Guiding Implicit Prosody with Delexicalized Melodies: Evidence from a Mismatch Task

Nicholas Van Handel, Matthew Wagers, Amanda Rysling (UC Santa Cruz)

In [1]'s "reading with delexicalized melodies" task, subjects heard low-pass filtered sentences, which lack segmental content but retain prosody, then replicated these melodies during silent reading of a target sentence. This method seems to hold promise for addressing when/how implicit prosody manifests in reading and how implicit prosody interacts with syntactic parsing [2, 3]. There is growing interest in extending this task [4, 5] and in using overt speech to guide reading [6]. Conclusions from this method depend on the extent to which subjects accurately replicate melodies in reading. Holding a sentence melody in memory is potentially difficult, but previous work has not explicitly assessed subjects' ability to project full melodies onto read sentences. Here, we report 4 match/mismatch tasks using more complex stimuli than [1], contrasting simultaneous and sequential presentation of the melody and written sentence.

Method. 36 items manipulated STRUCTURE (NP vs. Z) and MELODY (MATCH vs. MISMATCH); see Table 1. The NP/Z garden path [7, 8] was chosen because NP and Z have clear prosodic differences [9]. A native American English-speaking phonetician recorded all sentences. MISMATCH melodies cross-spliced NP and Z recordings, such that the boundary occurred in the wrong location. Accurate performance required subjects to remember the relative position of the boundary, providing a strong test of subjects' ability to replicate the melody; cf. [1], which only varied the presence of a boundary after the second word in a sentence. In Expt 1, melody and sentence were presented simultaneously, while in Expt 2, the sentence appeared after the melody. Subjects judged the melody as "Match" or "Mismatch" and rated their confidence on a 3-point scale. Responses were converted to a 6-point scale (1=confident "Mismatch"; 6=confident "Match") [10]. Bayesian cumulative link mixed models [11] were fit to ratings (fixed effects: STRUCTURE, MELODY, interaction; maximal random effects).

Expt 1. Simultaneous (n=65). Ratings are summarized in Figure 1. There were main effects of STRUCTURE, such that Z sentences were rated lower (-.69, [-1.02, -.36]), and MELODY, such that MISMATCH were rated lower (-3.03, [-3.55, -2.52]), with no interaction. The MISMATCH penalty confirms that subjects were sensitive to mismatches. The Z penalty suggests that it is harder to compare a melody to a written sentence when the latter contains a garden path.

Expt 2. Sequential (n=38). There was a main effect of MELODY, such that MISMATCH were again rated lower, (-.83, [-1.08, -.59]), but no effect of STRUCTURE, nor an interaction. The MISMATCH penalty shows that subjects distinguish MATCH from MISMATCH, but the effect size was small relative to Expt 1, with worse performance in the MATCH conditions in particular. Subjects also reported difficulty with the task; data from an additional 18 (32%) were excluded for giving higher ratings to MISMATCH melodies, indicating poor performance. Lack of a Z penalty may be the result of compressed ratings and poor performance.

Discussion. Expt 1 showed that subjects compared melodies and written sentences when presented simultaneously, but the Z penalty suggests that the melody did not override default parsing to prevent garden paths. This method may be appropriate to make certain phrasings available, but not to direct implicit prosody in first-pass reading. Expt 2, which required more memory load with sequential presentation, was less effective: while the effect of MELODY suggests a (limited) ability to replicate melodies, we are skeptical that subjects do so reliably enough for the melody to direct their implicit prosody. Poor performance with NP/Z raises doubts about generalizing the task to longer melodies or subtler cues. Preliminary results from replications with slower melodies, to make the task easier, show qualitatively the same effects (Expts 3 and 4, below). We note that [1]'s study included extensive production training, shorter sentences, and easier boundary placement conditions; the present study did not. We thus advise caution with any sequential method, closer to [1]'s original, as the assumption that subjects accurately read with a melody may not hold without [1]'s training and conditions.

STRUCTURE	MELODY	Item
NP	MATCH	[After Anne visited the British relatives %]$_{NP}$ [the cousins moved to the countryside.]$_{NP}$
	MISMATCH	[After Anne visited % the British relatives]$_{Z}$ [the cousins moved to the countryside.]$_{NP}$
Z	MATCH	[After Anne visited % the British relatives]$_{Z}$ [moved to the countryside.]$_{Z}$
	MISMATCH	[After Anne visited the British relatives %]$_{NP}$ [moved to the countryside.]$_{Z}$

Table 1. Sample NP/Z item.

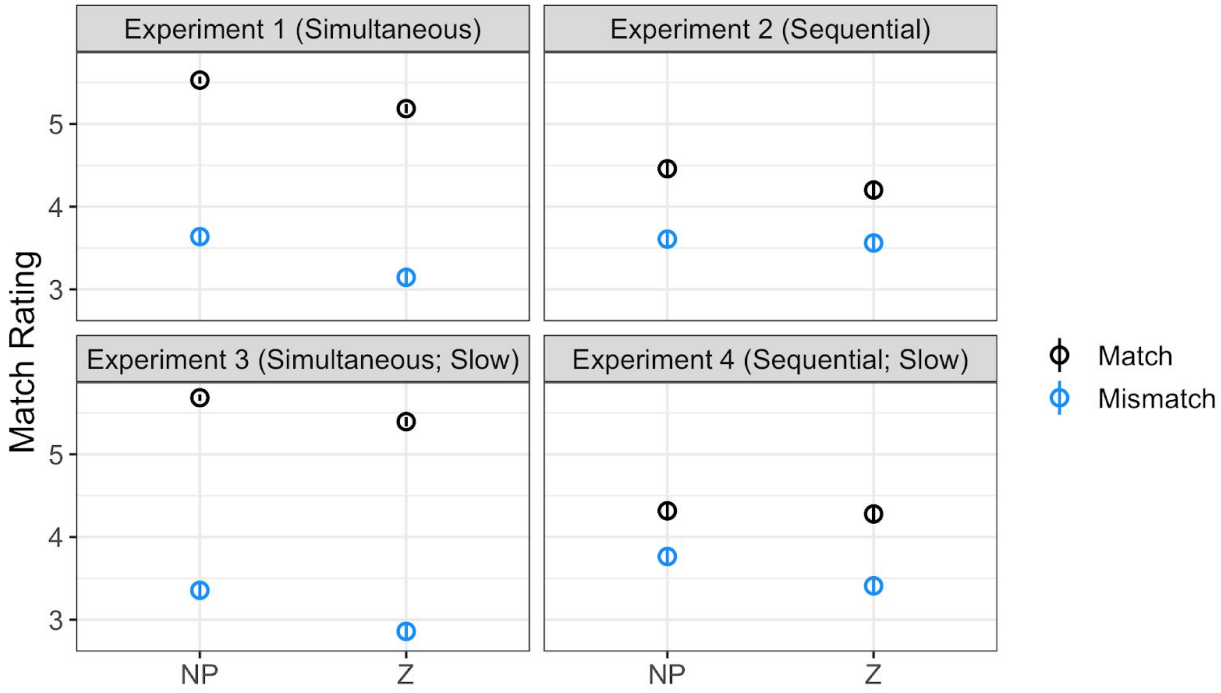

Figure 1. Mean ratings by experiment. Error bars indicate standard error of the mean.

References.

[1] Steinhauer, K., & Friederici, A.D. (2001). *Journal of Psycholinguistic Research*
[2] Fodor, J. (2002). *Speech Prosody 2002*
[3] Breen, M. (2014). *Language and Linguistics Compass*
[4] Luo, Y., Yan, M., & Zhou, X. *Journal of Experimental Psychology: Learning, Memory, and Cognition*
[5] Mills, J. (2020). *Proceedings of Speech Prosody 2020*
[6] Zhang, G., & Husband, E.M. (2019). *Poster at the 32nd CUNY Sentence Processing Conference*
[7] Frazier, L., & Rayner, K. (1982). *Cognitive Psychology*
[8] Frazier, L., Carminati, M.N., Cook, A.E., Majewski, H., & Rayner, K. (2006). *Cognition*
[9] Kjelgaard, M.M., & Speer, S.R. (1999). *Journal of Memory and Language*
[10] Macmillan, N.A., & Creelman, C.D. (2005).
[11] Bürkner (2017). *Journal of Statistical Software*