We need an integrated, multiple-predictor model of native language proficiency

Irina Sekerina
College of Staten Island, The City University of New York

In the keynote article, Dąbrowska reviews studies that she and her colleagues have conducted in recent years in support of the idea that there are different levels of proficiency in native language in monolingual adults. In the first section of the article, Dąbrowska reports the results from two Polish experiments on production of grammatical case morphology; in the second section, she turns to comprehension studies of syntactic constructions of different complexity in English. In all of them, Dąbrowska and colleagues did not find any differences between less-educated and highly educated participants for control items. However, less-educated speakers, both Polish and English, were significantly worse (and even below chance in some cases) in their accuracy of producing or comprehending the experimental items. In the final section of the article, Dąbrowska proposes a hypothesis of individual variation in knowledge of native language and argues that is caused by education-related differences, broadly defined.

I applaud Dąbrowska’s bold take on a topic that is considered to be politically sensitive in the U.S., especially in the context of education of minority children (Ramirez et al., 2005), and agree that it is time to shift attention in psycholinguistics from performance averaged across population groups (e.g., monolingual adults, children, second language learners, bilingual speakers, etc.) to individual differences, capacities, and deficits. However, while the contribution of education-related differences to native language proficiency is intriguing, it cannot be the whole story. I suggest that a focus on education as the main predictor of individual variation in native language is insufficient from two perspectives, theoretical and methodological.

Theory: The ‘third variable’ problem

One of alternative explanations that Dąbrowska considers and dismisses as not feasible is the idea that less-educated native speakers may have full competence of
We need an integrated, multiple-predictor model of native language proficiency

the grammar, but are unable to access it because of working memory (WM) limitations. Her arguments are based on two findings from her studies: (a) less-educated participants are significantly less accurate and slower than highly educated ones, even for simple syntactic constructions like passives that arguably do not impose a processing overload on WM, and (b) training on passives and sentences with quantifiers dramatically improves less-educated speakers’ performance on them. It should be noted that Dąbrowska does not provide any information on whether her less-educated and highly educated participants in any of the experiments were matched on their WM capacity, or even whether their WM was measured at all. All of them were “with no known learning difficulties”, but this is all that is reported about the cognitive abilities of the participants; thus, differences in WM cannot be ruled out as a plausible explanation.

In addition, WM is not a unitary construct; it is a multi-component system, verbal and non-verbal, that is simultaneously engaged in processing and storage of information (Conway et al., 2005), and supplies mental resources for executive abilities (Connor, MacKay, & White, 2000) and language processing (Swets et al., 2007). It is closely related to the capacity to allocate attention and speed of processing (Bayliss et al., 2003). Thus, WM, attention, and speed of processing constitute the pool of cognitive resources that contribute to individual differences in how people process linguistic tasks, such as the production of inflection markers and the comprehension of syntactic constructions. In the area of neuropsychology, this is known as cognitive reserve, i.e., efficiency, capacity, and flexibility brought about by individual differences in the networks and in cognition (Stern, 2009). People with larger or stronger cognitive reserve cope better with brain language-related pathology in the brain such as aphasia, show slower rates of age-related language decline (Obler & Pekkala, 2008), and enjoy a longer protection against the onset of Alzheimer’s disease (Mortimer, Snowdon, & Markesbery, 2003). Education level may be no more than a covariate of the cognitive reserve, rather than a predictor of individual variation in knowledge of native language. In experiments with native speakers, we need to include both factors, a composite measure of cognitive reserve and education level, and enter them as predictors into statistical models to assess their interaction.

Methodology: Accuracy measures and statistical modeling

Most of the English experiments reported by Dąbrowska measured participants’ performance scores, or offline accuracy, on production of correct case markers or on comprehension of syntactic structures. If the argument about education-related effects on the grammar is to be upheld, the performance pattern of less-educated participants that indicates their difficulty with comprehension must be
task-independent. Dąbrowska's less-educated participants were tested only on metalinguistic tasks, e.g., a force-choice picture-matching selection task in which they had to select one of the two pictures that went with the passives and sentences with quantifiers. However, a recent language acquisition study (O’Grady, Suguzi, & Yoshinaga, 2010) has shown that children’s accuracy of comprehension of sentences with the quantifier *every* paired with pictures depicting a partial many-to-one correspondence between two entities, fish and bowls (Dąbrowska’s examples (6)–(7)), critically depended on the task. The same children were tested twice, and their errors with quantifiers virtually disappeared (from 3.73 to 0.82 incorrect items out of 4) when they were switched from the sentence-picture matching task to an act-out task. We need to empirically demonstrate that the difficulty that less-educated native speakers exhibit in interpreting syntactic constructions remains in the act-out task and in looking-while-listening, a version of the Visual World eye-tracking paradigm employed with young children (Fernald et al., 2008).

Recording eye movements could be especially illuminating in our search of underlying causes of errors participants make in offline comprehension of sentences because they reveal online attention allocation strategies. Dąbrowska herself alludes to the role of attention in the article, but does not pursue this issue. In our recent eye-tracking study, children who looked longer at the extra fish while listening to the experimental sentences like *Every fish is in a bowl*, made significantly more errors in sentence-picture matching than children who did not (Sekerina & Sauermann, in prep.). Moreover, we found a similar pattern of looks in our control group of highly educated monolingual adults, students at the elite university in Moscow, although the low error rate of adults compared to children (4% versus 30%) was not sufficient enough for inferential statistical analysis. Thus, it is essential to establish that less-educated speakers consistently differ from highly educated ones across tasks and methodologies and that the latter retain their superior performance regardless of resource reduction and ‘nuisance’ variables.

Finally, Dąbrowska’s data are categorical, and were analyzed with the help of traditional statistical procedures, such as correlations and analysis of variance (ANOVA). These statistical tests are being criticized (Jaeger, 2008) because they can easily yield spurious significant differences in performance scores of groups of participants based on two levels of education, as in Dąbrowska’s case. The current trend in experimental psycholinguistics is to switch to advanced statistical modeling, i.e., logistic regressions with mixed effects that will include multiple predictors and their interaction as parameters.

To summarize, as important as it is to investigate education effects on native language proficiency, Dąbrowska recognizes that it is just one step in discovering complex, interactive factors that contribute individual variation in the grammar of monolingual native speakers. Our ultimate goal is to delineate a set of multiple
We need an integrated, multiple-predictor model of native language proficiency — neural, cognitive, and “personal” — for successful language and communication performance in native speakers that is supported by comprehensive empirical testing, wide range of methodology, and rigorous statistical modeling.

References


Author’s address

Irina Sekerina
College of Staten Island
Department of Psychology, 4S-102
2800 Victory Blvd.
Staten Island, NY 10314
USA
Irina.Sekerina@csi.cuny.edu

All rights reserved