

Using prosody to infer communicative contexts in non or minimally verbal autism

Laura Smorenburg¹, Elanie van Niekerk¹, Janna de Boer², Wolfram Hinzen³, Aoju Chen¹

¹*Utrecht University*, ²*Karakter: Academic Centre for Child and Youth Psychiatry*, ³*University of Pompeu Fabra*

Individuals with non- or minimally verbal autism (NV-ASD), who constitute approximately one-third of the autistic population [1], experience profound challenges in spoken communication due to limitations in both receptive and expressive language [2-3]. Given the limited effectiveness of many language-focused interventions for this group, it is important to examine whether nonverbal vocalisations carry communicative information through prosody. While previous work on verbal autism has often treated prosody as a diagnostic marker, the current study focuses on the communicative potential of prosody. Specifically, we investigate whether pitch-related prosodic features differentiate communicative contexts in NV-ASD vocalisations.

Prosody emerges early in typical development and supports the expression of affect and communicative intent prior to the acquisition of segmental speech. By seven months old, infants already mark some communicative functions such as discontent, satisfaction and requests using prosody [4]. This highlights the potential role of prosody in nonverbal vocalisations in signalling socially meaningful information.

In this pilot analysis, we analysed a subset of an existing dataset [3] comprising video recordings ($M = 31$ minutes, $SD = 7$) from six children and adolescents with NV-ASD (age: $M = 11;4$, $SD = 3;8$, verbal mental age assessed through the Peabody Picture Vocabulary Test-III [5]: $M = 1;8$, $SD = 1;4$). The naturalistic interactions in these videos were structured using the Autism Diagnostic Observation Schedule (ADOS: [6]) or the Adapted ADOS [7] and conducted at therapy centres. All vocalisations were segmented by exhalatory phases [8] and coded using an adapted preverbal vocalisation framework [9], excluding non-intentional sounds and singing. Vocalisations with overlapping speech or ambient noise that affected the pitch tracking were excluded and pitch height and pitch span, operationalised as F0 median and F0 span in semitones and normalised by speaker, were extracted in Praat [10]. Communicative context (Joint Attention, Request, Social Interaction, Stimming/Self-regulation, Non-directed, Unclear) was annotated from muted videos (adapted from [11;12]). Multinomial logistic regression was used to predict communicative context from median and span F0.

Median F0 significantly improved model fit relative to an intercept-only model (likelihood ratio test: $\chi^2(4) = 35.39$, $p < 0.001$), indicating that median F0 systematically relates to communicative context. Higher median F0 increased the log odds of Requests and Self-regulatory vocalisations relative to Non-directed vocalisations –although the standard errors for Requests indicate high uncertainty–, whereas Joint Attention and Social Interaction were more likely at lower median F0 (see Table 1). Despite this significant association, the model's overall classification accuracy was low (28.7%), with reliable identification only for Stimming/self-regulatory vocalisations (see Fig. 1). Adding F0 span did not further improve model fit (likelihood ratio test: $\chi^2(4) = 7.91$, $p = 0.095$).

These findings suggest that, in individuals with NV-ASD, median F0 may primarily reflect arousal-related dimensions of vocal behaviour rather than serving as a robust marker of specific communicative intents. While certain prosodic cues can be used for communicative purposes, prosody alone appears insufficient to distinguish communicative functions. Future work should explore the role of other prosodic, multimodal and contextual cues, to assess which (combination of) cues can more reliably support the interpretation of communicative intent within this population.

Table. 1. MLR log-odds coefficients and standard errors ($N = 698$, reference category = Non-Directed)

Context	Intercept ($\pm$ SE)	medianF0 ($\pm$ SE)
Joint Attention (JA)	-0.602 ± 0.134	-0.334 ± 0.131
Request (RQ)	0.100 ± 0.108	0.203 ± 0.107
Social Interaction (SI)	-1.205 ± 0.165	-0.243 ± 0.162
Self-regulation (ST)	0.142 ± 0.107	0.312 ± 0.106

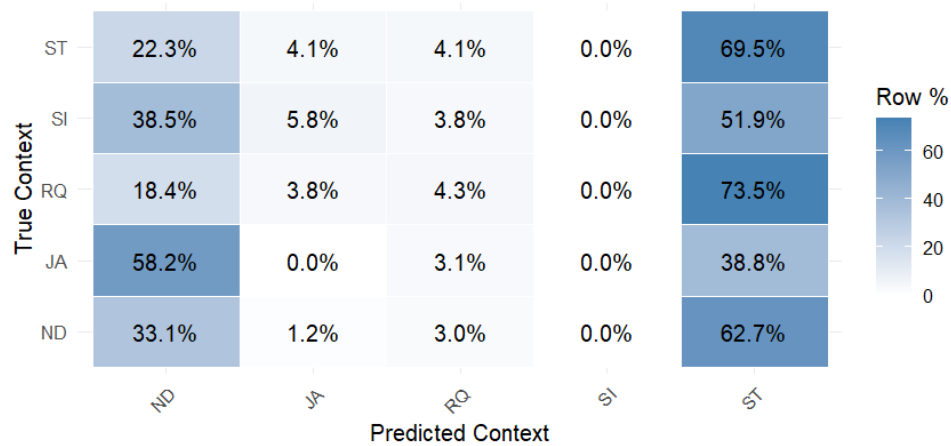


Figure. 1. Confusion matrix with classification accuracies (overall accuracy = 28.65%, chance level = 20%)

References

- [1] Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism research*, 6(6), 468-478. <https://doi.org/10.1002/aur.1329>
- [2] Chenausky, K., Brignell, A., Morgan, A., & Tager-Flusberg, H. (2019). Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism & Developmental Language Impairments*, 4. <https://doi.org/10.1177/2396941519856333>
- [3] Slušná, D., Rodríguez, A., Salvadó, B., Vicente, A., & Hinzen, W. (2021). Relations between language, non-verbal cognition, and conceptualization in non-or minimally verbal individuals with ASD across the lifespan. *Autism & Developmental Language Impairments*, 6. <https://doi.org/10.1177/23969415211053264>
- [4] Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: Evidence from Catalan-babbling infants. *Journal of child language*, 40(5), 919-944. <https://doi.org/10.1017/S0305000912000359>
- [5] Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-III* (3rd ed.). American Guidance Service.
- [6] Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, 30(3), 205-223.
- [7] Bal, V. H., Maye, M., Salzman, E., Huerta, M., Pepa, L., Risi, S., & Lord, C. (2020). The adapted ADOS: A new module set for the assessment of minimally verbal adolescents and adults. *Journal of Autism and Developmental Disorders*, 50(3), 719-729.
- [8] Oller, D. K., & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology: Toward a broader theory of development and disorders. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development* (pp. 509-536). Parkton, MD: York Press, Inc.
- [9] Buder, E. H., Warlaumont, A. S., Oller, D. K., Peter, B., & MacLeod, A. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. *Comprehensive perspectives on child speech development and disorders: Pathways from linguistic theory to clinical practice*, 4, 103-34
- [10] Boersma, Paul & Weenink, David (2026). Praat: doing phonetics by computer [Computer program]. Version 6.4.55, retrieved 15 January 2026 from <https://praat.org>
- [11] Maljaars, J., Noens, I., Jansen, R., Scholte, E., & van Berckelaer-Onnes, I. (2011). Intentional communication in nonverbal and verbal low-functioning children with autism. *Journal of Communication Disorders*, 44(6), 601-614.
- [12] Wetherby, A. M. (1986). Ontogeny of communicative functions in autism. *Journal of autism and developmental disorders*, 16(3), 295-316.