Maddie Meyers
Earl Franklin Fellowship Research Proposal

Caregiver reconstruction of children's errors: the preservation of structure in language

How do you ask a group of people where they are going in Spanish? In Spain, the answer depends on the group: you might ask "Donde van ustedes?" of a group of work colleagues, but to address your friends, you use the informal "Donde váis vosotros?" instead. In Mexican Spanish, this distinction has disappeared, and the "ustedes" form is used exclusively. Why did Spanish change in this way, simplifying and shedding the formal second person plural? Why do languages change at all, aside from acquiring new vocabulary? One working theory is that languages evolve, like biological organisms, to adapt to two dynamic competing pressures: one, to be easily transmitted and learned (and hence simple), and another, to be an effective system for communication (and hence informative; Lupyan & Dale, 2010). Children are often the actors who drive language evolution (Senghas, 2003), yet they differ from adults in their 1) cognitive capabilities, namely, memory systems (Kempe et al., 2015), 2) interests and early vocabularies, and 3) conversation partners. Therefore, children's developing cognitive systems prevent aspects of language that are difficult to learn and remember from being passed on—pushing languages towards simplicity (Hudson Kam & Newport, 2005; Senghas, 2003). But, languages that become too simple can lose the ability to be effective for communication (Kirby et al., 2014). What enables languages to retain their communicative utility in the face of these learnability pressures?

I propose to test a novel hypothesis for the maintenance of structure in language: Communicative inference by caregivers. Even the youngest children are not passive learners of language—they are active participants, engaging in conversations with their parents. These adults are experts both in the language and in the children themselves—they understand the child's intuitions, personality, and context. Caregivers play an important interpretive role in these interactions by their ability to understand the intended target of children's errorful productions (Chouinard & Clark, 2003). They may reconstruct their child's language in numerous ways--through explicit or implicit correction, or simply through modeling correct use of the language over time (Hudson Kam & Newport, 2005). These reconstructions may be a mechanism by which more structure is retained in language than children could sustain alone.

Iterated Learning

One way to study language learning in the lab is to use the iterated learning paradigm. Created to study the effects of simplicity and informativity on inter-generational language evolution, this method is useful for testing the proposed hypotheses (Kirby et al., 2014). In an iterated learning paradigm, one participant is trained on a randomly-generated language—for example, a set of words created by arbitrarily pairing syllables together. The participant is later asked to recall the language, and their responses are given as training input for the next subject, thus creating a transmission chain (see Figure 1). This iterated process mimics the transmission of language across generations, with each participant unintentionally changing the language through their memory biases. A study by Kempe et al. (2015) compared child (5-8) and adult performances on an iterated learning task using a novel dot-pattern paradigm (showed in Figure 1). Their results found that structure emerged faster in children than adults—that is, the children's patterns simplified much faster than the adult's, allowing them to be easier to reproduce earlier in the transmission chain. This study provides evidence of the importance of looking at both children and adults in an iterated learning paradigm, as they have different cognitive skills which affect their performance in language-learning tasks (Kempe et al., 2015). However, language evolution can never be fully grasped using this paradigm with only separate adult or child earning chains, because language learning does not occur only within the same age group (horizontal transmission), or only across age groups (vertical transmission), but it occurs in both directions. In a true language-acquisition situation, a child receives a great deal of language input from their caregiver and uses it to interact with their peers throughout life, eventually growing into a new teacher-caregiver.

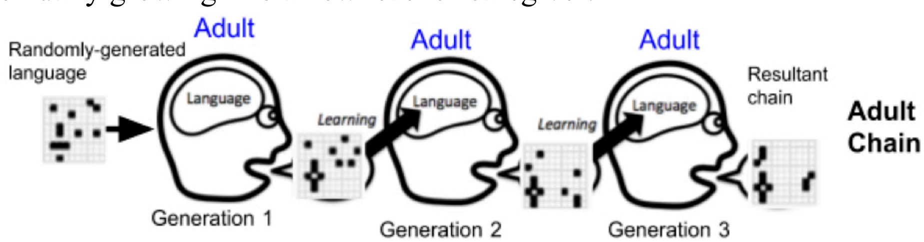

Figure 1 shows an example of the iterated learning paradigm, using dot-patterns stimuli from Kempe et al. (2015)

Methods

To investigate the complexity of adult-child language learning interactions, I will be conducting three iterated learning chains--one with only adults, and another with children (ages 4-6). In the third (dyad) chain, adults will be trained on the language, and children will reconstruct each the target. The child's recalled item will be given to the adult, who will be instructed to correct any errors they find. Instead of using randomly-generated words, I will be using the same kid-friendly dot paradigm introduced by Kempe et al. (2015) (Figure 1). While these dot patterns may not be equivalent to syllables or words, with them we are able to study the same language-learning and memory biases, while removing any interference from a subject's prior language experience. In learning languages, children and adults must infer rules and patterns, just as in the dot-pattern reconstruction task (Kempe et al., 2015). Therefore, the iterated learning paradigm using dot patterns is a simple, effective way of studying how subjects' learning biases affect productions over time.

Data Collection

I plan to create an online iPad task for data collection. iPad tasks have many advantages over other research methods, including the paper-and-sticker task used by Kempe et al. (2015) because the use of an iPad reduces the completion time of the study and is engaging for young children (Frank et al., 2016). By using this task, I hope to collect a greater number of trials—10 from each of 90 children and adults—from a younger set of children—4-6-year olds—than has been collected previously (Kempe et al., 2015). Once the stimuli have been created using JavaScript, I plan to pilot the study with subjects at the Museum of Science and Industry (MSI). After making any necessary adjustments, I hope to collect the duration of the child data at the museum, and the adult data will be collected online using Amazon Mechanical Turk. It is possible to collect the parent data from any adult, as they will have the same amount of proficiency in this novel language as a caregiver would have.

Children in either condition (child chains or child-adult dyad chains), and adults in the adult chains will be shown an 8x8 grid with 10 targets. They will be given an attention task—to click on all of the targets—and will then be shown a visual mask—an image meant to clear any lingering visual effects of the previous image. After the mask, the child will be asked to recreate the pattern they had seen in the previous image by dragging the targets to locations on the grid. The children will complete 10 trials and will receive no feedback on their responses. Each iterated chain will be 5 participants long, with the first participant receiving a randomly-generated pattern, and each subsequent participant receiving the previous subject's responses as input. All transmission chains will begin with the same randomly-generated seed patterns. The dyad chains will have slightly different methodology (see Figure 2). Children will complete the identical task as in the child chains. However, adults in these chains will first be trained on the same input as was given to the child in their generation. Instead of recalling the patterns, these adults will be given the responses of the children and will be asked to correct any errors they see in the child's reconstruction. The adult's "fixed" version of the pattern will be used as training input for the next dyad (Figure 2).

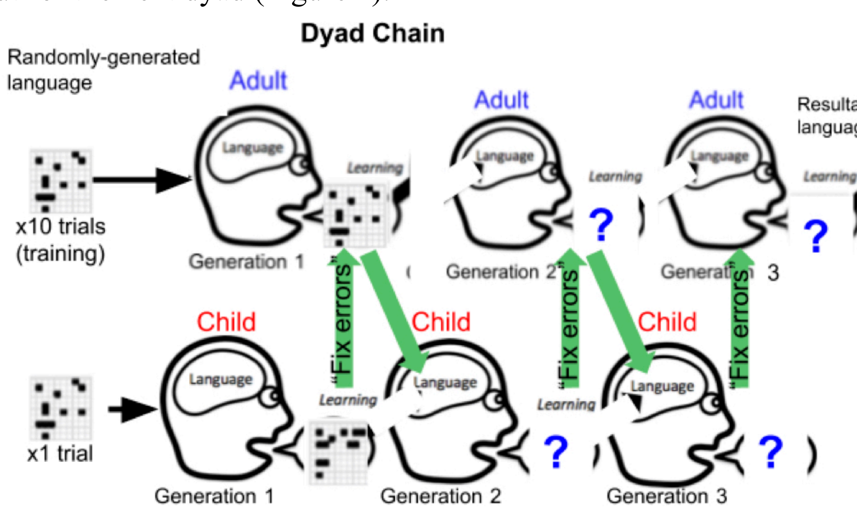

Figure 2 shows the proposed methods for dyad chains.

Data Analysis

I plan to use methods introduced by Kempe et al. (2015) in order to analyze how the dot patterns change over time, in terms of their algorithmic complexity, transmission accuracy, identifiability, and clustering. Algorithmic complexity refers to the amount of processing time necessary to generate a particular structure (dot pattern) and will be measured using the Block

Decomposition Method (Kempe et al., 2015; Zenil et al., 2014). Trends in all features will be analyzed for significance using both Page's *L* trend test and ANOVA tests.

Experimental Hypotheses

While I expect all three chains to decrease in complexity due to the subject's inherent memory constraints, I expect a unique effect in the dyad reconstruction chain in line with the previously proposed hypothesis. Figure 3 shows the hypothesized trends in algorithmic complexity. A child may make greater simplifications in the dot patterns due to their even more limited memory systems and stronger learnability biases—replicating the findings of Kempe et al. (2015). However, in the dyad chains, I expect to see a generation-by-generation effect where children reduce complexity of the language, but the parent reintroduces this complexity by fixing some of the child's errors. It is unclear how much parents can help retain the descriptiveness of language—perhaps they can bring these constructions to mimic their own in complexity. However, we expect them to reintroduce descriptiveness in the language and therefore protect against children's simplicity biases.

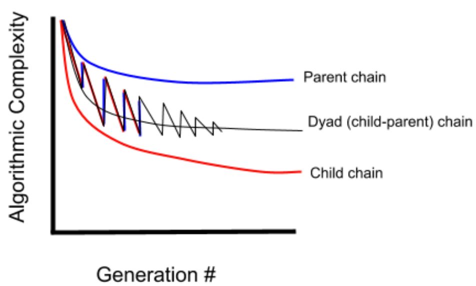

Figure 3 shows the projected results for the proposed study, where all chains show decreases in algorithmic complexity (simplification), but the adults in the dyad chains are reconstructing children's productions to approach their own.

Significance and Implications

The completion of the proposed study will contribute to literature on language acquisition and language change. Importantly, the dynamic contributions of both parents and children are often overlooked in studies of language acquisition, as many isolate either child-directed language (CDL) or child-produced language, instead of looking at them simultaneously.

The theories explored in this study are not only relevant to language development but could possibly be extended to other domains of development, including social, emotional, and moral development. Do caregivers play the same important role in teaching their children cultural values? If children do not practice a cultural value "correctly", or at all, and parents do not reconstruct this mistake, the culture may also simplify or change (Henrich, 2004), but would it change in the same way as a language system?

This study attempts to show not only how caregivers are necessary to help reconstruct an individual child's errors, but also how they are pivotal on a large-scale to protect language as a whole from simplifying to disuse. Thus, correcting a child's errors over time serves a dual purpose: to aid the child, and to protect the language system.

References

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. Journal of child language, 30(3), 637-669.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. Trends in cognitive sciences, 19(10), 603-615.

Henrich, J. (2004). Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *American Antiquity*, *69*(2), 197-214.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. Language learning and development, 1(2), 151-195.

Kempe, V., Gauvrit, N., & Forsyth, D. (2015). Structure emerges faster during cultural transmission in children than in adults. Cognition, 136, 247-254.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. Current opinion in neurobiology, 28, 108-114.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. PloS one, 5(1), e8559.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. Cognitive Development, 18(4), 511-531.

Zenil, H., Soler-Toscano, F., Dingle, K., & Louis, A. A. (2014). Correlation of automorphism group size and topological properties with program-size complexity evaluations of graphs and complex networks. *Physica A: Statistical Mechanics and its Applications*, *404*, 341-358.