Bilingualism: Language and Cognition 20 (4), 2017, 700–701 © Cambridge University Press 2016 doi:10.1017/S1366728916000912

L2 Proficiency matters in comparative L1/L2 processing research

GREGORY D. KEATING
San Diego State University

(Received: August 11, 2016; final revision received: August 19, 2016; accepted: August 19, 2016; first published online 26 September 2016)

Grammatical processing in a second language (L2) that is learned in adulthood has been shown to differ from processing in a first/native language (L1). Clahsen and Felser's (2006) landmark article provided the first comprehensive account of these differences. According to their shallow structure hypothesis (SSH), L2 learners, unlike L1 speakers, do not compute abstract, hierarchical representations during online sentence comprehension; instead, they rely on lexical, semantic, and pragmatic information to build 'good enough' representations. However, native-like processing is attainable – with sufficient L2 proficiency – for word-level processing and morphosyntactic feature processing between locally related words. Clahsen and Felser's article spurred a prolific volume of research over the last decade. Some findings support the SSH, whereas others favor the competing claim that L1/L2 differences result from capacity-based limitations.

In this keynote article, Cunnings (2016) provides the first rigorous attempt to critically review this body of work and to offer alternate explanations for the disparate findings. His account includes three claims about L2 learners' grammatical processing:

1. L2 learners construct fully-specified syntactic parses, contra the SSH;
2. L1/L2 differences that persist at high levels of proficiency result from L2 learners' increased susceptibility to interference during memory retrieval; and
3. L2 learners rely more heavily on discourse-based cues to memory retrieval than L1 speakers.

Cunnings' proposal promises to push L2 processing research in new directions. First, it introduces a model of working memory not previously considered in L2 processing research. Second, by characterizing L1/L2 differences in terms of interference during memory retrieval, it positions L2 processing research to contribute to discussions of the human sentence comprehension mechanism taking place in related fields where interference is a primary construct of interest, such as mainstream psycholinguistics (e.g., Van Dyke & Johns, 2012) and communicative disorders (e.g., Sheppard, Walenski, Love & Shapiro, 2015). The primary weakness of the article is that most of its claims are supported by data from existing L2 processing studies that were not designed to test for interference. However, I expect this article to motivate the research needed to investigate the claims directly – in fact, I'm already designing a study to test Cunnings' claim that interference likely affected the processing of gender agreement in my previous work (Keating, 2010)!

Like the SSH, Cunnings' proposal attempts to account for L1/L2 processing differences that "persist at high levels of proficiency." As our field conducts new research to further this endeavor, it must correct an unsound yet widespread methodological practice: using data from non-advanced learners as evidence for persistent L1/L2 processing differences. For example, many of the studies taken as support for shallow L2 processing tested learner groups that mixed upper intermediate and advanced learners (as determined by Oxford Placement Tests). The same is true of many of the studies that Cunnings cites in support of his claims. As one example, Cunnings draws on the findings of Pan and Felser (2011) and Pan, Schimke, and Felser (2015) to support the claim that L2 learners may rely more heavily on discourse-based cues to memory retrieval than L1 speakers. However, both studies included L2 learners who scored 5.5 on the IELTS, which indicates partial command of the language. The performance of non-advanced learners is unsuitable for making, much less testing, claims about L1/L2 differences at high proficiency levels. This limitation is not unique to the studies cited in Cunnings' article; it applies to many, if not most, published L2 processing studies. Why is this problem so pervasive?

Participant availability aside, I believe the problem stems, in part, from data collection practices. In most L2 processing studies, potentially eligible participants are recruited into one testing session during which

Address for correspondence:
Gregory D. Keating, Department of Linguistics and Asian/Middle Eastern Languages, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182
gkeating@mail.sdsu.edu

https://doi.org/10.1017/S1366728916000912 Published online by Cambridge University Press

they complete all tasks – background questionnaires, proficiency tests, and online processing tasks. In this approach, a participant's proficiency isn't usually determined until after all tasks are completed, at which point researchers may be more reluctant to discard otherwise good data sets that took significant time to collect. In lieu of investing more time to recruit additional participants – who may turn out to be non-advanced – researchers proceed to include the non-advanced learners with the advanced ones and muddy the data. This practice has consequences: heterogeneous samples muddy the conclusions we make about the native-likeness of L2 processing and about which theoretical accounts best explain the observed phenomena.

In our lab, we've begun to address this problem by conducting multi-session studies. In the first session, participants complete background questionnaires and proficiency tests. Only those participants that meet predetermined background and proficiency criteria get invited to participate in subsequent sessions to complete the online tasks of interest. We are still assessing the outcomes of this approach, but we believe the strategy is yielding more homogeneous participant groups while also reducing overall data collection time and participant fees.

Research on the effects of interference in L2 processing has the potential to advance our field, but to do so it must also do what much existing L2 processing research has not: restrict participant inclusion to homogeneously advanced L2 learners.

References

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Cunnings, I. (2016). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*. doi:10.1017/S1366728916000675, Published online by Cambridge University Press, June 20, 2016.

Keating, G. D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. VanPatten & J. Jegerski (Eds.), *Research in second language processing and parsing* (pp. 113–134). Amsterdam/Philadelphia: John Benjamins.

Pan, H., & Felser, C. (2011). Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading. *Lingua, 121*, 221–236.

Pan, H., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism, 19*, 298–313.

Sheppard, S. M., Walenski, M., Love, T., & Shapiro, L. P. (2015). The auditory comprehension of *Wh*-questions in aphasia: support for the intervener hypothesis. *Journal of Speech, Language, and Hearing Research, 58*, 781–797.

Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass, 6*, 193–211.

https://doi.org/10.1017/S1366728916000912 Published online by Cambridge University Press