Bilingualism: Language and Cognition 20 (4), 2017, 683–684 © Cambridge University Press 2016 doi:10.1017/S1366728916001061

L2 processing as noisy channel language comprehension

RICHARD FUTRELL
Department of Brain & Cognitive Sciences, MIT, Cambridge
EDWARD GIBSON
Department of Brain & Cognitive Sciences, MIT, Cambridge

(Received: August 23, 2016; final revision received: August 26, 2016; accepted: August 26, 2016; first published online 22 September 2016)

The thesis in this paper is that L2 speakers differ from L1 speakers in their ability to do memory storage and retrieval about linguistic structure. We would like to suggest it is possible to go farther than this thesis and develop a computational-level theory which explains why this mechanistic difference between L2 and L1 speakers exists. For this purpose, we believe a noisy channel model (Shannon, 1948; Levy, 2008; Levy, Bicknell, Slattery & Rayner, 2009; Gibson, Bergen & Piantadosi, 2013) could be a good start. Under the reasonable assumption that L2 speakers have a less precise probabilistic representation of the syntax of their L2 language than L1 speakers do, the noisy channel model straightforwardly predicts that L2 comprehenders will depend more on world knowledge and discourse factors when interpreting and recalling utterances (cf. Gibson, Sandberg, Fedorenko, Bergen & Kiran, 2015, for this assumption applied to language processing for persons with aphasia). Also, under the assumption that L2 speakers assume a higher error rate than L1 speakers do, the noisy channel model predicts that they will be more affected by alternative parses which are not directly compatible with the form of an utterance.

A noisy channel model of language comprehension basically posits that, when a comprehender is perceiving or remembering an utterance, she does error detection and correction on the utterance. More precisely, the comprehender considers that the utterance as she perceives it may have been affected by some noise process: maybe the speaker made speech errors; maybe she misheard the words; maybe when she is remembering the utterance as stored in memory, she is remembering incorrectly. She then tries to correct possible errors in the perceived utterance according to what is most likely in terms of syntactic probabilities, what is most likely in terms of discourse factors such as plausibility, and what she thinks is the probability that an error occurred (the noise rate). For example, upon hearing (or remembering) the utterance "the mother gave the candle the daughter", a rational comprehender might conclude that the utterance as perceived was just a corrupted version of "the mother gave the candle to the daughter" – in that case the comprehender would interpret the utterance nonliterally (Gibson et al., 2013). The noisy channel model is easily formalized mathematically and has enjoyed wide applicability and deep study in fields other than natural language, such as artificial intelligence and electrical engineering (Shannon, 1948; Levy, 2008).

The noisy channel model might offer a high-level explanation for some of the findings which are explained in this paper in more mechanistic terms. The hypothesis would be that L2 speakers have different models of language than L1 speakers, or they might assume different noise rates than L1 speakers do. On the other hand, it seems reasonable to hypothesize that they have the same or similar discourse and semantic knowledge as L1 speakers.

A major set of findings discussed in this paper suggests that L2 speakers rely more on discourse cues such as topicality when resolving anaphora than on linguistic cues. In a noisy channel model, the comprehender considers the probability of the utterance according to syntactic knowledge and according to semantic and discourse knowledge, and does error correction accordingly. The theory predicts non-literal interpretations of utterances when the veridical form of the utterance has low probability under either kind of knowledge. Suppose that L2 speakers have a less precise probability model of the syntax of the relevant language than L1 speakers do, but their knowledge of semantic and discourse factors is roughly the same. For example, L2 speakers may assign probability .2 to a syntactic structure to which L1 speakers assign probability .01, and may assign probability .8 to a syntactic structure to which L1 speakers assign probability .99. These probabilities would reflect an L2 speaker's increased uncertainty about the syntax of the language. Then, when interpreting utterances, the L2 speakers will rely on non-syntactic knowledge more than L1 speakers do: if the utterance is syntactically low probability, then the L2 speaker is less likely to correct it, and if it is syntactically high probability, the L2 speaker is less likely to maintain it. Thus syntactic probability has less effect on comprehension for an L2 speaker, so other sources of knowledge, such as

Address for correspondence:
Richard Futrell, Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences, 43 Vassar Street, Room 46-3037, Cambridge Massachusetts, United States 02139
futrell@mit.edu

https://doi.org/10.1017/S1366728916001061 Published online by Cambridge University Press

discourse and semantic factors, will have a proportionally stronger effect. Rather than positing that the fundamental difference between L2 speakers and L1 speakers is their retrieval ability, the noisy channel model locates the causally relevant difference in the probabilistic knowledge of language, which then affects retrieval accuracy.

Cunnings also discusses results (such as Jacob & Felser, 2016) showing that L1 and L2 speakers are influenced by incorrect initial parses of sentences, with L2 speakers showing more influence; a noisy-channel model seems particularly attractive in these cases. It could be that, as readers are reading word-by-word, they have a noisy representation of the previous input in memory, which they attempt to correct using their knowledge sources. For instance, Cunnings discusses an example from Christianson, Hollingworth, Halliwell & Ferreira (2001): English speakers sometimes interpret the sentence "While Anna dressed the baby that was small and cute spit up in the bed" as meaning that Anna dressed the baby. It is possible that, by the end of the sentence, the readers misremember it as "While Anna dressed the baby that was small and cute(, it) spit up in the bed"; if the latter version of the sentence is more probable syntactically and semantically, then the readers might even rationally believe that the writer made a mistake and omitted "it". The noisy channel model for very similar cases is worked out and supported with reading time data in Levy (2011). In this case, the result that L2 users are more prone to interference from garden path parses would follow naturally under the assumption that L2 speakers assume a higher noise rate in their input than L1 speakers do, making them more likely to do corrections on their input.

References

Christianson, K., Hollingworth, A., Halliwell, J. & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42,* 368–407.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, 110*(*20*), 8051–8056.

Gibson, E., Sandberg, C., Fedorenko, E., Bergen, L., & Kiran, S. (2015). A rational inference approach to aphasic language comprehension. *Aphasiology,* DOI: 10.1080/02687038.2015.1111994

Jacob, G. & Felser, C. (2016). Reanalysis and semantic persistence in native and nonnative garden-path recovery. *Quarterly Journal of Experimental Psychology 69*(*5*), 907–925.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. *Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing* (Association for Computational Linguistics, Stroudsburg, PA), pp. 234–243.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, 106*(*50*), 21086–21090.

Levy, R. (2011). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. In *Proceedings of the 49th annual meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 1055–1065).

Shannon, C. (1948). A mathematical theory of communication. *Bell Systems Technical Journal, 27,* 623–656.

https://doi.org/10.1017/S1366728916001061 Published online by Cambridge University Press