Explaining individual differences in linguistic proficiency

Ewa Dąbrowska

INTRODUCTION

The vast majority of the 14 commentaries are thoughtful and constructive and raise a variety of important points. Due to space limitations, I will not be able to respond to all the points in as much detail as they deserve. Some of the apparently critical remarks, I believe, are due to misunderstanding. I therefore begin my response by clarifying some issues. I then discuss some alternative explanations of the data suggested by the commentators, including the possibility that the results described in the keynote article are attributable to performance factors, and respond to commentators who disagree with the central claim of the keynote article, viz., that there are substantial individual differences in adult native speakers' knowledge of the grammar of their language. I conclude by discussing some broader issues and suggestions for further research.

SOME CLARIFICATIONS

1. Education-related differences are not "the whole story"

Several researchers point out that education-related differences "cannot be the whole story" (**Sekerina** this issue: XX). Sekerina also warns about the third variable problem, and **Hadley and Rispoli** point out that some of the observed differences in grammatical knowledge may be due to biological factors. On a related note, **Schulz** finds the concept of educational background "irritatingly vague", and argues that most of the differences described in the keynote article are group, rather than individual, differences, while **Reuland** points out that the low academic attainment (LAA) label applies to individuals with very different abilities.

All of these observations are fully compatible with views expressed in the keynote article. I never claimed that education was "the whole story", or even a significant causal factor. The studies described in the article used education (operationalized as the number of years of schooling) as a grouping variable. Possible causes were discussed in the section on reasons for individual differences, and include environmental factors such as the amount and quality of linguistic experience as well as learner-internal factors, including IQ, language aptitude, and need for cognition (the degree to which an individual enjoys effortful cognitive ability).

It is also worth pointing out that not all of the constructions discussed in the keynote article showed education-related differences. Those that did, however, all followed the same pattern: highly educated participants performed at ceiling; less educated participants varied in ability, with performance ranging from chance (and, in some cases, below chance) to ceiling. Thus, we are dealing with both group and individual differences.

2. Explicit instruction is not necessary for learning

Serratrice points out that explicit instruction is not necessary for learning. I did not argue that it is – only that it *may be helpful*, at least for some constructions.

What I suggested may be necessary is that the learner attend to both form and meaning at the same time – and explicit explanation may facilitate this.

Serratrice also points out that the training in the Street and Dąbrowska (2010) study involved explicit instruction as well as exposure to the trained construction in a disambiguating context, so we don't know whether the improvement in performance was a result of teaching, exposure, or a combination of the two, and she suggests that the study should have had an additional control condition: exposure to the trained form in a disambiguating context, but without explicit instruction. Until this manipulation is performed, Serratrice argues, we cannot be sure what caused learning. This is absolutely correct; however, the purpose of the study was not to investigate the role of explicit instruction in learning, but simply to determine whether learning would occur at all.

3. The existence of individual differences does not entail that the UG hypothesis is false

Phillips observes, correctly, that the existence of individual variation does not constitute an argument against Universal Grammar, and concludes that the main argument in the paper is a non-sequitur. However, I did not claim that the existence of individual differences in language attainment entails that the UG hypothesis is false: all I said is that it undermines one of the arguments for UG, namely the convergence argument. I am nonplussed at how Phillips arrived at his interpretation of my argument, since the keynote article states very clearly that

> "The results summarised here suggest that the convergence argument [for Universal Grammar] is based on a false premise: there are, in fact, considerable differences in how much speakers know about some of the basic constructions of their native language. This does not necessarily mean that Universal Grammar does not exist: one can argue in favour of innate constraints on language learning on other grounds, for instance, poverty of the stimulus." (XX)

Phillips also argues (again, correctly) that whether or not different learners are exposed to different input is irrelevant to the convergence argument: "If learners consistently reach the same conclusions based on the same input utterances …. then it suggests that there are constraints on the conclusions that they draw" (XX). Yes – *if* they consistently reach the same conclusions – but this is precisely what is at issue!

ALTERNATIVE EXPLANATIONS

Reuland suggests that the results reported in the keynote article could be accommodated by appealing to Avrutin's (2006) notion of "weak syntax", originally developed to explain the performance of Broca's aphasics. According to Avrutin, the syntactic component in Broca's aphasics is unimpaired but weak; this results in slower processing, and, consequently, patients often resort to nonsyntactic means when processing utterances. Thus, Broca's aphasics' "weak syntax" is supposed to account for not just agrammatic comprehension of structures like passives, but also their slow, effortful and telegraphic speech.

There are a number of problems with Reuland's proposal. He does not explain why syntax should be "weak" in speakers who have not suffered brain damage. Furthermore, while some LAA participants have relatively poor comprehension, their spontaneous production is not like that of aphasics. Related to this, there is no evidence of an across-the-board impairment in morphosyntactic knowledge in these speakers: they merely have problems with some specific constructions. Last but not least, it is not clear how the "weak syntax" hypothesis could accommodate the results of the training studies. Reuland suggests that training enhances automatization and hence the speakers' syntax becomes stronger. This is extremely implausible: automatization is a slow process, and is thus unlikely to have occurred during a five-minute training session involving just six exemplars of the target construction.

Reuland also suggests that the relatively poor performance observed in some participants in the Polish dative study may be explained by lack of lexical rather than morphological knowledge: the participants, he argues, may know the structure of the paradigm, but just lack the knowledge that the masculine singular dative ending is -*owi*, the neuter ending is -*u*, etc. However, such an explanation cannot account for the observed results (see Dąbrowska 2008). The Polish participants do know what the dative ending are, as they readily supply them with real words as well as some nonce words. They are also able to correctly identify the gender of nonce nouns. What they lack is the knowledge that the same ending applies across-the-board to all nouns belonging to a particular class – i.e., knowledge about the structure of the paradigm.

Vainikka and Young-Scholten suggest that the differences observed in Polish speakers' knowledge about the genitive could be attributable to differences between spoken and written language: the grammars of more educated participants may conform to the norms of written Polish, while those of the less educated participants may reflect spoken language. Two points need clarifying here. First, the differences observed in the genitive study were not education-related: all the participants were students from the same class in the same school, and thus had the same level of education and similar social backgrounds. Education-related differences *were* found for the Polish dative; these however, are not attributable to different norms for spoken and written language, because exactly the same rules operate in both varieties.

THE COMPETENCE/PERFORMANCE DISTINCTION

Several commentators observe that evidence for grammatical competence is always indirect, and emphasize the need to exercise caution when interpreting experimental results. **Sekerina** and **Phillips** point out that children who fail on tasks tapping knowledge of a particular construction sometimes succeed when tested using a different method. Adults tend to be less sensitive to task differences; however, Sekerina and Phillips are right to insist that the results described in the keynote article should be replicated using different methods. Note that this has already been done for the passive, and the results of the different studies (Dąbrowska and Street 2006, Street and Dąbrowska 2010, in press) are very similar; however, further research is clearly required with other constructions.

Schulz points out that it is important to rule out the possibility that participants didn’t understand the task, and also suggests that the poor performance on implausible sentences in the Dąbrowska and Street (2006) study may be explained by pragmatic factors. The first objection is unjustified: as explained in the article, the LAA participants performed at ceiling in control conditions. The second point is valid;

however, pragmatics cannot explain the differences in performance on implausible actives and passives. Note, too, that later studies showed similar results with reversible sentences.

Phillips makes a similar point, but draws stronger conclusions. He contends that the evidence that I present bears on participants' "type (ii) knowledge", i.e. their "skill or efficiency at constructing specific representations or interpretations" (XX), rather than "type (i) knowledge", i.e., the representations that a speaker *can* construct. In other words, he seems to be arguing that the difficulties that some participants experienced in the studies described in the article are facts about performance, and hence irrelevant to claims about linguistic competence. **Reuland** seems to hold an even more radical view when he asserts, bizarrely, that "chance behavior indicates that [speakers] know the rules but cannot always apply them" (XX3).

One entire section of the keynote article was devoted to arguing that the observed differences must be at least partly attributable to differences in competence. Phillips and Reuland are clearly not convinced by it. What they do not state in their contributions is what *would* count as evidence about "type (i) knowledge": they appear to believe that performance data can never be used to falsify claims about competence.

The problem, of course, is that – since competence cannot be tested directly – our only evidence about it comes from studies of performance. As is well known, performance can be affected by grammatically irrelevant factors such as attention or cooperativeness; but the solution is to design experiments which control for such confounds, rather than giving up the commitment to linguistics as an empirical science.

ACQUISITION BY FIAT

Several commentators disagree with the central claim of the keynote article, namely, that there are substantial differences in individual speakers' knowledge about their language. **Roeper** objects that I do not provide "a single example of an alternative grammar that speakers arrive at", or even evidence that speakers respond consistently, arguing that "if another grammar were present, we would expect consistent alternative behavior" (XX1). I must begin by pointing out that I do, in fact, provide an example of alternative grammars: for the Polish genitive singular inflection, it is clear that different speakers have different rules (some use *-a* with all masculine nouns; some use *-a* with animates and *-u* with inanimates; and some have more specific rules based on either semantic or phonological criteria). However, it is true that most of the examples discussed in the keynote article involved a different situation, namely, cases where individual speakers' grammars are arguably *incomplete*, that is to say, lack a particular rule or principle (as opposed to *divergent*, i.e., containing a different rule or principle – cf. Sorace 1993). The reason for this is that when speakers' grammars are different, this may be due to dialect differences – and it is very difficult to demonstrate that a particular variant is *not* associated with a particular language variety, be it social or stylistic. Roeper asserts that in order to demonstrate alternative grammars, I would need to demonstrate "consistent alternative behaviour". This is simply incorrect: if one speaker's grammar lacks a construction, rule or constraint that is present in another speaker's grammar, then they have different grammars. Note, too, that in L2 research, inconsistent performance is regarded as the hallmark of incomplete acquisition.

Roeper also suggests that the low academic attainment participants' performance could have been affected by "personality variables linked to social class", and that "it might be that education affects a person's 'grammaticality judgment attitude' skewing the results in one direction or another without revealing a grammatical difference" (XX1). It is not clear what relevance the last point has for the research described in the keynote article, since none of the studies actually used grammaticality judgments. But let's assume for the sake of argument that personality affects performance on inflection tasks and grammatical comprehension tasks. It is still not clear why personality variables should affect performance on passives, sentences with quantified NPs, and nouns from high-density neighbourhoods, but not actives and other constructions which were used in the control conditions. Furthermore, I fail to see how personality variables linked to social class could explain the change in performance after training observed in Street and Dąbrowska (2010) and Chipere (2001) – unless a five-minute training session can have lasting effects on personality. Most importantly, Roeper cannot simply dismiss a whole raft of studies by asserting that some unspecified personality variables might have affected performance: he needs to make some concrete and falsifiable statements about what these personality variables might be, how they affect performance, and why they affect some constructions but not on others.

Finally, **Roeper** asserts that I use some "extremely indirect and obscure correlational claims" to "argue for the superiority of one group – once again educated, wealthier people – over others" (XX). I will leave it to readers to decide for themselves whether my claims are indeed "extremely indirect and obscure"; they are certainly not purely correlational (the two training studies discussed in the paper were true experiments). And I am certainly not arguing for the superiority of the richer and more educated: as explained in the keynote article, the LAA participants were able to learn the relevant constructions very quickly when given appropriate input.

I would also like the world to be a fair place and everybody to have equal opportunities. Unfortunately, this is not the case – as demonstrated by the work on socio-economic status and brain development reviewed by **Pakulak**. Roeper's attempt to claim the moral high ground by denying the existence of the differences described in the article is not only scientifically questionable; it is also socially irresponsible, since the evidence suggests that we can do a great deal to help less privileged children (cf. Neville et al. 2011, Hackman and Farah 2009, Hackman et al. 2010).

Schulz takes a rather different tack. She begins by discussing four possible scenarios:

(1) speakers of different regional varieties have somewhat different grammars;

(2) a particular speaker may have two different grammars (e.g. for the standard and the regional variety);

(3) different speakers may have different variants of a particular rule; and

(4) some speakers may lack a particular rule.

Scenarios (1) and (2) are not relevant to the discussion, since the research described in the keynote article dealt with differences between speakers that could not be explained by appealing to dialect differences: for instance, it is not the case that *The boy was kissed by the girl* means 'The girl kissed the boy' in some dialects of English and 'The boy kissed the girl' in other dialects. Schulz acknowledges that scenarios (3) and (4) could be real, but suggests such a state of affairs would demonstrate that the relevant speakers simply have a minimally *different* grammar rather than a grammar that is impoverished or incomplete. In other words, she appears to accept the central claim of the keynote article – the existence of individual differences in speakers' knowledge of core grammatical constructions (although she claims that such differences are quite small) – but denies incomplete acquisition, arguing that a grammar that lacks a particular rule is "equally complete" (XX2).

Can a grammar that lacks a particular rule be regarded as "equally complete" and merely simpler, and possibly more elegant? There is a sense in which this is a reasonable suggestion: a two-bedroom house is a complete house with two bedrooms – not an incomplete version of a three-bedroom house. On the other hand, there is a real sense in which a grammar that does not allow speakers to correctly interpret passive sentences or simple sentences with quantifiers can be regarded as incomplete. A second language learner who has not mastered the passive construction would be uncontroversially regarded as having incomplete mastery of English grammar. There is no principled reason for applying different criteria to native speakers.

BROADER ISSUES

Several commentators point out that it is important to spell out what exactly the low-performing participants do not have. For instance, with respect to the results on the comprehension of passives, **Reuland** makes the following observation:

> "There is no such thing as a 'passive rule'. The formation of passives in English involves at least three independent rules/processes: case suppression for the object, suppression of the external thematic role, and movement of the object into the subject position…. Hence, any claim about non-acquisition of passives should

be specified as a claim about a specific subprocess involved in passives, and be tested against the consequences it predicts." (XX1)

It is clear from this passage that Reuland confuses reality with a particular theoretical account of it. *In some versions of generative grammar*, the formation of passives involves case suppression for the object, movement, etc.; but non-derivational theories account for passives in a very different way. In Construction Grammar, for instance, the passive is considered a construction (i.e., a form-meaning pairing) in its own right, not something that is derived by transforming a structure with an active-like word order.

Different theories will make different predictions about what other constructions may be affected in individuals who have not mastered the passive, and Reuland is absolutely right that these predictions need to be spelled out explicitly. But the purpose of the keynote article was not to determine whether Construction Grammar accounts for the facts better than Minimalism, or vice versa, but to establish that individual differences exist – the necessary first step to subsequent investigations of relationships between structures. For this reason, I have tried to describe the results in a maximally atheoretical way. Adherents of various frameworks can easily formulate alternative hypotheses about what exactly the low-performing individuals have not acquired.

Sekerina and **Schulz** both suggest that we should go beyond merely documenting differences and attempt to determine their causes. I couldn't agree more: but the first step is to establish that such differences do exist – as demonstrated by some of the comments, this is still controversial. Future research will need to investigate the reasons for individual differences much more rigorously, examine the interrelationships between cognitive, affective, and environmental factors, and determine whether specific factors affect different constructions in different ways. As Schulz points out, there is a large number of studies on individual differences in SLA which could serve as models for such research.

One particular area which is likely to be particularly rich in theoretical implications is the effect of literacy on language development. **Vainikka and Young-Scholten** point out that the grammars of literate L2 learners may be fundamentally different from those of naturalistic learners. The research reviewed in the keynote article suggests that analogous differences may be found in the grammars of L1 speakers. Several commentators offer supporting evidence. **Birdsong** reviews a number of studies suggesting that good readers process spoken language faster and more accurately; and **Sparks** points out that L1 print exposure predicts both L1 and L2 achievement.

Hulstijn makes a similar point, arguing that we need to distinguish between basic and non-basic language. The former includes core grammatical constructions and vocabulary items shared by all unimpaired native speakers and relies on the Language Acquisition Device, while the latter also implicates higher cognitive processes and is not shared by all speakers.

While it would certainly be useful to know which aspects of language are shared by (nearly) all native speakers and which are not, establishing the contents of ‘Basic Language Cognition’ will be very difficult in practice. It is not clear, for example, how one could provide a non-circular definition of language impairment, or what would count as mastery of a particular construction – would we require consistent correct performance, or would a speaker who performed just above chance also qualify? The basic/nonbasic distinction will almost certainly turn out to be a matter of degree rather than a strict dichotomy.

Even more problematic is the suggestion that "non-basic" aspects of language rely on different cognitive abilities. As **Ellis** points out, basicness is at least to some degree a function of frequency: the most frequent constructions are likely to be mastered by all learners. True, frequency itself depends on factors such as salience, functionality, and complexity, but it is also partly a matter of linguistic convention: in some languages (e.g. French), the basic word order within the noun phrase is noun-adjective, with adjective-noun available as a marked (i.e., nonbasic) variant, while in others (e.g. Polish), the opposite is the case.

DeKeyser raises several points connected with age effects in acquisition. He suggests that the existence of variability in L1 does not undermine the critical period hypothesis, but rather helps to explain some findings which at first appear to be problematic for the CPH. High-aptitude late L2 learners do very well on some tests because they rely on explicit knowledge; therefore, DeKeyser argues, second language researchers interested in age effects should "avoid structures for which quite a bit of variability has been documented; otherwise it is a foregone conclusion that the ranges of L1 and L2 variation are going to overlap" (XX3). Three points immediately come to mind in connection with this observation. First, it is not a foregone conclusion that the ranges of variation are going to overlap: if a structure is so difficult that even native speakers struggle with it, it would not be unreasonable to expect it to be unlearnable by most L2 learners. Secondly, DeKeyser's recommendation to avoid structures which are variable even in L1 makes sense if the purpose of one's inquiry is to demonstrate the existence of critical period effects. But if the purpose of our inquiry is to understand the nature of L1 and L2 speakers' knowledge, then clearly we must look at both types of structures. Last but not least, given that there are considerable differences in native speakers' performance on fairly basic grammatical structures, there may not be very much left to study if we avoid those structures for which there is variability!

DeKeyser also observes that "critical period" effects are most likely to show up on tasks tapping implicit/procedural knowledge and processing. However, it is very difficult to determine whether the differences are due to age effects per se, or to proficiency. Comparing non-native speakers with native speakers who are matched for linguistic proficiency may shed light on this issue. One recent study which did this (Pakulak and Neville 2011) found different ERP responses in native and non-native speakers, which would support DeKeyser's position. However, it is clear that more research is required. Pakulak and Neville tested very simple phrase structure violations in NPs (**at my his farm*); it remains to be seen whether similar effects are observed for more complex constructions.

If consistent performance is regarded as the hallmark of proceduralized knowledge (cf. Paradis 2009), then the variable performance observed in low academic attainment L1 speakers suggests that not all first language grammatical knowledge is proceduralized, at least not in all speakers. Furthermore, as pointed out in the keynote article, there are considerable differences in age of acquisition of non-basic constructions in L1: some learners may know all there is to know about the passive by age 4, while others may not master the construction until late childhood or adolescence – or never master it.

This raises an interesting question: if an L1 learner acquires a particular construction relatively late in development, is her representation of this construction more like that of second language learners than like that L1 learners who mastered it early? As **Pakulak** points out, while adult linguistic attainment correlates with education, early experience may explain more variance. This would have important consequences for work on critical period effects, as well as obvious social and educational implications.

This again suggests that the differences between L1 and L2 knowledge and processing may be a matter of degree. Language comprehension and production, even in native speakers, involves a combination of highly automatic and more effortful, controlled processes (Novick et al. 2005, Ye and Zhou 2009); interpreting the same construction may rely primarily on automatic processes in some speakers, and primarily on controlled processes in others (cf. Novick et al. 2005). Thus, differences between L1 and L2 speakers may amount to differences in the number of constructions that they have proceduralized. Future research on ultimate attainment, whether in the first or second language, will need to allow for variation not just between speakers, but also between constructions. Grammars are patchworks of constructions, not monolithic blocks.

REFERENCES

Avrutin, S. (2006). Weak syntax. In K. Amundt & Y. Grodzinsky (Eds.), *Broca's Region*. Oxford: Oxford University Press.

Chipere, N. (2001). Native speaker variations in syntactic competence: Implications for first language teaching. *Language Awareness, 10*, 107-124.

Dąbrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language, 58*, 931-951.

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*, 604-615.

Hackman, D., & Farah, M. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Science, 13*, 65-73.

Hackman, D., Farah, M., & Meaney, M. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience, 11*, 651-659.

Neville, H. J., Stevens, C., Klein, S., Fanning, J., Bell, T., Isbell, E., et al. (2011). Improving behavior, cognition and neural mechanisms of attention in lower SES children. *Society for Neuroscience, 37*.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective and Behavioural Neuroscience, 5*, 263-281.

Pakulak, E., & Neville, H. (2011). Maturational constraints on the recruitment of early processes for syntactic processing. *Journal of Cognitive Neuroscience, 23*, 2752-2765.

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins.

Sorace, A. (1993). Incomplete vs divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research*, *9*, 22–47.

Street, J., & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua, 120*, 2080-2094.

Street, J., & Dąbrowska, E. (in press). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholinguistics*.

Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience and Biobehavioral Reviews, 33*, 1168-1177.