Prediction cannot be directly trained: An extension to Jones and Westermann (2021)

Sam Jones and Gert Westermann

Lancaster University

Word count: 1930

Author Note

This work was supported by Economic and Social Research Council (ESRC) International Centre for Language and Communicative Development (LuCiD) [ES/S007113/1 AND ES/L008955/1]. We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Sam Jones, Department of Psychology, Lancaster University, Lancaster, United Kingdom, LA1 4YF. Email: sam.jones@lancaster.ac.uk. Telephone: +44 (0) 1524 593698.

Abstract

In January 2021 we published a viewpoint article entitled ‘Predictive processing and developmental language disorder’ (DLD) in the *Journal of Speech, Language, and Hearing Research*. The current commentary provides an important extension to this work. Specifically, we aim to head off the suggestion that a child’s ‘predictive capacity’ may be trained independently of improving the quality of their long-term speech representations.

Keywords: developmental language disorder (DLD), predictive processing, speech-language pathology

Prediction cannot be directly trained:
An extension to Jones and Westermann (2021)

In January 2021 we published a viewpoint article entitled 'Predictive processing and developmental language disorder' (DLD) in the *Journal of Speech, Language, and Hearing Research* (S. D. Jones & Westermann, 2021). In this article, our aim was to introduce the predictive processing framework to a perhaps unfamiliar readership, and to consider how this framework may help re-focus our understanding of the challenges facing children with language learning difficulties.

In our target article, we cited evidence that children with well-developed language skills implicitly anticipate the sorts of linguistic features that they later expect to hear, whether these features are acoustic-phonetic, lexical, syntactic, or semantic (Blank & Davis, 2016; Borovsky et al., 2012; Davis & Johnsrude, 2003; S. D. Jones & Westermann, 2021; Mani & Huettig, 2012; Sohoglu et al., 2012). Active, top-down anticipation of this sort may enable the child to get ahead of the curve and to rapidly resolve perceived ambiguities, supporting efficient speech comprehension. A striking example of this advantage can be seen in tasks involving sentences containing distorted words. Here, top-down anticipatory processing enables adult listeners to accurately decode distorted words, on the basis of perceived sentential context and prior language knowledge (Blank & Davis, 2016; Davis et al., 2005; Sohoglu et al., 2012).

Where language develops more slowly, as it does in children diagnosed with DLD, the effective anticipation of upcoming speech will be necessarily compromised, leaving the child less well prepared to navigate the noise that characterizes natural speech and giving rise to apparently laboured language comprehension (Borovsky et al., 2012, 2012; Hestvik et al., 2022; Mani & Huettig, 2012). Rather than exploiting online anticipatory processing during exposure to the features of an unfolding sentence, a child with language learning difficulties may be relatively more dependent on post hoc sentence element integration and ambiguity resolution (S. D. Jones & Westermann, 2021).

There is a strong possibility that the predictive processing framework can enrich our understanding of the challenges facing children with language learning difficulties. Very different assumptions follow, for instance, from the albeit compatible positions that speech comprehension appears laboured in DLD because; (i) cognitive deficits, for instance commonly assumed working memory capacity limitations (Archibald & Gathercole, 2006), affect the efficiency of processing subsequent to speech making contact with the basilar membrane, which has long been the dominant view within the field (e.g., Montgomery & Evans, 2009), or (ii) because deficits in long-term language memory prevent the child from fully engaging in the top-down anticipation of unfolding speech.

However, one line of discussion that we have encountered since the publication of our target article has caused us some concern. Specifically, on a number of occasions we have encountered the suggestion that, since top-down anticipation forms an integral feature of well-developed language processing (Sohoglu et al., 2012), the communication skills of a child who is struggling with language may be boosted by training that child's 'predictive capacity', independently of improving the quality of their long-term speech representations. This is a direction we cautioned against in our target article, notably in our discussion of an intervention programme developed by Plante et al. (2014) (see S. D. Jones & Westermann, 2021, p. 184), but which we believe deserves further attention.

Despite numerous important points of disagreement, theoretical frameworks invoking a notion of prediction are seemingly united in the position that the implicit expectation of an upcoming percept, such as a noisy word in a spoken sentence, is the product of (i) an active, multimodal sensory state, for example the perception of an unfolding speech string in a given communicative context, and (ii) long-term probabilistic knowledge of the ways in which speech sounds, words, and structures co-occur in associated contexts (e.g., Sohoglu et al., 2012). Real problems arise, therefore, if we attempt to detach predictive processing from activated long-term memory, and to treat prediction both as a functionally discrete faculty and, crucially, as a potential target of clinical intervention.

As a field, it is not the first time that we have made this mistake. Numerous studies have pursued the hypothesis that the proximal cause of DLD is a capacity limitation in a working memory system of the form first proposed by Baddeley and Hitch (1974). The claim that this system, specifically the *phonological loop* buffer component of working memory, was both functionally discrete from long-term speech memory and capacity limited in children with language learning difficulties (e.g., Archibald & Gathercole, 2006), led to the emergence of empirical research and commercial packages of intervention that claimed to be able to boost working memory capacity and in doing so confer gains in communication skills and wellbeing (Alloway et al., 2013; Spencer-Smith & Klingberg, 2015). Working memory training has, however, proved an abject failure, with little compelling evidence that training effects either last over time or transfer across tasks (Melby-Lervåg & Hulme, 2013). As we have written elsewhere, our view is that the absence of any convincing effect here reflects the likelihood that much of the explainable variance in working memory task performance (e.g., in nonword repetition) reflects differences in the precision of activated long-term speech representations, and in associated skills such as motor planning and articulation, and not in the capacity of a functionally discrete working memory buffer system (G. Jones et al., 2020; S. D. Jones & Westermann, 2022).

The move towards working memory training began with a body of research that functionally isolated and attributed a causal role to the phonological loop in early language difficulties. And there is some evidence that we are approaching similar territory with respect to predictive processing. In a recent empirical study, Hestvik et al. (2022) found no neural signature of prediction error during anomalous sentence processing among children with DLD, suggesting that these children were not actively anticipating the upcoming syntactic features of the sentences to which they were exposed. On this basis, Hestvik et al. (2022) characterise DLD as a "syntactic prediction impairment" and attribute a causal role to atypical predictive behaviour, writing that; "this lack of a prediction error signal can interact with language acquisition and result in DLD" (p. 1).

Our own view is rather different. We do not see DLD as a "syntactic prediction deficit" but instead as a deficit principally in long-term speech representation, at all levels of linguistic analysis (e.g., acoustic-phonetics, words, and constructions), which is attributable to an as yet poorly understood constellation of factors including atypical auditory processing (Bishop & McArthur, 2005). Successful predictive language processing is, in our view, the automatic and inevitable consequence of successful language learning, that is, of implicitly knowing what sorts of sounds, words, or constructions tend to co-occur in a given communicative context, and the resulting pre-emptive, top-down activation of this information in an associated context. Reciprocally, prediction error feedback helps to fine tune long-term speech representations in the event of a mismatch between an individual's mental model of their speech environment and the speech that they actually perceive. Atypicality in the active anticipation of upcoming speech is, under this view, the inevitable by-product of low-quality long-term speech and language representations, and should be expected in any area in which language skills are weak, not only in syntax (S. D. Jones & Westermann, 2021, p. 182).

Indeed, undeveloped anticipatory processing skills (inferred by Hestvik et al., 2022, in the absence of a neural signature of prediction error) would be expected in any individual who is unfamiliar with the target structure of the target language being tested, including younger children without neurodevelopmental disorder (Friederici, 2006) or second language learners (controlling, of course, for cross-linguistic similarity). In testing only age-matched control children, Hestvik et al. (2022) do not rule out the possibility that the atypical predictive behaviour observed in their sample is the by-product of low language familiarity, and perhaps adopt a causal position accordingly. In our target article, however, we cited evidence continuous with the view that speech prediction emerges naturally and incrementally as the individual reaches ever higher standards of linguistic awareness (S. D. Jones & Westermann, 2021, p. 182). It was emphasized, for instance, that a neural signature commonly associated with syntax-driven prediction error emerges only when language skills are relatively well developed (see Friederici, 2006, for review). This is an important insight, because it may prevent us from automatically invoking language-independent explanations (e.g., attentional or working memory deficits) upon observing that speech processing and comprehension appear laboured in DLD. Such performance deficits may, instead, be the inevitable consequence of an immature mental model of the speech environment. A child who struggles with language may not actively anticipate upcoming linguistic features not because of an impaired prediction faculty, but because of well-recorded deficits in long-term speech representation.

While low language familiarity means that the advantages of top-down anticipatory processing (e.g., robustness to noise, active feature integration, and rapid ambiguity resolution) may be relatively out of reach for a child with speech and language problems, this does not mean either that a discrete prediction deficit plays a primary causal role in language learning difficulties or, vitally, that prediction should form a target of clinical intervention. This latter claim would, in our view, put us in the impossible position of attempting to ‘fix’ an emergent phenomena (i.e., prediction) while ignoring the constituent underlying processes (i.e., multimodal sensory processing and activated long-term memory). Some form of predictive capacity training may feasibly deliver limited gains in speech skills because the tasks used may be likely to involve structured exposure to speech. However, as in the working memory literature, we would expect such gains to be fragile, showing little evidence of longevity or transfer across tasks relative to the evidence-based methods of improving long-term speech representation quality that already form an important part of the speech and language therapist's toolkit (Melby-Lervåg & Hulme, 2013; Rinaldi et al., 2021).

Careful consideration of this issue is essential because, as noted above, we have been here before, with numerous programmes of research and intervention established on the conviction that the phonological loop buffer system within working memory can be trained independently of long-term speech representations to confer transferable and long-lasting language gains. This track record illustrates how the reification of an emergent phenomena in translational research can result in the ineffective use of resources and a potential collapse in both the confidence of the individual undergoing intervention and trust in professionals when speech and language gains are not seen due to a child being put through support programmes of questionable efficacy.

Predictive processing remains a highly active research area, and as with all things in science it is possible that we will need to revise our view in light of new data. However, the current best evidence suggests that, despite implicating dissociable neural substrates (Ficco et al., 2021), activated long-term memory forms a functionally indivisible component of top-down anticipatory processing. On navigating the world as it unfolds through time, and generating and propagating prediction error signals, the brain can *only* look to its current sensory state and to associated, previously encoded memory traces. A rich mental model of the speech environment is required in order to engage in and benefit from the automatic anticipation of upcoming speech, and such a model is, by definition, deficient in children diagnosed with DLD, as well as those with other forms of language difficulty. Our focus as researchers and practitioners should remain on improving the quality of the long-term speech representations formed by children with language learning difficulties through the continued development and delivery of evidence-based methods (Rinaldi et al., 2021). Gains in anticipatory processing would then be expected to follow as the natural corollary of gains in long-term linguistic awareness.

References

Alloway, T. P., Bibile, V., & Lau, G. (2013). Computerized working memory training: Can it lead to gains in cognitive skills in students? *Computers in Human Behavior*, *29*(3), 632–638. https://doi.org/10.1016/j.chb.2012.10.023

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders*, *41*(6), 675–693. https://doi.org/10.1080/13682820500442602

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol. 8)* (pp. 47–89). Academic Press.

Bishop, D. V. M., & McArthur, G. M. (2005). Individual differences in auditory processing in specific language impairment: A follow-up study using event-related potentials and behavioural thresholds. *Cortex*, *41*(3), 327–341. https://doi.org/10.1016/S0010-9452(08)70270-3

Blank, H., & Davis, M. H. (2016). Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLOS Biology*, *14*(11), e1002577. https://doi.org/10.1371/journal.pbio.1002577

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*(8), 3423–3431. https://doi.org/10.1523/jneurosci.23-08-03423.2003

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, *134*(2), 222–241. https://doi.org/10.1037/0096-3445.134.2.222

Ficco, L., Mancuso, L., Manuello, J., Teneggi, A., Liloia, D., Duca, S., Costa, T., Kovacs, G. Z., & Cauda, F. (2021). Disentangling predictive processing in the brain: A meta-analytic study in favour of a predictive network. *Scientific Reports*, *11*(1), 16258. https://doi.org/10.1038/s41598-021-95603-5

Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron*, *52*(6), 941–952. https://doi.org/10.1016/j.neuron.2006.12.002

Hestvik, A., Epstein, B., Schwartz, R. G., & Shafer, V. L. (2022). Developmental Language Disorder as Syntactic Prediction Impairment. *Frontiers in Communication*, *6*, 637585. https://doi.org/10.3389/fcomm.2021.637585

Jones, G., Justice, L. V., Cabiddu, F., Lee, B. J., Iao, L.-S., Harrison, N., & Macken, B. (2020). Does short-term memory develop? *Cognition*, *198*, 104200. https://doi.org/10.1016/j.cognition.2020.104200

Jones, S. D., & Westermann, G. (2021). Predictive processing and developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *64*(1), 181–185. https://doi.org/10.1044/2020_JSLHR-20-00409

Jones, S. D., & Westermann, G. (2022). Under-resourced or overloaded? Rethinking working memory and sentence comprehension deficits in developmental language disorder. *Psychological Review*, *Advance online publication.* http://dx.doi.org/10.1037/rev0000338

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 843–847. https://doi.org/10.1037/a0029284

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291. https://doi.org/10.1037/a0028228

Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*(2), 269–288. https://doi.org/10.1044/1092-4388(2008/07-0116)

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A. M., & Burton, R. (2014). Variability in the language input to children enhances learning in a treatment context. *American Journal of Speech-Language Pathology*, *23*(4), 530–545. https://doi.org/10.1044/2014_AJSLP-13-0038

Rinaldi, S., Caselli, M. C., Cofelice, V., D’Amico, S., De Cagno, A. G., Della Corte, G., Di Martino, M. V., Di Costanzo, B., Levorato, M. C., Penge, R., Rossetto, T., Sansavini, A., Vecchi, S., & Zoccolotti, P. (2021). Efficacy of the Treatment of Developmental Language Disorder: A Systematic Review. *Brain Sciences*, *11*(3), 407. https://doi.org/10.3390/brainsci11030407

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, *32*(25), 8443–8453. https://doi.org/10.1523/JNEUROSCI.5069-11.2012

Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a Working Memory Training Program for Inattention in Daily Life: A Systematic Review and Meta-Analysis. *PLOS ONE*, *10*(3), e0119522. https://doi.org/10.1371/journal.pone.0119522