The principles of scientific inquiry

Caroline F. Rowland

Language Development Department, Max Planck Institute for Psycholinguistics, The Netherlands / ESRC LuCiD Centre, University of Liverpool

Cognitive science is a multi-disciplinary endeavour. In many ways, this is an advantage; the strongest theories of L1 and L2 acquisition are supported by a consensus of evidence across disciplines such as psychology, linguistics, neuroscience and computational science (Monaghan & Rowland, 2017). However, multidisciplinarity can also lead to confusion because different disciplines approach similar research questions from diverse perspectives, and bring with them very different, often unspoken, assumptions.

Given this, an important factor that unites all the cognitive sciences is adherence to the principles of Popperian deductive science (Popper, 1959). For Popper, as for most modern scientists, the criterion distinguishing science from non-science is falsifiability. A theory is only scientific if it makes predictions that are potentially incompatible with empirical observations. Theories that are compatible with all observations, either because they are endlessly modified to accommodate such observations, or because they are so vague as to be consistent with all possible observations, are unscientific. For many cognitive scientists, there is something frustrating about a theory that makes no falsifiable predictions, a frustration shared by Popper with respect to the psychoanalytic theories of Freud and Adler, which "do not exclude any physically possible human behaviour" (Popper, 1974, p. 985).

Thus, I am pleased that Yang (2018) shows an admirable commitment to all four principles of Popperian scientific reasoning in this article. First, his models adhere to the principle that theories have to be internally consistent; they cannot contain any mutually incompatible statements (i.e. contradictions). For example, the Variational model (Yang, 2002) assumes that (a) the child entertains a number of grammars in the learning space, (b) these grammars are associated with probabilities or weights, and (c) learning takes place as the child makes changes to the probabilistic distribution of the grammars in response to the input. The

consequent prediction is, thus, that (d) "the amount of unambiguous evidence for parameter values in child-directed input corpora [will] correlate with the developmental time course of the parameters" (p. 669). All of these statements are logically compatible.

The second principle requires that the logical form of the theory must be explicit and formulated in such a way that its basic statements correspond to experience. In other words, the theory must describe a possible real world. Here, there are some areas in which Yang's assumptions seem underspecified (e.g. how the Variational model solves Pinker's (1984) bootstrapping problem), but overall much of what Yang proposes is explicit enough to be testable against experience. In particular, his mathematical models describe plausible learning mechanisms – mechanisms that are possible given what we already know about learning in the brain. For example, the Variational model incorporates a learning parameter that captures the fact that some children are slower to process linguistic information than others (Fernald, Perfors, & Marchman, 2006).

The third principle states that it must be possible to compare, favourably, the new theory to existing ones. Again, Yang's theories adheres to this principle. For example, he shows that the Variational model provides a better explanation of the null subject stage in English acquisition, in which children omit obligatory subjects (e.g. _ *want cookies*), than traditional parameter setting theories. These theories predict, wrongly, that children's language will resemble that of speakers of pro- or topic-drop languages during this stage. The Variational model, however, is compatible with the pattern of subject omission in children's speech.

Fourth and finally, a scientific theory must be testable "by the empirical application of the conclusions derived from it" (Thornton, 2017: Section 4). It must make predictions about the behaviour of the world that can be tested against observations in the world. If the predictions are supported by the evidence, the theory is corroborated. If not, the theory cannot be completely correct. Again, Yang adheres to this principle (at least in part, see below) because his predictions are clearly stated and tested against empirical evidence.

For example, in Section 3.2, Yang (2018) summarises work from 2013, in which he tested the predictions of two theories of determiner acquisition. The early productivity theory proposes that children's "combinatorial productivity [in their use of determiners] is on full display from the earliest testable stage of multi-word combinations" (2018, p. 677). The logical prediction of this hypothesis is that young children's productivity with determiners, defined as their ability to combine determiners *a* and *the* freely with nouns, will be as productive as that of adults. For example, if 19% of the nouns used by adults are paired with both *a* and *the* in a speech sample, then approximately 19% of nouns used by children should occur with both *a* and *the* in a similar size sample (combinatorial productivity).

The lexical specificity theory, however, claims that children start out with lexically specific knowledge of how to combine *a* and *the* with certain nouns (often in semi-formulaic phrases such as *where's the X, that's a Y)*, and only later develop the understanding that determiners can combine with nouns productively (Yang cites Pine & Lieven, 1997, but see also Pine, Freudenthal, Krajewski, & Gobet, 2013; Pine & Martindale, 1996). The logical prediction of this hypothesis is that young children's determiner use will be more restricted than that of adults (e.g. if 19% of an adult's nouns occur in a sample with both determiners, significantly fewer of the children's nouns will occur with both determiners).

Testing these two hypotheses is made difficult by the fact that both adult and child speech follows a Zipfian distribution, in which a small number of words occur very often, but most words are rare even in large corpora. This means that even adults may show limited combinatorial productivity, because the chances of these low frequency nouns occurring with both *a* and *the* in a given speech sample are small. Using data from corpora of children's and child directed speech, Yang (2013) showed that there is no significant difference in the productivity of children's and adult's determiner usage, once we control for the expected distribution of nouns and determiners. This evidence supports the early productivity theory's prediction that young children's use of determiners *a* and *the* is adult like, and not the prediction of the lexical specificity theory. Interestingly, the prediction does not hold for adult L2 learners of English, who show more limited combinatorial productivity (see Section 5.2, Table 4, p. 691). Thus, in accordance with Popper's fourth principle, Yang concludes that the extension of the theory to L2 acquisition does not hold, at least in the theory's current form.

However, with regards to this fourth principle of falsifiability, I would urge Yang to go further, and directly address evidence which other authors have provided on this question. In the spirit of Popperian scientific enquiry, theories need to be evaluated against all the evidence against them. There are, thus, some surprising omissions in this article.

In particular, I was surprised there was no discussion of Pine et al.'s (2013) detailed work on determiner acquisition, which built on Yang's (2013) own careful mathematical analysis. Pine et al. agreed with Yang that the expected productivity scores should take into account the Zipfian distribution of language. However, they also argued that Yang's analysis underestimated the productivity of adult speech because it did not control for the fact that adults know more nouns than children: "it is not based on a controlled set of nouns, but on an adult corpus that includes between 5 and 16 times as many different nouns as any of the child corpora being analysed" (Pine et al., 2013, p. 349). Given the Zipfian distribution of nouns and determiners in speech, this means that "the proportion of nouns that occur with low frequency [and are thus unlikely to show overlap] in the adult

corpus is likely to be considerably higher than the proportion of nouns that occur with low frequency in the children's corpora" (Pine et al., 2013, p. 349). Thus, they argued, Yang's analyses substantially underestimates the productivity of adult speech, invalidating the comparison with child speech.

In a detailed set of analyses, Pine et al. (2013) showed that failing to control for noun identity does, indeed, depress adult productivity scores (study 1). They then compared child and adult speech, restricting the analysis to the nouns that both adults and children produced, thus removing the confound. Contrary to Yang's (2013) conclusions, children's determiner use was significantly less productive than that of adults (study 2). There was also a developmental cline in productivity, with children's determiner use becoming more productive with age (study 3), again contrary to the prediction of the early productivity account.

Of course, this is only a single paper (though note that Meylan, Frank, Roy, and Levy (2017) replicate the findings of Pine et al., study 3). Popper himself acknowledged that a single counter-example is not sufficient to falsify a theory in practice. Non-corroboration is not necessarily falsification because no one observation is free from error (see also Lakatos, 1978). Thus, we should not dismiss a theory on the basis of one study alone.

However, Pine et al.'s (2013) evidence is not discussed at all in the keynote article, bar a passing reference in a footnote. I do not think Popper would approve. Even if we disagree with the conclusions of our critics, the solution is to engage with their evidence, both on an intellectual level using logical argumentation, and on an empirical level in replication attempts. Yang's three mathematical models have extremely interesting implications for research on both L1 and L2 acquisition. However, we must be careful to weigh up all the evidence when we evaluate our models. "Those among us who are unwilling to expose their ideas to the hazard of refutation do not take part in the scientific game" (Popper, 1959, p. 280).

Funding

Acknowledgements

References

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Lakatos, I. (1978). The methodology of scientific research programmes. In J. Worrall & E. Currie (Eds.), *Philosophical papers ; vol. 1* (p. viii, 250). Cambridge, UK: Cambridge University Press.

Meylan, S. C., Frank, M. C., Roy, B. C., & Levy, R. (2017). The emergence of an abstract grammatical category in children's early speech. *Psychological Science*, 28(2), 181–192. https://doi.org/10.1177/0956797616677753

Monaghan, P., & Rowland, C. F. (2017). Combining language corpora with experimental and computational approaches for language acquisition research. *Language Learning*, 67, 14–39. https://doi.org/10.1111/lang.12221

Pine, J. M., & Martindale, H. (1996). Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language*, 23(2), 369–95. https://doi.org/10.1017/S0305000900008849

Pine, J. M., & Lieven, E. V. M. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics*, 18(02), 123–138. https://doi.org/10.1017/S0142716400009930

Pine, J. M., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition*, 127(3), 345–360. https://doi.org/10.1016/j.cognition.2013.02.006

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA.: Harvard University Press. https://doi.org/10.1007/s13398-014-0173-7.2

Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.

Popper, K. R. (1974). *The philosophy of Karl Popper*. (P. A. Schilpp, Ed.). La Salle, Illinois: Open Court Publisher.

Thornton, S. (2017). Karl Popper. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford University. Retrieved from https://plato.stanford.edu/archives/sum2017/entries/popper/

Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford University Press.

Yang, C. D. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, 110(16), 6324–6327. https://doi.org/10.1073/pnas.1216803110

Yang, C. (2018). A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism* 8(6), 665–706.

Address for correspondence

Caroline F. Rowland
Max Planck Institute For Psycholinguistics
Wundtlaan 1, 6525 XD Nijmegen
The Netherlands

Caroline.Rowland@mpi.nl

Publication history