

Authors' response

BEN AMBRIDGE, EVAN KIDD, CAROLINE F. ROWLAND AND
ANNA THEAKSTON

Our target paper argued for the ubiquity of frequency effects in acquisition, and that any comprehensive theory must take into account the multiplicity of ways that frequently occurring and co-occurring linguistic units affect the acquisition process. The commentaries on the paper provide a largely unanimous endorsement of this position, but raise additional issues likely to frame further discussion and theoretical development. Specifically, while most commentators did not deny the importance of frequency effects, all saw this as the tip of the theoretical iceberg. In this short response we discuss common themes raised in the commentaries, focusing on the broader issue of what frequency effects mean for language acquisition.

LANGUAGE ACQUISITION AND LANGUAGE PROCESSING

The first theme arising from the commentaries concerned the intimate relationship between acquisition and language processing. While acquisition researchers are ultimately concerned with how children come to acquire linguistic knowledge, this process is inevitably constrained by the fact that input must be filtered through the parser (Bates & MacWhinney, 1982; Fodor, 1998; Seidenberg & MacDonald, 1999). What this means is that language acquisition is, in fact, a *PRODUCT* of the language processing system, and is thus constrained by the limitations of that system's capabilities. As O'Grady and Ellis & Ogden point out, this fact provides us with constraints on theory building that acquisition researchers sometimes forget. Concepts such as "the power law of learning", "contingency", ambiguity, and "prototypicality" (Ellis & Ogden) can be used to determine how much of the input signal is likely to be processed by the child's language processor, which will in turn determine how much that child learns from any one instance of input. Conceptualizing acquisition in these terms links acquisition to "sixty years of psycholinguistic research into ... fluent language processing" (Ellis & Ogden) and, importantly for our focus here, can link acquisition to processing theories based on learning theory that "generate the observed frequency effects for free" (Ellis & Ogden). For instance, this approach links frequency effects to demonstrations of powerful distributional

learning abilities in young infants (Arnon, Diessel), and directs us to domains and phenomena that further constrain our theories (e.g., memory, O'Grady). As Ellis & Ogden point out, language is a complex adaptive system; dynamic processes such as language acquisition, real-time language processing, and language variation and change (as discussed by Diessel) tell us that it couldn't be any other way. Frequency effects are one signature of this dynamism.

MECHANISMS FOR ACQUISITION

However, the story only begins here, and inevitably the devil is in the detail. In particular, we currently lack consensus on the theoretical constructs that constitute the child's processing system and how it generates frequency effects. Once again, situating language acquisition within the broader field of cognitive science provides important clues to the nature of the language learning mechanism. Ellis & Ogden point us toward exemplar-based theories of cognition, situating language acquisition within classic domain-general approaches to cognition (Ebbinghaus, 1913 [1885]; see also Reber, 1993), the footprints of which we see in functionalist theories of acquisition (e.g., MacWhinney & Bates, 1989; Tomasello, 2003) and language (e.g., Bybee, 2010; Goldberg, 2006; Diessel, this volume). Adding to the focus on learning theory, Arnon identified PREDICTION as integral; specifically, a distributional learner processing language in real time will come to make predictions about upcoming content based on her past learning history, and inevitable errors in prediction lead to representational change in the system. Prediction as an explanatory construct has gained traction in the adult psycholinguistic literature (e.g., Hale, 2006; Jaeger, 2010; Pickering & Garrod, 2013), and is an integral component of Chang, Dell, and Bock's (2006) connectionist model of syntax, which explicitly links acquisition to adult sentence production. If children are doing prediction-for-learning (Chang, Kidd & Rowland, 2013), then this may be an important driver of representational change in the developing linguistic system.

Turning to the processing cost associated with syntactic acquisition (i.e. why some structures are more difficult than others), we agree with O'Grady that the role of frequency is best revealed by studying phenomena where internal constraints work antagonistically with frequency information (consistent with our INTERACTION thesis). He explicitly identifies Working Memory (WM) as one such internal constraint on information processing that places computational limits on acquisition and therefore the shape of the linguistic system. While we have no doubt that information processing is capacity-limited, there is reason to believe that drawing a sharp distinction between WM and long-term linguistic representations, where

frequency is likely to be encoded (Diessel), is something of a false dichotomy (see Jones, Gobet & Pine, 2008). For instance, WM effects in adult processing of object RCs disappear following increased exposure to complex sentences (Wells, Christiansen, Race, Acheson & MacDonald, 2009), and priming studies suggest that increased exposure to object RCs make them more easily accessible (Hutton & Kidd, 2011; Nitschke, Kidd & Serratrice, 2010). Such findings are difficult to account for in a model that treats WM and long-term representations as independent (for further discussion, see Kidd, 2013).

Explicating the dynamic mechanisms of language acquisition has important consequences for explaining individual differences in language acquisition (Rowe, Schwartz). As Rowe points out, frequency in the input does not equate to guaranteed and uniform intake for all children. This is evident in separate points made by Shirai and Schwartz: Shirai argues that a relatively frequent association between telic verbs and past tense marking translates to an overwhelming preference for the same association in acquisition in English, whereas Schwartz has shown that acquisition varies with exposure schedule (an effect not exclusive to language; Dunlosky, Rawson, March, Nathan & Willingham, 2013), which differs for children with language impairment. Therefore, while variations in input frequency attributable to environmental influences (e.g., SES) and properties of social interaction (e.g., parental responsiveness, joint attention) significantly affect acquisition, these effects interact with internal neurocognitive constraints and pre-existing biases. This means that a full explanation of individual differences requires an understanding of how the language processing system filters the input, which, in turn, requires an understanding of the theoretical constructs that constitute that processing system. This brings our discussion of mechanism full circle. O'Grady identifies WM as one candidate for constraints on the language processing system, but there are other mechanisms that directly bear on a child's ability to learn from the input. One promising idea is that the child's ability to find patterns in linguistic material (and thus her ability to do statistical learning; Aslin & Newport, 2012; Romberg & Saffran, 2010) is beginning to be linked to individual differences in typical language development (Kidd, 2012) and to language impairment (e.g., Hsu and Bishop, 2010, argue that impaired statistical learning skills may prevent children with SLI developing abstract syntactic representations).

However, in a dynamic language system, individual differences ultimately arise from a complex interplay between external factors and internal constraints on acquisition over time. For instance, while the quality of the linguistic environment will exert an independent effect on uptake, the efficiency of a child's learning mechanism will also be a PRODUCT of her environment up to that point (e.g., 18-month-olds who have richer input

process linguistic stimuli more quickly; Hurtado, Marchman & Fernald, 2008). Therefore, as Rowe correctly points out, other levels and kinds come into play, such that it is likely that frequency has non-linear effects on uptake. Capturing these processes in a comprehensive theory of language acquisition is no small task. Here we follow Ellis & Ogden in championing the role of computational models in this research effort (e.g., Twomey, Chang & Ambridge, 2014), and follow Shirai and Edwards, Beckman & Munson in stressing the importance of cross-linguistic research in teasing apart factors, such as verb frequency and verb semantics, that are often confounded in commonly researched languages like English (Shirai) and in determining which effects are input driven and which reflect universal constraints (Edwards *et al.*).

HOW SPECIAL IS LANGUAGE?

A discussion of frequency effects and the learning mechanisms that produce them raises the broader theoretical question: How special is language in comparison to other cognitive domains? The commentaries by Arnon, Diessel, and Ellis & Ogden link frequency effects in language acquisition to similar effects in the broader cognitive science literature, and point to domain-general learning mechanisms that could, in principle, explain frequency effects in acquisition. However, Yang proposes that frequency effects have been ignored in the past with “ample justification”, arguing that there is a more limited role for frequency effects in language acquisition than in other areas of learning. We disagree. Our main point is that there are genuine frequency effects across all levels of language that need to be incorporated into our theories of acquisition. Such effects go beyond the phenomena we specifically discussed; as Edwards *et al.*, showed, our framework for discussing frequency effects is also useful in explaining early phonological development. We contend that an accurate categorization of these effects will lead us closer to a better understanding of the learning mechanisms that support acquisition.

However, this research effort will not be fruitful unless we are all clear on the role of frequency in acquisition and unless we all take into account the full range of data in doing so. The difference between Yang's and our own interpretation of the literature demonstrates that this is not necessarily a straightforward task. Contra Yang, we argue that the data for over-regularization errors show clear effects of surface frequency and neighbourhood density in preventing or producing error (see, amongst others, Albright & Hayes, 2003; Alegre & Gordon, 1999; Ambridge, 2010; Kidd & Lum, 2008; Maratsos, 2000; Marchman, 1997; Marchman, Wulfeck & Ellis Weismer, 1999; Maslen, Theakston, Lieven & Tomasello, 2004; Matthews & Theakston, 2006; Prasada & Pinker, 1993). With

respect to generalizations of irregular forms, Ambridge (2010) demonstrated that children show clear evidence of an ability to generalize irregular patterns (to novel forms), contra the claim that “[g]eneralizations of irregular forms are virtually absent in child English” (Yang). There is also ample evidence that children and adults store high-frequency phrases as chunks (Bannard & Matthews, 2008) or partially abstract representations (e.g., Pine, Freudenthal, Krajewski & Gobet, 2013), and that the errors they make with these expressions are best explained in terms of the frequency of the phrase and not the individual lexical items (e.g., Kirjavainen, Theakston & Lieven, 2009; Kirjavainen, Theakston, Lieven & Tomasello, 2009). In addition, stored lexical strings (e.g., *gimme*) are different from utterances where separate elements are combined (e.g., *give you*) on dimensions other than the frequency of their variable elements (i.e., *me* versus *you*); one signature feature is that they are typically phonologically reduced (Bybee, 2010). Insofar as the frequency of lexical strings predicts children’s performance in experiments manipulating frequency while holding categorical rule application constant (Bannard & Matthews, 2008), there is something to explain (see also Arnon). Finally, work by us and our colleagues has identified frequency effects in children’s acquisition of sentences containing structure dependence (Ambridge, Rowland & Pine, 2008; see also Lewis & Elman, 2001; Realı & Christiansen, 2005), and in the provision of optional infinitives (e.g., Freudenthal, Pine & Gobet, 2010; Räsänen, Ambridge & Pine, 2014). We contend that, far from proving inverse or anti-frequency effects, these data demonstrate that looking in the right place is an important first step in elucidating the power and limitations of learning in language acquisition.

The shadow of broader debates in cognitive science will always be with us, but we are obliged to reach some consensus when the evidence overwhelmingly points us in one direction. The evidence, which is corroborated by our commentators, is that frequency effects are real and pervasive in acquisition. The challenge for any theory of acquisition is to explain them. We hope that our target paper and the collection of commentaries serve to sharpen future research and theoretical development.

REFERENCES

- Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90(2), 119–61.
- Alegre, M. & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40(1), 41–61.
- Ambridge, B. (2010). Children’s judgments of regular and irregular novel past tense forms: new data on the dual- versus single-route debate. *Developmental Psychology* 46(6), 1497–504.
- Ambridge, B., Rowland, C. F. & Pine, J. M. (2008). Is structure dependence an innate constraint? New experimental evidence from children’s complex question production. *Cognitive Science* 32, 222–55.

- Aslin, R. N. & Newport, E. L. (2012). Statistical learning: from learning items to generalizing rules. *Current Directions in Psychological Science* **21**, 170–6.
- Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning – the effect of familiarity on children's repetition of four-word combinations. *Psychological Science* **19**, 241–8.
- Bates, E. & Mac Whinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (eds.), *Language acquisition: the state of the art* (pp. 173–218). New York: Cambridge University Press.
- Bybee, J. (2010). *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Chang, F., Dell, G. S. & Bock, K. (2006). Becoming syntactic. *Psychological Review* **113**, 234–72.
- Chang, F., Kidd, E. & Rowland, C. F. (2013). Prediction in processing is a by-product of language learning. *Behavioral and Brain Sciences* **36**, 350–1.
- Dunlosky, J., Rawson, K. A., March, E. J., Nathan, M. J. & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest* **14**, 4–58.
- Ebbinghaus, H. (1913 [1885]). *Memory: a contribution to experimental psychology*. New York: Teachers College, Columbia University.
- Fodor, D. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research* **27**, 339–74.
- Freudenthal, D., Pine, J. M. & Gobet, F. (2010). Explaining quantitative variation in the rate of optional infinite errors across languages: a comparison of MOSAIC and the Variational Learning Model. *Journal of Child Language* **36**, 643–69.
- Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science* **30**, 609–42.
- Hsu, H. J. & Bishop, D. V. M. (2010). Grammatical difficulties in children with specific language impairment (SLI): Is learning deficient? *Human Development* **53**, 264–77.
- Hurtado, N., Marchman, V. A. & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science* **11**, F31–9.
- Hutton, J. & Kidd, E. (2011). Structural priming in comprehension of relative clauses: in search of a frequency by regularity interaction. In E. Kidd (ed.), *The acquisition of relative clauses: typology, processing, and function* (pp. 227–242). Amsterdam: John Benjamins.
- Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology* **61**, 23–62.
- Jones, G., Gobet, F. & Pine, J. M. (2008). Computer simulations of developmental change: the contributions of working memory capacity and long-term knowledge. *Cognitive Science* **32**(7), 1148–76.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology* **48**, 171–84.
- Kidd, E. (2013). The role of verbal working memory in children's sentence comprehension: a critical review. *Topics in Language Disorders* **33**, 208–23.
- Kidd, E. & Lum, J. (2008). Sex differences in past tense overregularization. *Developmental Science* **11**, 882–9.
- Kirjavainen, M., Theakston, A. & Lieven, E. (2009). Can input explain children's me-for-I errors? *Journal of Child Language* **36**(5), 1091–114.
- Kirjavainen, M., Theakston, A., Lieven, E. & Tomasello, M. (2009). Children's omission of infinitival-to. *First Language* **29**(3), 315–41.
- Lewis, J. D. & Elman, J. (2001). Learnability and the statistical structure of language: poverty of stimulus arguments revisited. In Barbora Skarabela, Sarah Fish, & Anna H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (Vol. 1, pp. 359–370). Somerville, MA: Cascadilla Press.
- MacWhinney, B. & Bates, E. (1989). *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

- Maratsos, M. (2000). More overregularisations after all: new data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language* **27**, 183–212.
- Marchman, V. A. (1997). Children's productivity in the English past tense: the role of frequency, phonology and neighborhood structure. *Cognitive Science* **21**(3), 283–304.
- Marchman, V., Wulfeck, B. & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: a study of the English past tense. *Journal of Speech, Language, and Hearing Research* **42**, 206–19.
- Maslen, R. J. C., Theakston, A. L., Lieven, E. V. M. & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language, and Hearing Research* **47**, 1319–33.
- Matthews, D. E. & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: the effects of frequency, phonology, and competition. *Cognitive Science* **30**(6), 1027–52.
- Nitschke, S., Kidd, E. & Serratrice, L. (2010). First language transfer and long-term structural priming in comprehension. *Language and Cognitive Processes* **25**, 94–114.
- Pickering, M. J. & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Science* **36**, 329–47.
- 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), 345–60.
- Prasada, S. & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* **8**, 1–56.
- Räsänen, S. H., Ambridge, B. & Pine, J. M. (2014). Infinitives or bare stems? Are English-speaking children defaulting to the highest-frequency form? *Journal of Child Language* **41**(4), 756–79.
- Reali, F. & Christiansen, M. H. (2005). Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science* **29**, 1007–28.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: an essay in the cognitive unconscious*. Oxford: Oxford University Press.
- Romberg, A. R. & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science* **1**(6), 906–14.
- Seidenberg, M. S. & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science* **23**, 569–88.
- Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Twomey, K., Chang, F. & Ambridge, B. (2014). Do as I say, not as I do: a lexical distributional account of English locative verb class acquisition. *Cognitive Psychology* **73**, 41–71.
- Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J. & MacDonald, M. C. (2009). Experience and sentence comprehension: statistical learning and relative clause comprehension. *Cognitive Psychology*, **58**, 250–71.