Is prediction necessary to understand language?

Falk Huettig
(MPI Nijmegen, Netherlands)

LANGUAGE, COGNITION AND NEUROSCIENCE, 2016
VOL. 31, NO. 1, 19–31
http://dx.doi.org/10.1080/23273798.2015.1072223

Routledge
Taylor & Francis Group

Is prediction necessary to understand language? Probably not

Falk Huettig[a,b] and Nivedita Mani[c]

[a]Max Planck Institute for Psycholinguistics, P.O. Box 310, AH, Nijmegen 6500, The Netherlands; [b]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands; [c]Psychology of Language Group, University of Göttingen, Göttingen, Germany

ABSTRACT

Some recent theoretical accounts in the cognitive sciences suggest that prediction is *necessary* to understand language. Here we evaluate this proposal. We consider arguments that prediction provides a unified theoretical principle of the human mind and that it pervades cortical function. We discuss whether evidence of human abilities to detect statistical regularities is necessarily evidence for predictive processing and evaluate suggestions that prediction is necessary for language learning. We point out that not all language users appear to predict language and that suboptimal input makes prediction often very challenging. Prediction, moreover, is strongly context-dependent and impeded by resource limitations. We also argue that it may be problematic that most experimental evidence for predictive language processing comes from "prediction-encouraging" experimental set-ups. We conclude that languages can be learned and understood in the absence of prediction. Claims that all language processing is predictive in nature are premature.

ARTICLE HISTORY
Received 15 December 2014
Accepted 7 July 2015

KEYWORDS
Cognition; language processing; learning; prediction; predictive coding

Nivi Mani

Prediction in language processing

- there has been a wealth of research on the importance of prediction for language comprehension (e.g., Altmann & Mirkovich, 2009; Dell & Chang, 2014; Federmeier, 2007; Huettig, 2015; Kutas, DeLong, & Smith, 2011; Pickering & Garrod, 2007, 2013)

- many researchers explicitly or implicitly appear to support the notion that prediction is necessary to understand language (in line with recent proposals that prediction is an or *the* fundamental principle of human information processing, e.g., Clark, 2013; Friston, 2010)

Prediction in language processing

- we are in favor of an intermediate view: prediction contributes to understanding in many situations because it provides a 'helping hand' for dealing with specific situations

- language understanding, we conjecture, however, does *not* always involve prediction and as such is not necessary for language processing

- languages can be learnt and understood in absence of prediction

Prediction in language processing

BRAIN RESEARCH 1626 (2015) 118–135

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Review

Four central questions about prediction in language processing

Falk Huettig[a,b,*]

[a]Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[b]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Potential arguments that prediction is necessary for language processing

1.1 Prediction provides a unified theoretical framework for the cognitive sciences

Clark (2013):“brains … are essentially prediction machines” ... prediction “offers a distinctive account of neural representation, neural computation, and the representation relation itself" and a "deeply unified account of perception, cognition, and action”.

Do we really need a deeply unified principle underlying all functioning of the human mind?
- Occam's razor?

- there can be no grand principle of brain function because a complex organ such as the brain almost certainly uses a diverse set of principles (Anderson & Chemero, 2013)

Potential arguments that prediction is necessary for language processing

1.1 Prediction provides a unified theoretical framework for the cognitive sciences

- young children in particular, often focus on extending competences and engage in learning by exploration rather than prediction (Sloman, 2013)

- classic effects in the attention literature (e.g. Carrasco et al., 2004) for which a predictive framework makes either false predictions or offers no explanation (Block & Siegel, 2013)

- Clark's unified framework lacks too many implementational details and architectural commitments to be evaluated seriously

→ general framework remains to be tested thoroughly (theoretically as well as empirically) and is currently too underspecified

Potential arguments that prediction is necessary for language processing

1.2 Prediction pervades cortical function

- brain is fundamentally engaged in predictive coding and computes prediction errors which are assumed to bias our minds towards making correct inferences (Friston, 2010)

- involves the minimizing of prediction error through recurrent or reciprocal interactions among levels of cortical hierarchy

- higher hierarchical levels are thought to create forward models of lower level (cortical or subcortical) activity

- lower level activity is assumed to only contain the prediction error (often called the 'surprisal', i.e., the extent to which the predictions are disconfirmed) between 'predicted' activity and actual activity at lower levels

- prediction error is supposed to be used to update the forward models of lower level cortical activity

Potential arguments that prediction is necessary for language processing

1.2 Prediction pervades cortical function

- one interesting proposal is that oscillatory activity during language processing provides a measure of such predictive coding

- Friston et al. (2015) suggest that alpha and beta oscillatory activity reflects the forward models of lower level (cortical or subcortical) activity (i.e. the predictions) whereas gamma oscillatory activity indicates processing of prediction errors to update the predictions (see also Bressler & Richter, 2015; Engel & Fries, 2010)

- observed oscillatory activity is inconsistent, some studies found higher power in the gamma frequency range for highly predictable words than for semantically anomalous words (e.g., Hald et al., 2006; Penolazzi et al., 2009), others have found higher gamma power for world knowledge *violations* and no increase in gamma oscillations for semantically correct sentences (Hagoort et al. 2004)

→ currently available experimental evidence does not provide particularly strong support that prediction pervades cortical function at least as far as language processing is concerned

Potential arguments that prediction is necessary for language processing

1.3 Humans are adept in detecting sequential statistical regularities in language input

- connectionist approaches to structure extraction have provided compelling accounts that language learners are skillful in detecting statistical relationships in language input (e.g. Elman et al., 1990)

- even very young language learners are skillful in detecting statistical relationships in the input (e.g., Saffran et al., 1996)

→ results could also be interpreted as indexing the ease of infants' recognition of frequently co-occuring syllables, independent of any prediction-based processing

Potential arguments that prediction is necessary for language processing

1.3 Humans are adept in detecting sequential statistical regularities in language input

- performance in a statistical learning task correlates positively with sensitivity to word predictability when perceiving degraded spoken sentences (Conway et al., 2010; see also Misyak et al. 2010)

→ correlational evidence that individuals who are good at detecting statistical relationships in implicit learning tasks are also good at predicting language input, but no *direct* experimental evidence available that unequivocally links the detection of sequential statistical regularities to mechanisms of predictive language processing

- see also Tremblay et al. (2013): random input can lead to the formation of better representations of items than regular input

Potential arguments that prediction is necessary for language processing

1.4 Without prediction there would be no learning

- the fact that prediction may play an important role in language learning does not necessitate that language learning *always* involves prediction

- "prediction is not the major goal of the language learner" (Elman, 1990, p.193)

Chang, Kidd, & Rowland (2013): prediction in language processing is a by-product of language learning

- syntactic structure is learned because the learner's syntactic representations are gradually adjusted in order to be able to predict sentences

- structural priming in adults occurs because these error-based learning mechanisms stay on in proficient adult language users

- prediction in adult language processing, according to this view, is a consequence of language learning

Potential arguments that prediction is necessary for language processing

1.4 Without prediction there would be no learning

but

- infants (Pelucchi et al. 2009) and adults (Perruchet & Desaulty, 2008) track backward statistics in fluent speech

- backward transitional probabilities often are more informative than forward statistics

→ a clear example of how language learning can take place in the *absence* of prediction since backward transitional probabilities cannot be used for prediction

→ no study conducted so far has *directly* tested whether children can learn new words/grammars without prediction

Potential arguments that prediction is necessary for language processing

1.5 There is a wealth of experimental evidence that people predict in language processing

but

- most of this evidence for prediction however is not relevant for answering the question about the *precise importance* of prediction for language understanding

- vast majority of studies on predictive language processing have used sentences in which the target word was extremely predictable, i.e., very high cloze probability sentences

→ further research with low cloze probability items is required to answer the question of whether prediction is *necessary* to understand language

Arguments in line with the notion that prediction provides 'a helping hand' but is not necessary for language processing

2.1 Not everybody predicts

- variation in the amount of prediction of course does not necessarily mean absence of prediction

but:
often complete lack of evidence for any prediction in some participants/populations: Mani & Huettig (2012) find that children with low productive vocabulary scores do not fixate a related target image *cake* in a strongly predictive context, e.g., “The boy eats the…”

→ a wide range of participants who show either reduced or no anticipation of upcoming language input, but who are, nevertheless, competent language users

→ suggests that while prediction may be important to language comprehension, it does not always involve prediction

Arguments in line with the notion that prediction provides 'a helping hand' but is not necessary for language processing

2.2 Suboptimal input makes prediction less (rather than more) likely

- prediction is a powerful tool that listeners can use especially when required to compensate for noisy input (e.g. Pickering & Garrod, 2007)

- increased top-down semantic influences in the interpretation of implausible sentences in noise (e.g., Gibson, Bergen & Piantadosi, 2013)

but:

- noisy or reduced speech input often makes no difference or prediction even less likely (Mitterer & Russell, 2013; Brouwer, Mitterer, & Huettig 2013)

→ when listeners are exposed to casual speech containing many phonological reductions they may often be unable to predict because they are more uncertain what they have just heard

→ prediction can be very challenging if the input on which to base predictions is poor

Arguments in line with the notion that prediction provides 'a helping hand' but is not necessary for language processing

2.3 Prediction is strongly context-dependent

Huettig & Guerra (in prep.): visual world experiment with short or extensive visual preview and sentences presented either in a slow or a normal speech rate

- slow speech resulted in prediction in all experiments but a normal speech rate only afforded prediction if participants had an extensive preview of the visual referents

→ prediction is an important aspect but not a necessary characteristic of language processing

Arguments in line with the notion that prediction provides 'a helping hand' but is not necessary for language processing

2.4 Prediction is (frequently) impeded by resource limitations

- Christiansen and Chater (in press) have recently argued that processing speech input is severely limited, resulting in a “Now or Never” bottleneck: "only an incremental, predictive language system ... can deal with the onslaught of linguistic input, in the face of severe memory constraints of the now-or-never bottleneck".

but:

- an incremental predictive system also imposes important constraints and limits on prediction in language processing

- memory constraints and sheer speed of incoming input mean that often there is simply not enough time or not enough resources available for prediction to occur (cf. Federmeier et al., 2010; Huettig & Janse, 2016)

→ further research is needed to assess the extent to which prediction in language processing is impeded by resource limitations

Arguments in line with the notion that prediction provides 'a helping hand' but is not necessary for language processing

2.5 Much experimental evidence comes from 'prediction-encouraging experimental set-ups'

Visual world experiments:

- visual stimuli presented in eye-tracking experiments on prediction may provide critical scaffolding for the finding of such effects

- visual stimuli pre-activate spoken words (McQueen & Huettig, 2014)

EEG:

- most electrophysiological present sentences word by word in a (often slow) manner far removed from normal reading situations

- many studies measure the electrophysiological sign of anticipation (e.g., a reduced N400 ERP component) during the target word only (and not before): integration or prediction?

The way forward

- more focus on understanding why prediction effects are not found in some studies

- if prediction effects are not found in certain populations, to what extent do these populations also suffer from impoverished language skills or general cognitive skills

- if prediction effects are scaffolded by certain tasks, or certain kinds of stimuli or working memory demands, then to what extent is such scaffolding provided in natural conversation

- how does language processing in natural conversation proceed without such scaffolding (and consequently without predictive processing)

- if research continues to suggest that prediction is necessary for language processing, e.g., with regard to language acquisition, or the learning of statistical regularities, it is critical that this work more accurately outlines the precise contribution of prediction to these processes

Conclusions

- there are significant constraints for claims that prediction is necessary for language understanding

- claims that all language processing is predictive in nature are premature

- sometimes, processing words when they occur may be more efficient and economical than predicting them