The effectiveness of speech training on speech prosody of children with autism spectrum disorder

Bruce Xiao Wang[13], Sarah Si Chen[2345], Fang Zhou[2], Angel Chan[2], Tempo Tang[2]

[1]Department of English and Communication, Hong Kong Polytechnic University, Hong Kong, Hong Kong.
[2]Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University, Hong Kong, Hong Kong.
[3]Research Institute for Smart Ageing, Hong Kong Polytechnic University, Hong Kong.
[4]Research Centre for Language, Cognition, and Neuroscience, The Hong Kong Polytechnic University
[5]The Hong Kong Polytechnic University – Peking University Research Centre on Chinese Linguistics

1. Introduction

Autism spectrum disorder (ASD)

- A neurodevelopmental disorder
- Deficits in social communication and interaction [1]
- Difficulties in perceiving and producing reciprocal prosodic cues (e.g., focus marking)

Speech prosody

- Important communicative functions, e.g., affective, pragmatic and syntactic [2]; Changes in the prosody leads to change in sentence meaning [3]
- Focus: From a functional perspective, focus refers to an emphasis on some part of a sentence as motivated by a particular discourse situation.

Short-term lab perceptual training

- Effectively modifies perceptual mechanisms and Improves perception of speech prosody typically developing population.
- Improves speech production [6,7].

2. Aim

- To test the effectiveness of speech training on English prosodic focus marking produced by Cantonese-speaking children with autism spectrum disorder (CASD)

3. Methods

Participants

	Gender	Age	IQ	CASL
CASD	4F;12M	9.62	101.63	63.38
ETD	4F;12M	9.86	108.88	109.44

Stimuli: 15 SVO target sentences (with pictures) grouped into conversation pairs.

Focus Type	Precursor Questions	Target Sentence
Broad	What do you see in the picture?	**[Eve is buying the ring]$_F$**.
Narrow initial	Who is buying the ring?	**[Eve]$_F$** is buying the ring.
Narrow medial	What is Eve doing to the ring?	Eve is **[buying]$_F$** the ring.
Narrow final	What is Eve buying?	Eve is buying the **[ring]$_F$**.
Contrastive initial	Mary is buying the ring?	**[Eve]$_F$** is buying the ring.
Contrastive medial	Eve is wearing the ring?	Eve is **[buying]$_F$** the ring.
Contrastive final	Eve is buying the toy?	Eve is buying the **[ring]$_F$**.

Pre- and post-production tests for the CASD group: 15 target sentences for the CASD group.

Speech Training: 6 out of 15 sentences

Three Phases

- Phase 1

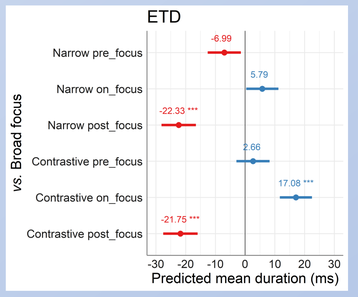
  participants were instructed to distinguish focus marking types (i.e., broad, narrow and contrastive) upon hearing only congruous pairs
- Phase 2
- congruous (odd pairs) & incongruous (even pairs)
- Phase 3

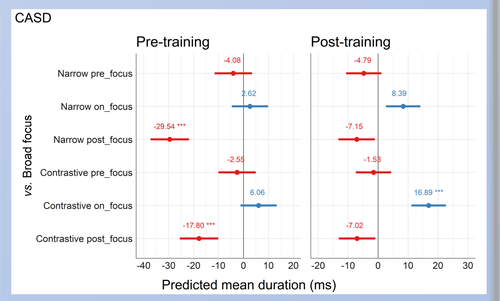
  randomized congruous & incongruous pairs

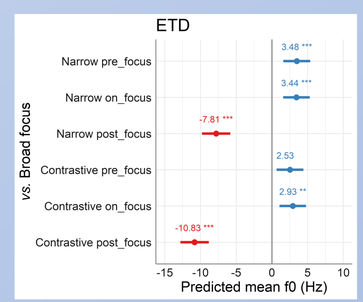
Linear Mixed-Effects Models (LMM) [4]

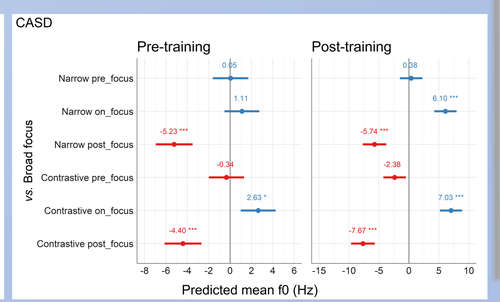
- Response variables: duration, f0, f0 range, intensity
- Fixed effect: Focus condition (**broad focus**, narrow pre/on/post focus and contrastive pre/on/post focus)
- Random effect: word, participant, word type (subject, verb and object)

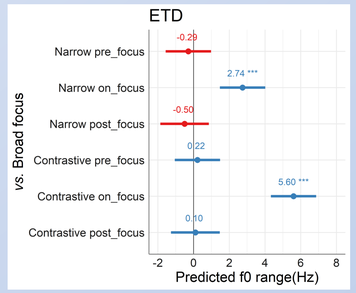
4. Results

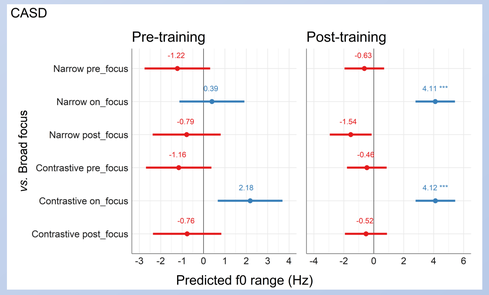


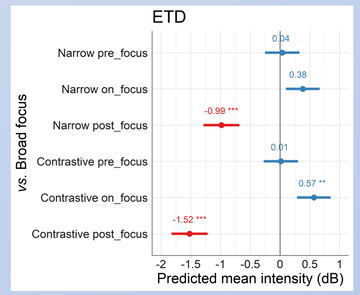


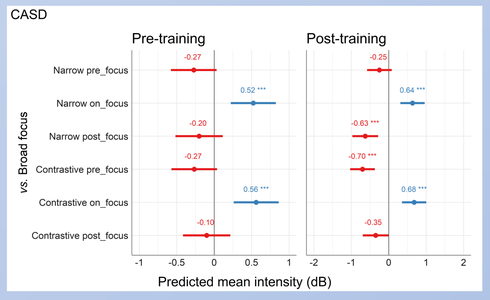












CASD focus marking patterns in post-training production, but not in pre-training production

- Duration: Contrastive on-focus expansion
- Mean F0: on-focus expansion
- F0 range: on-focus expansion
- Intensity: narrow post-focus compression

Prosodic focus marking patterns produced by CASD in the post-training production

- more similar to those produced by ETD, i.e., Both CASD (post-production) and ETD had,

- Increased duration for contrastive on-focus words
- Increased mean f0 & f0 range for on-focus words on focus expansion
- Lowered intensity for words under narrow post-focus conditions

- more similar to healthy native American English adult speakers

- F0 increasing for on-focus words and lowering for post-focus words reported in native American English speakers [5]

Overall, short-term lab perceptual training improves prosody production among ASD children.

6. Conclusion

- CASD used more acoustic cues to signal sentence prominence in the post-training production.
- CASD used duration and f0 cues more than intensity in signaling sentence prominence.
- CASD produced prosodic focus marking patterns more similar to the ETD group after speech training
- Speech training improved the production of prosodic focus marking of CASD.

References
[1]DSM, *American Psychiatric Association, 'Diagnostic and Statistical Manual of Mental Disorders (DSM-5) 5th ed.'* Washington DC: American Psychiatric Association Publishing, 2013. Accessed: Apr. 03, 2023. [Online]. Available: https://www.psychiatry.org:443/psychiatrists/practice/dsm [2]S. Peppé, J. McCann, F. Gibbon, A. O'Hare, and M. Rutherford, 'Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism,' *J. Speech Lang. Hear. Res.*, vol. 50, no. 4, pp. 1015–1028, Aug. 2007, doi: 10.1044/1092-4388(2007/071). [3]A. Cruttenden, *Intonation*. Cambridge University Press, 1997. [4] Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, **67**(1), 1–48. doi:10.18637/jss.v067.i01. [5]Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *J. Phon.*, vol. 33, no. 2, pp. 159–197, Apr. 2005, doi: 10.1016/j.wocn.2004.11.001. [6] Wayland, R. P. & Li, B. Effects of two training procedures in cross-language perception of tones. *Journal of Phonetics* **36**, 250–267 (2008). [7] Chen, S. *et al.* The effects of perceptual training on speech production of Mandarin sandhi tones by tonal and non-tonal speakers. *Speech Communication* **139**, 10–21 (2022).