

Cue-specific prosodic training effects on syntactic disambiguation in Korean children with ASD and neurotypical peers

Jiyoung Lee¹, Sahyang Kim², Taehong Cho¹

¹Hanyang University, ²Hongik University

Prosody is essential in resolving syntactic ambiguity not only in production but also in perception [1, 2]. Listeners utilize prosodic cues, such as pauses and variations in pitch, to interpret the intended structure under ambiguity. While this prosody-syntax mapping develops gradually in typically developing children (TD) [3], children with autism spectrum disorder (ASD) often show atypical prosody in production and perception [4, 5], which may affect their use of prosodic cues for syntactic disambiguation. Yet, some evidence suggests that certain prosodic cue use may be learnable through training [6]. This raises our main questions: whether short-term perceptual training enhances prosody-syntax mapping in Korean ASD children (**ASD**) compared to TD children (**TD**) and neurotypical adults (**Adults**), and whether universal temporal vs. language-specific F0 cues yield different training effects across groups.

Syntactic ambiguity was elicited using Korean particle ‘**-(i)rang**’ in null-subject sentences ‘**N1-(i)rang N2-eul** verb’, where **-(i)rang** licenses comitative or coordinative readings (Fig. 1). The ambiguity was resolved via IP boundaries with a final pitch rise (≈ 10 semitones) and/or a 300-ms pause at or following N1-**-(i)rang** for a comitative reading or N2-**eul** for a coordinative reading. 16 ASD ($M_{\text{age}} = 8.69$), 16 TD ($M_{\text{age}} = 9.2$; age-matched to ASD: Wilcoxon, $W = 87$, $p = .126$), and 14 adults ($M_{\text{age}} = 24.52$) performed a picture-matching perception task across four cue conditions (no-cue, pitch-only, pause-only, and pitch-and-pause) in a pretest-four-day training-posttest design (Fig. 1). Accuracy was analyzed using Bayesian mixed-effects models with weakly informative priors (normal (0, 2.5) for fixed and random effects).

As developmentally expected, accuracy differed credibly across groups ($\text{ASD} < \text{TD} < \text{Adults}$), with adults showing the highest and ASD the lowest accuracy. Overall accuracy improved after training, and a credible difference across prosodic cue conditions was also found ($\text{No-cue} < \text{Pitch-only} < \text{Pause-only} \approx \text{Pitch-and-Pause}$; all 95% CIs exclude zero, except Pause-only vs. Pitch-and-Pause). Full model results are in Table 1. Here we focus on the Group $\times$ Session $\times$ Condition interaction (Tab. 2; Fig. 2). The interaction revealed that training effects differed across groups only when pitch cues were present. Specifically, ASD showed little to no pre-to-post improvement compared to TD and adults in Pitch-only and Pitch-and-Pause conditions, reflecting their restricted capacity to utilize pitch cues for syntactic resolution. In contrast, all groups improved comparably in Pause-only condition, suggesting that ASD, like TD and adults, learned to utilize temporal cues for syntactic disambiguation after training. In No-cue condition, no meaningful training effect was observed across all groups.

The findings suggest that short-term training can enhance prosody-syntax mapping, though not uniformly across groups and cue conditions. When any prosodic cue was available, whether single or combined, adults showed credibly higher accuracy than both child groups, reflecting well-established prosody-syntax mapping regardless of training. TD children also demonstrated adult-like use of pause and pitch cues for syntactic disambiguation after training. In contrast, ASD children showed credible training effects only with temporal pause cues and little to no improvement when pitch cues were present, which is in line with previously reported atypical prosodic processing in ASD [4, 5]. For ASD, language-specific F0 patterns may be treated as low-level acoustic noise rather than meaningful linguistic cues, making it more challenging to link pitch-based prosody to syntactic structure. This cue-specific patterns suggest that ASD children's difficulty may lie not in prosody-syntax mapping itself, but in using language-specific pitch cues for syntactic disambiguation. The results therefore point to a developmental divergence between TD and ASD children in how they integrate universal temporal vs. language-specific prosodic cues into prosody-syntax mapping, and suggest that prosody-syntax mapping in ASD is not globally impaired but cue-dependent.

- **Comitative reading:**
figures on the left
in a & b

- **Coordinative reading:**
figures on the right
in a & b

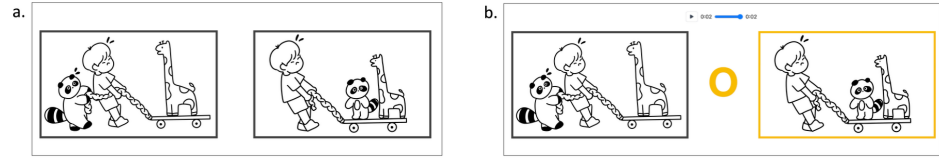


Figure 1. Example screen for the perception task (a) and perception training with trial-by-trial feedback (b). Picture pairs depicting the two interpretations were shown side by side as response options, with left-right positions randomized across trials.

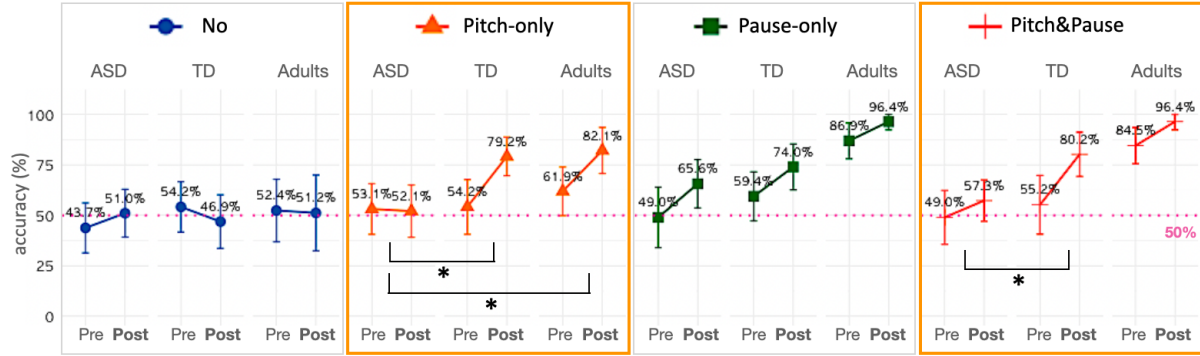


Figure 2. Mean accuracy (%) with SD for ASD, TD, and Adults across four conditions (no-cue, pitch-only, pause-only, pitch-and-pause). Pink dotted line indicates chance level (50%). * indicates a credible group difference in pre-to-post change (95% CI excludes zero). In the no-cue condition, no pitch or pause cue was provided. Without these cues, listeners tended to choose the coordinative interpretation regardless of the intended meaning. Since half of the no-cue stimuli were intended as comitative, overall accuracy remained at chance level (50%).

Table 1. Credible effects from Bayesian mixed-effects model (95% CI excludes zero); † posterior probability = .96 (directional test). Effect coding was used for all predictors. Type indicates interpretation type (comitative vs. coordinative); effects involving Type are not discussed here due to space limitations.

Full model: Accuracy ~ Group * Session * Condition * Type + (1 Participant)			
Effect	Estimate	Est.Error	95% CI
GroupTD_vs_AS	0.51	0.14	[0.24, 0.78]
GroupAdult_vs_Children	1.40	0.18	[1.05, 1.76]
SessionPost_Pre	0.72	0.13	[0.47, 0.98]
Conditionpitch_effect	0.36	0.13	[0.11, 0.62]
Conditionpause_effect	0.99	0.13	[0.74, 1.26]
Conditioninteraction	-0.57	0.26	[-1.07, -0.06]
GroupTD_vs_AS:Conditionpitch_effect	0.47	0.23	[0.02, 0.91]
GroupAdult_vs_Children:Conditionpause_effect	1.94	0.34	[1.29, 2.64]
SessionPost_Pre:Conditionpause_effect	0.69	0.26	[0.17, 1.21]
GroupTD_vs_AS:TypeCom_Cor	0.59	0.23	[0.15, 1.03]
Conditionpitch_effect:TypeCom_Cor	1.29	0.25	[0.80, 1.78]
Conditionpause_effect:TypeCom_Cor	1.89	0.26	[1.39, 2.40]
GroupTD_vs_AS:SessionPost_Pre:Conditionpitch_effect	1.45	0.45	[0.57, 2.35]
GroupAdult_vs_Children:Conditionpitch_effect:TypeCom_Cor	1.27	0.63	[0.06, 2.53]
GroupTD_vs_AS:Conditionpause_effect:TypeCom_Cor	1.40	0.44	[0.53, 2.27]
GroupAdult_vs_Children:Conditionpause_effect:TypeCom_Cor	2.47	0.66	[1.22, 3.81]
GroupAdult_vs_Children:SessionPost_Pre:Conditioninteraction:TypeCom_Cor	-3.01†	1.76	[-6.45, 0.41]

Table 2. Post-hoc pairwise comparisons for the Group × Session × Condition interaction. Only credible contrasts (95% CI excluding zero) are reported. β = difference-in-differences in log odds.

Group × Session × Condition				
Group	Session	Condition	β	95% CI
ASD - TD	Pre - Post	NoCue - PitchOnly	-1.81	[-3.03, -0.64]
ASD - Adult	Pre - Post	NoCue - PitchOnly	-1.43	[-2.79, -0.02]
ASD - TD	Pre - Post	NoCue - Pitch&Pause	-1.64	[-2.91, -0.39]
ASD - TD	Pre - Post	PitchOnly - Pitch&Pause	1.27	[0.08, 2.47]
Group × Session Condition				
Condition = PitchOnly	Group	Session	β	95% CI
	ASD - TD	Pre - Post	1.22	[0.41, 2.11]
	ASD - Adult	Pre - Post	1.14	[0.23, 2.10]
Condition = Pitch_and_Pause	Group	Session	β	95% CI
	ASD - TD	Pre - Post	1.04	[0.10, 1.94]

References

- [1] Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. the Journal of the Acoustical Society of America, 90(6), 2956-2970.
- [2] Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. Journal of psycholinguistic research, 25(2), 193-247.
- [3] Foley, M., Gibbon, F. E., & Peppé, S. (2011). Benchmarking typically developing children's prosodic performance on the Irish-English version of the Profiling Elements of Prosody in Speech-Communication (PEPS-C). Journal of Clinical Speech and Language Studies, 18(1), 19-40.
- [4] McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. International Journal of Language & Communication Disorders, 38(4), 325-350.
- [5] Grice, M., Wehrle, S., Krüger, M., Spaniol, M., Cangemi, F., & Vogeley, K. (2023). Linguistic prosody in autism spectrum disorder—An overview. Language and Linguistics Compass, 17(5), e12498.
- [6] Chen, S., Wang, B. X., Hong, Y. T., Zhou, F., Chan, A., Tang, P. Y., Li, B., Wen, C. Y., Cheung, J., Liu, Y., & Chen, Z. M. (2024). Acoustic changes in speech prosody produced by children with autism after robot-assisted speech training. Proceedings of INTERSPEECH 2024, 4683-4687.