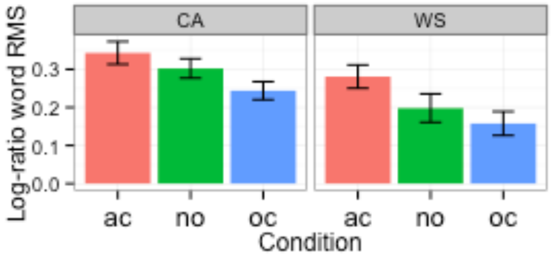
The two tested groups showed unexpectedly similar changes in the duration and intensity of words according to the sequence types. First, the duration analysis showed that the object noun that was produced at the end of an utterance was consistently longer than the preceded animal noun, regardless of the sequence in both the CA and the WS groups (as shown with the negative log ratio values in Figure 2). In both groups, the durational difference was smallest in the animal contrast (AC) sequence, and largest in the object contrast (OC) sequence. The mixed effects model revealed a main effect of sequence (β=.06, t=3.83, p<.01), and no effect of group (β=-.04, t=-1.6) nor the interaction between sequence and group (β=.02, t=0.58). Further analysis revealed that in the CA group, the log ratio of the duration in OC was significantly lower than AC (t=-2.95, p<.05) and NO (t=-2.5, p<.05), while AC and NO did not differ from each other (t=.43). In the WS group, the only significant difference in log ratio of duration was found between AC and OC (t=2.19, p<.05).

Figure 2: Change in relative **duration** of words: ac: animal contrast; oc: object contrast; no: no contrast



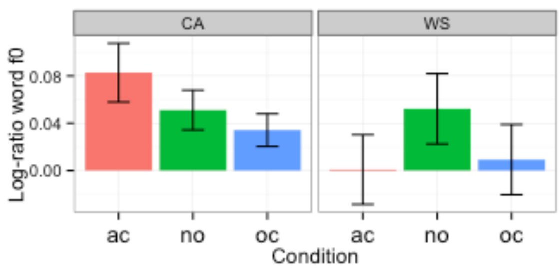
The intensity measure revealed that in both groups, the animal noun that appeared earlier in the phrase had consistently higher intensity than the following object noun regardless of the sequence (as shown with the positive log ratio values in Figure 3). In both groups, the intensity difference was largest in the AC sequence, and smallest in the OC sequence. Again, the mixed effect model revealed a robust effect of sequence (β=.11, t=4.45, p<.01) but neither the effect of group (β=-.09, t=-1.6) nor the interaction between group and sequence (β=.02, t=0.4). Further analysis showed that in both groups, the only significant difference among the sequence types was between AC and OC (CA: t=-2.68; WS: t =-3.5).

Figure 3: Change in relative **intensity** of words



Interestingly, the F0 measure showed a very distinct pattern between the two groups (Figure 4). The CA group showed the consistently higher mean F0 for the animal noun than for the object noun regardless of the sequence, and this difference was largest in the AC sequence and smallest in the OC sequence. In the WS group, however, the mean F0 of the animal noun was higher than that of the object noun only in the NO sequence. When there was a contrast on either the animal or on the object, the mean F0 of the two words within a phrase did not differ much. The mixed effects model showed no sequence effect (β=16.1, t=1.6), no group effect (β=-6.03, t=-0.4), and no interaction between the two (β=-19.05, t=-0.9). Further analysis revealed that the log ratio of F0 was marginally higher for AC than OC in the CA group (t=-1.88, p<.1), whereas it had no significant differences among the sequence types in the WS group.

Figure 4: Change in relative **F0** of words



3. DISCUSSION

The present data demonstrate interesting similarities and differences in prosodic expression of contrast between the two tested groups. First, the WS group performed similar to the CA-matched group in their use of duration and intensity in expressing contrast. In both groups, the relative duration of the object noun was reliably longer when the contrast was on the object than when it was on the animal, and the relative intensity of the animal noun was reliably higher when the contrast was on the animal than when it was on the object. Thus, speakers seemed to use different cues to mark contrast according to the position of the target word in the phrase; words that appear earlier in the phrase were marked with higher intensity, while words that appear at the end of utterance were marked with longer duration.

In contrast with the previous reports on the focus expression in WS, speakers in the WS group, who received no explanation of contrast in the instruction of the task, produced fairly consistent changes in their speech signals according to the visual cues. Because the animals and objects were repeated equally within each slide and thus multiple contrastive relations existed in the visual context of each trial, the reliable changes in the prosodic cues in the present data suggest that the participants with WS were sensitive to the *discourse-based* contrastive status of words and could express the contrast appropriately using intensity and durational cues. In the focus output task of PEPS-C, a participant is asked to request a card that matches his card instead of what the experimenter suggests (e.g., “What about a green bike?” to elicit “I want a WHITE bike.”). Although this task aims to elicit contrast in a natural discourse context, inappropriate responses may be produced due to a poor understanding of the rules of the game, or failure to attend to the experimenter’s utterances, and not necessarily due to the general insensitivity to contrast or the incapability to use prosody to express contrast. Prosodic skills, which are linked to the multiple levels of meta-linguistic skills, should be evaluated carefully with the consideration of factors that may affect the responses in each particular task.

Along with the similar use of intensity and duration, the present data revealed a clear difference in the use of F0 between the CA and the WS group. With the natural pitch declination within an utterance, producing a large pitch range expansion is expected to be easier in the phrase-initial than in the phrase-final position. The CA group exhibited the predicted changes in the F0 values: the mean F0 is always higher for the animal than for the object and the difference becomes larger when the animal is contrasted than when the object is contrasted. The WS group showed the trend of F0 declination only in NO sequence, and the F0 range for the animal and the object nouns became similar for both of the contrastive sequences. Thus, it is possible that the past low ratings of prosodic skills in WS were due to the detection of atypical F0 changes. Although more detailed acoustic measurements such as the duration and the intensity of stressed syllables, the vowel quality, and the F0 peak alignment would better characterize the speech in WS, the loss of clear segmental boundaries in some speakers challenges such analyses. Additional speech data and further exploration of analysis strategies are needed for the better understanding of prosody in WS.

7. REFERENCES

[1] Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St. George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12* (Suppl.), 7-29.

[2] Brock, J. (2007). Language abilities in Williams syndrome: A critical review. *Development and Psychopathology, 19,* 97-127.

[3] Catterall, C., Howard, S., Stojanovik, V., Szczerbinski, M., & Wells, B. (2006). Investigating prosodic ability in Williams Syndrome. *Clinical Linguistics and Phonetics*, 20(7–8), 531–538.

[4] Karmiloff-Smith, A., Brown, J. H., Grice, S., & Paterson, S. (2003). Dethroning the myth: Cognitive dissociations and innate modularity in Williams syndrome. *Developmental Neuropsychology, 23,* 227-242.

[5] Martens, M. A., Wilson, S. J., & Reutens, D. C. (2008). Research Review: Williams syndrome: a critical review of the cognitive, behavioral, and neuroanatomical phenotype. *The Journal of Child Psychology and Psychiatry. 49*(6), 576-608.

[6] Mervis, C. B., &John, A. E. (2010). Cognitive and behavioral characteristics of children with Williams syndrome: Implications for intervention approaches. *American Journal of Medical Genetics Part C (Seminars in Medical Genetics), 154C*, 229-248.

[7] 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 & Phonetics*, *17*(4-5), 345–354.

[8] Stojanovik, V. (2010). Understanding and production of prosody in children with Williams syndrome: A developmental trajectory approach. *Journal of Neurolinguistics, 23*, 112-126.

[9] Stojanovik, V. Setter, J. E., & van Ewijk, L. (2007). Intonation abilities of children with Williams syndrome: A preliminary investigation. *Journal of Speech, Language, and Hearing Research, 50,* 1606-1617.