Infants' history of distributional learning in real time

Casey Lew-Williams
Princeton University

Phillips and Ehrenhofer (2015), henceforth P&E, are doing the field a favor. They link adult psycholinguistic research not only to developmental psycholinguistic research, but to research on language learning that emerges from the psychological tradition. A range of scientists — those who get out of bed in the morning to study language, to study development, and/or to study the cognitive science of learning — will find important insights embedded throughout the authors' discussion of how language processing informs language acquisition. The title of the article broadly refers to language acquisition, but the content is focused primarily on the level of syntax and semantics. This spotlight is by no means problematic, but it invites my main response: for a complete account of the relation between language processing and language acquisition, we need elaboration on the role of infancy in staging the emergence of language processing.

In the first months of life, language processing happens as infants gradually accumulate information about distributions of sounds, syllables, and words. 'Statistical learning' is at best a meta-theoretic perspective, and at worst an atheoretical lab observation, but we have preliminary evidence for its relevance to semi-natural learning contexts (e.g., Frank, Tenenbaum, & Gibson, 2013; Hay, Pelucchi, Graf Estes, & Saffran, 2011; Kurumada, Meylan, & Frank, 2013; Lew-Williams, Pelucchi, & Saffran, 2011). The target article culminates in a discussion of why (typically developing) children are well-rounded and high-achieving language learners, and to do so, the authors turn to research on adult second language learners. Adult learners, the P&E appropriately suggest, have early successes in language processing that lock them into routines and block them from learning new information (see Arnon & Ramscar, 2012). This is followed by a consideration of children's relatively weak memory abilities, inferior reanalysis abilities, and immature prediction abilities as likely culprits for the lack of a comparable trajectory in early development. While children's limitations in revising initial interpretations

Linguistic Approaches to Bilingualism 5:4 (2015), **494–498**. DOI 10.1075/lab.5.4.09lew
ISSN 1879–9264 / E-ISSN 1879–9272

represent a critical part of the story, another angle is needed to complement the narrative: infants' and toddlers' early successes in language processing may not block but instead open doors to learning new information. They may not have the same difficulties as adult learners because they get to exploit a history of co-occurrence information; what infants do for a living is engage in a temporally protracted process of input-based pattern extraction, both within and across the auditory and visual systems. And at some point on a currently unspecified timetable, they graduate beyond simple patterns and form generalizations over the input (see Marcus, Fernandes, & Johnson, 2007). For adult learners, complex sentences come in at a non-optimal pace (perhaps up to 5 words per second), and language learning in traditional classroom contexts hinges on the testimony of an instructor. But for child learners, the input is often friendly enough to guide them toward reasonably efficient interpretations of the signal. New information and new structures are not necessarily a burden; instead, novelty iteratively rests on a wealth of accumulated distributional learning.

In this sense, it seems inaccurate to conclude that "children's language processing abilities give them no obvious advantage over adult learners" (p. 413). Young children *do* have severe limitations in their abilities to form predictions over complex sequences of words, for a variety of reasons: they have two instead of 20 years of language experience, they have immature nervous systems, and they lack an optimal incorporation of explicit strategizing into the task of learning. But their language processing abilities also have a historically based statistical advantage over adult learners. While they are only capable of using the most common adjectives (Fernald, Thorpe, & Marchman, 2010), verbs (Fernald, Zangl, Portillo, & Marchman, 2008), and articles (Lew-Williams & Fernald, 2007) to 'get ahead of' the input (prediction abilities that may only be apparent in distilled lab simulations), their attention to local relations between sounds, syllables, and words gradually gives rise to parsing and adult-like prediction. P&E suggest that accurate sentence analysis is a first step in ensuring eventual proficiency (p. 431), but this statement skips over the foundational 'skill-building' that happens in infancy. More attention is needed to the fact that young children spend many months discovering relevant structure in patterns of sounds, syllables, and words in the input. Learning about the distributional properties of words is just the beginning of becoming a native listener, and over time, infants and toddlers apply their knowledge of distributions to tasks such as recognizing referents and processing words in combination. We now have growing evidence for this continuity (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Graf Estes, Evans, Alibali, & Saffran, 2007; Lany, 2014; Misyak, Christiansen, & Tomblin, 2010).

Importantly, different caregivers provide different access to the units of language and the relations between them (Hart & Risley, 1995; Hoff & Naigles, 2002;

Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2008; Weisleder & Fernald, 2013). But in considering why some linguistic phenomena are hard to learn and why children generally fare better than adults, P&E downplay the potential usefulness of scrutinizing variability in the input. They propose that delayed learning of a specific linguistic phenomenon probably does not result from variability in exposure to that phenomenon. Why? Because we see surprising consistency across children in age of mastery, which is problematic for input-based accounts and necessitates the idea that learner-internal changes "could allow them to confirm or disconfirm regularities that they were previously unable to test" (p. 438). This is an excellent insight about the dynamic interplay between cognition and experience with language, and there is certainly no one-to-one mapping between amount of input and timing of learning a specific linguistic phenomenon. But there should still be a correspondence between the two, and the deployment of advanced cognitive skills may hinge on having received rich input. Children from wealthy families get approximately 3.5 times more language experience per unit time, on average, relative to children from poor families (Hart & Risley, 1995). Moreover, according to my crude estimates of time-spent-in-the-classroom, adult second language learners may fare even worse. These input estimates do not necessarily speak to cases like anaphora, but dissociating global from specific input is odd, given that global input is comprised of diverse exposure to specific linguistic phenomena. And thus it is premature to conclude that variation in learning speed is less dependent on input variability than on learner-internal variability. Researchers interested in infant language, socioeconomic status, early intervention, and the achievement gap have barely begun to tackle the importance or irrelevance of overall input quantity. The field is currently preoccupied with debating the high-level issue of input quantity vs. quality as the engine behind children's divergent learning trajectories, and very few studies have focused specifically on syntax. One exception is a study conducted by Huttenlocher and colleagues (2002), which showed that preschool teachers who use a greater proportion of multi-clause sentences and a greater number of noun phrases per sentence have students who — one year later — demonstrate greater syntactic growth. Dissecting the relations between specific input features and specific linguistic constructs is likely to be a fruitful direction for future work, not one that should be left behind.

The 'less is eventually more' account takes the contrast of mature/immature processing resources in a highly productive direction, but children's limitations need to be considered alongside what we know about the origins of detecting language structure in infancy. Infants may not be better than adults at finding structure in patterned input, but they do receive overwhelmingly more practice. With this practice comes the opportunity to 'miss' information and learn from mistakes.

Newborns miss everything, toddlers mis-predict, and children mis-analyze, because child-directed input is fast, sometimes imprecise, and often novel. While systematic mis-parsing may be destructive for the learner in some ways, erroneous processing — broadly construed — also helps create change to the system (Chang, Dell, & Bock, 2006). The 'less is eventually more' account speculates that these limitation-based errors help young children avoid damage and ultimately shine. But the extended time-scale (the 'eventualness') of the account will benefit by extending its reach back to infancy as the starting point of language learning. Infants and toddlers receive lots of language input, they digest its structure gradually over time, and they confront and learn from erroneous predictions at every turn.

Throughout their discussion, P&E suggest a need for more research that directly compares child and adult learners, more evaluations of the real-world relevance of lab findings, and more focus on the nature of young children's predictive abilities. But they intentionally and wisely withhold suggestions for methodological outreach. Their main point is to understand how language is learned in real time, and to let the field deploy its machinery as it wishes.

References

Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, *122*, 292–305. DOI: 10.1016/j.cognition.2011.10.009

Chang, F., Dell, G.S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*, 234–272. DOI: 10.1037/0033-295X.113.2.234

Conway, C.M., Bauernschmidt, A., Huang, S.S., & Pisoni, D.B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*, 356–371. DOI: 10.1016/j.cognition.2009.10.009

Fernald, A., Thorpe, K., & Marchman, V.A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive Psychology*, *60*, 190–217. DOI: 10.1016/j.cogpsych.2009.12.002

Fernald, A., Zangl, R., Portillo, A., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I.A. Sekerina, E.M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: Online methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins. DOI: 10.1075/lald.44.06fer

Frank, M.C., Tenenbaum, J.B., & Gibson, E. (2013). Learning and long-term retention of large-scale artificial languages. *PLoS One*, *8*, e52500. DOI: 10.1371/journal.pone.0052500

Graf Estes, K.M., Evans, J., Alibali, M.W., & Saffran, J.R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*, 254–260. DOI: 10.1111/j.1467-9280.2007.01885.x

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hay, J.F., Pelucchi, B., Graf Estes, K., & Saffran, J.R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*, 93–106. DOI: 10.1016/j.cogpsych.2011.06.002

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*, 418–433. DOI: 10.1111/1467-8624.00415

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337–374. DOI: 10.1016/S0010-0285(02)00500-5

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L.V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*, 343–365. DOI: 10.1016/j.cogpsych.2010.08.002

Kurumada, C., Meylan, S.C., & Frank, M.C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition*, *127*, 439–453. DOI: 10.1016/j.cognition.2013.02.002

Lany, J. (2014). Judging words by their covers and the company they keep: Probabilistic cues support word learning. *Child Development*, *85*, 1727–1739. DOI: 10.1111/cdev.12199

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*, 193–198. DOI: 10.1111/j.1467-9280.2007.01871.x

Lew-Williams, C., Pelucchi, B., & Saffran, J.R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, *14*, 1323–1329. DOI: 10.1111/j.1467-7687.2011.01079.x

Marcus, G.F., Fernandes, K.J., & Johnson, S.P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, *18*, 387–391. DOI: 10.1111/j.1467-9280.2007.01910.x

Misyak, J.B., Christiansen, M.H., & Tomblin, J.B. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, *2*, 138–153. DOI: 10.1111/j.1756-8765.2009.01072.x

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, *5*.

Rowe, M.L. (2008). Child-directed speech: Relation to socioeco- nomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*, 185–205. DOI: 10.1017/S0305000907008343

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *20*, 1–10.

Author's address

Casey Lew-Williams
Department of Psychology
Princeton University
Princeton, NJ 08544
USA

caseylw@princeton.edu