Do Individual Differences in Working Memory Capacity Predict Cross-Speaker Variation in Planning Scope? Some Prosodic Tests

Jason Bishop[1,2]

[1]CUNY-College of Staten Island (USA), [2]CUNY-Graduate Center (USA)
jbishop@gc.cuny.edu

While several phonetic and phonological patterns indicate that speech production planning unfolds in prosodic phrase-sized chunks [1] rather than one- or two-word sequences [2], research in both phonetic science and psycholinguistics has begun to ask how flexible this planning is, and to what extent it may reflect demands both external and internal to speakers [3]. The present study explored the extent to which speakers' planning may vary in relation to the latter, exploring individual differences in working memory capacity (WMC) as a source of speaker-internal—and thus speaker-specific—constraints on planning. In particular, we tested the hypothesis that speakers with higher WMC engage in more extensive/longer-range planning, exploring two prosodic variables.

The first was speakers' average phrase length, counted in syllables, over a sample of connected speech. We predicted that speakers with higher WMC, if they engage in longer-range planning, would tend to produce longer prosodic phrases (intonationall-defined). Notably, this question has a parallel in the literature on implicit prosody (i.e., prosody generated internally during silent reading), where sentence processing tasks have suggested that the size of readers' implicit prosodic phrases varies along with their WMC [4]. To our knowledge, however, the question of whether WMC predicts variation in the length of speakers' overtly-produced phrases has not been explored.

The second variable was silent pause duration. Pause durations are known to increase as the length of an upcoming phrase increases [5,6] and so are widely assumed to, in part, reflect speakers' planning. Interestingly, however, [7] demonstrated that pause duration is not only predicted by the length of an upcoming phrase, but the phrase's internal structure. In particular, and somewhat counterintuitively, speakers tend to produce shorter pauses when an upcoming phrase is structurally complex. For example, on average, speakers of English and German will produce shorter pauses before an Intonational phrase (IP) that contains two or more smaller intermediate phrases (ip) than before an IP that contains just one ip. [7,8] suggested this reflects the options afforded by complex phrases; complex phrases can be approached more incrementally, while more monolithic structures cannot (see **Fig.1**). The question we asked was whether WMC predicts variation in the extent to which speakers avail themselves to this more incremental planning option. It was predicted that, because speakers with higher WMC should be more likely to plan larger chunks (and thus less likely to engage in the more incremental planning option), their pause durations should be less affected by the complexity of an upcoming IP. Relative to speakers with lower WMC, then, speakers with higher WMC should have longer, not shorter, pause durations before complex IPs.

The hypotheses relating WMC to these two prosodic variables were tested on a corpus of read speech collected in the context of another study [9]. The corpus was based on a production study with a large group of native English speakers from the US (N=100) who completed a standard measure of WMC and who read the same 160-word passage aloud (in a sound-attenuated booth, using a Shure SM10A headworn mic, digitally recorded at 41kHz). The phrase structure of the produced passages was identified using ToBI conventions for Mainstream American English [10]; the length (in syllables) of all speakers' fluent ips and IPs were then counted, and the durations of all pauses preceding fluent IPs were measured. Mixed-effects linear regression was then used to model phrase length and pause duration, with the goal of identifying any effect of WMC that could be related to speakers' planning in the ways described above.

Results indicated the following. First, WMC was to some extent predictive of speakers' phrase lengths; speakers with higher WMC tended to produce longer IPs and (IP-medial) ips (see **Fig.2**), although the effect was only significant when phrase levels were collapsed (*est=-.0417, SE=.0136, t=-2.99 p<.01*). One interpretation of this finding (anticipated in [8]) is that speakers can manage

their WMC-related limitations in different ways, with some speakers adjusting the lengths of their ips, others their IPs, and yet others a combination of both. Second, WMC was also found to predict the sensitivity of speakers' pause durations to the complexity of an upcoming IP; when the length of an upcoming IP was held at its mean, there was no main effect of its complexity on pause durations (*est=.006, SE=.001, t=.355,p>.1*). Instead, the effect of complexity depended on speakers' WMC; longer pause durations preceded complex phrases, but only for speakers with higher WMC (*est=.040, SE=.027, t=2.03, p<.05*), consistent with more advanced—and less incremental—planning by these speakers. We discuss the implications of these findings for speech production planning; for models of how prosodic structure interacts with other components of grammar and language production; and for the use of individual differences approaches in the investigation of these issues.

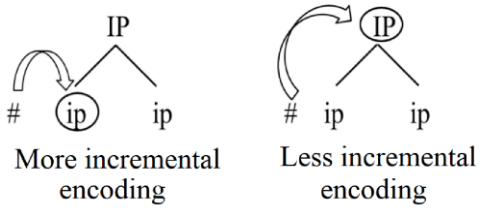

Fig.1 Planning options in complex prosodic structures (based on [8]); speakers can plan at lower nodes or (less incrementally) at higher ones. Planning strategies will be reflected in pause durations (indicated by '#' in the figure).

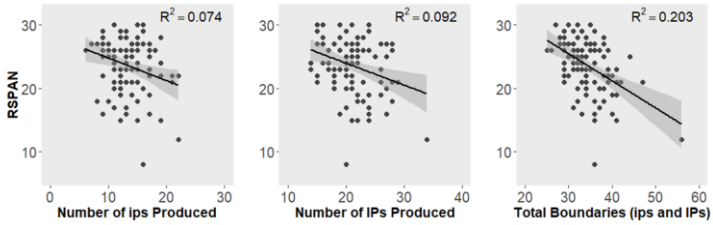

Fig.2 Relationship between speakers' WMC (as measured by reading spans) and the number of phrase breaks produced in the recorded passage. The strongest relation is between speakers' WMC and the total number of boundaries, collapsing phrase levels (ip and IP).

References

[1] Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics*,101, 112–156.

[2] Levelt, W.J. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

[3] Wagner, V., Jescheniak, J., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 423−440.

[4] Swets, B., Desmet, T., Hambrick, D., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: a psychometric approach. *J. Experiment Psychology: General*, 136, 64-81.

[5] Ferreira, F. (1993). Creation of prosody during sentence production. *Psycholinguistic Review*, *100*, 233–253.

[6] Whalen, D., & J. Kinsella-Shaw (1997). Exploring the relation of inspiration duration to utterance duration. *Phonetica* 54, 138–152.

[7] Krivokapić, J. (2007). The planning, production, and perception of prosodic structure. Ph.D. dissertation, University of Southern California.

[8] Krivokapić, J. (2012). Prosodic planning in speech production. In S. Fuchs, M. Weihrich, D. Pape, & P. Perrier (eds.): *Speech planning and dynamics*. Peter Lang. pp. 157–190.

[9] Bishop, J. (resubmitted). Exploring the similarity between implicit and explicit prosody: prosodic phrasing and individual differences. Ms, City University of New York.

[10] Beckmann, M., & Ayers-Elam, G. (1997). Guidelines for ToBI labeling (Version 3). Unpublished ms. Ohio State University.