Quantifying individual differences in native and non-native sentence processing: Evidence from garden-paths

Ian Cunnings & Hiroki Fujita (University of Reading)
i.cunnings@reading.ac.uk

The role that individual differences may play in theoretical debate surrounding language acquisition and processing has been widely debated (for review, see [6]). This is particularly evident in research comparing native (L1) and non-native (L2) sentence processing, where debate has questioned whether individual differences influence how nativelike L2 processing can become [2,5]. However, despite increased interest in individual differences, the question of whether commonly used psycholinguistic tasks consistently measure individual differences has received little empirical attention. This is unfortunate given that robust cognitive tasks often make poor individual differences measures, a finding dubbed 'the reliability paradox' [3]. Whether reading tasks reliably measure individual differences in L1 comprehension has only recently begun to be assessed, and initial findings suggest tasks like self-paced reading may not consistently measure individual differences [7]. We aimed to examine this issue from the perspective of L1/L2 processing, using garden-path effects as a test-case to examine the quantifiability of individual differences in L1 and L2 processing.

We conducted a web-based self-paced reading experiment in Ibexfarm examining the co-ordination ambiguity [4]. Participants were 64 L1 English speakers, who self-identified English as their only L1, and 64 L2 English speakers from different L1 backgrounds, who self-identified a language other than English as their L1. Participants read temporarily ambiguous sentences like (1a) and unambiguous controls (1b). To test offline comprehension, participants answered questions that tapped whether the initially assigned misinterpretation ("Ken washed the cat and the dog") lingers following disambiguation in garden-path sentences [1].

Analysis using mixed-effects models of log-transformed reading times yielded a significant ambiguity effect at the disambiguating verb ("played"; estimate = 0.04, SE = 0.01, t = 4.79, p < .001), with longer reading times for ambiguous than unambiguous sentences. There was also a significant ambiguity effect in comprehension accuracy (estimate = -1.04, SE = 0.11, t = -9.26, p < .001), with lower accuracy following ambiguous than unambiguous sentences, and a significant interaction between ambiguity and group (estimate = -0.28, SE = 0.08, t = -3.54, p < .001), such that the difference between unambiguous and ambiguous sentences was larger for L2 than L1 readers (see [9]). To assess how consistently self-paced reading measures individual differences, we calculated the split-half reliabilities for the ambiguity effect for L1 and L2 readers separately in both reading times and comprehension accuracy, using the method described by [7]. We randomly split the data in half, fit separate mixed-effects models to each half of the data, and correlated the extracted by-subject random effects from the two models fit to each half of the data. We repeated this process 100 times and then calculated the mean split-half correlation. These analyses revealed high split-half correlations for by-subject intercepts in reading (r > .900) and comprehension accuracy (r > .750) in both groups, indicating the task consistently ranked participants as either slower or faster (or more/less accurate) readers overall. However, split-half correlations for the ambiguity effect were all low, for both reading times (r < .200) and comprehension accuracy (r < .420).

Together, these results suggest that while self-paced reading may consistently measure individual differences in overall reading speed and comprehension accuracy, it does not consistently measure individual differences in the size of garden-path effects. The offline comprehension measure may provide a slightly better index of individual differences in garden-path effects than online reading, though this still did not reach established thresholds to be considered a consistent measure (r > .700, [8]). These results are similar to [7], who conducted their study in a typical lab setting with L1 participants only. Together, our results suggest there is either little consistent variability in the size of garden-path effects between individuals to begin with, or that the task is unable to consistently measure them. Either way, before drawing any strong theoretical conclusions about the role of individual differences in language acquisition and processing, these findings highlight the need to develop tasks and materials that can consistently tap individual differences in the first place.

(1) Example experimental stimuli (n = 24, plus 56 fillers)

(a) Ambiguous
Ken washed the dog and the cat in the garden played with a ball.

(b) Unambiguous
Ken washed the dog while the cat in the garden played with a ball.

Question: Did Ken wash the cat?

Figure 1. *Reading times at the disambiguating verb (left panel) and comprehension accuracy (right panel).*

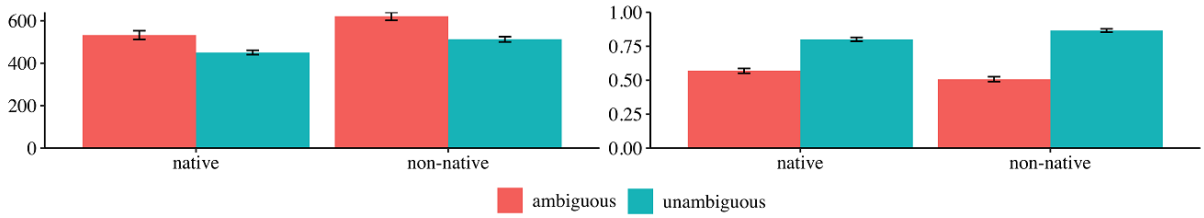
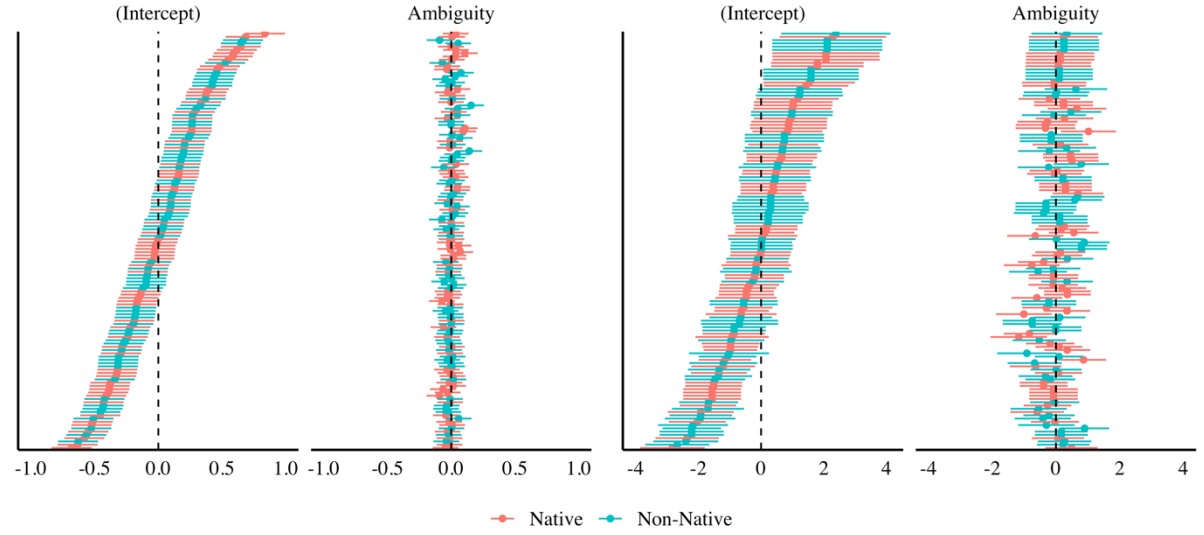


Figure 2. *By-subject random effects at the disambiguating verb (left panel) and in comprehension accuracy (right panel), illustrating consistent individual differences in overall reading speed and comprehension accuracy (intercepts) but not ambiguity.*



References

[1] Christianson et al. (2001). *Cognitive Psychology*, *42*, 368–407.
[2] Clahsen & Felser (2018). *Studies in Second Language Acquisition*, *40*, 693–706.
[3] Hedge et al. (2018). *Behavior Research Methods*, *50*, 1166–1186.
[4] Hoeks et al. (2002). *Journal of Memory and Language*, *46*, 99–119.
[5] Hopp (2018). *Second Language*, *17*, 5–27.
[6] Kidd et al. (2018). *Trends in Cognitive Sciences*, *22*, 154–169.
[7] James et al. (2018). *Journal of Memory and Language*, *102*, 155–181.
[8] Nunnally (1978). *Psychometric theory* (2nd ed.).
[9] Pozzan & Trueswell (2016). *Bilingualism: Language and Cognition*, *19*, 636–643.