Simplifications made early in learning can reshape language complexity

Kenny Smith[*1]

[*]Corresponding Author: kenny.smith@ed.ac.uk
[1]Centre for Language Evolution, The University of Edinburgh, Edinburgh, UK

Languages spoken in larger populations seem to be relatively simple (Wray & Grace, 2007; Trudgill, 2011). One possible explanation is that this is a consequence of the simplifying influence of non-native speakers: adult learners tend to reduce complexity during learning, and large languages tend to have a higher proportion of non-native speakers. This hypothesis, that languages adapt to their social niche (Dale & Lupyan, 2012), receives some statistical support from typological studies which show negative correlations between number of non-native speakers and morphological complexity (e.g. Lupyan & Dale, 2010; Bentz & Winter, 2013; Sinnemäki, 2020, but see ongoing debate in Koplenig, 2019; Kauhanen et al., 2023; Koplenig, 2023). It has also been subjected to experimental tests using artificial language learning techniques, exploring the impact of simplifications made by non-native-like early learners on morphological complexity (Atkinson et al., 2018; Berdicevskis & Semenuks, 2022). Here I report a series of experiments combining their methods, which reconciles the apparent conflict between their results and indicates that the presence of non-native-like early learners in a population can lead to gradual simplification of morphology.

In Experiment 1 I replicate Atkinson et al.'s Experiment 1 using crowdsourced participants (N=94), finding that learners trained on a morphologically complex miniature language simplify its morphology early in learning (e.g. after only 2 bocks of training), but given adequate exposure (e.g. 8 blocks of training) accurately learn the target language. In Experiment 2 I replicate Atkinson et al. Experiment 2: contrary to our original finding, an order of magnitude more data (N=522) suggests that Experiment 2 learners who receive data featuring simplified morphology (produced by early learners from Experiment 1) do show a modest reduction in the morphological complexity of the language they themselves produce. This shows that simplifications made by adult learners can result in simplification of a population's language, although the very small effect size is consistent with Atkinson et al.'s suggestion that learning from the mixed output of multiple individuals reduces or nullifies the simplifications seen in the output of individuals.

However, Atkinson et al. (2018) Experiment 2 (and Experiment 2 here) simulate only a single generation of transmission; other work emphasises the influ-

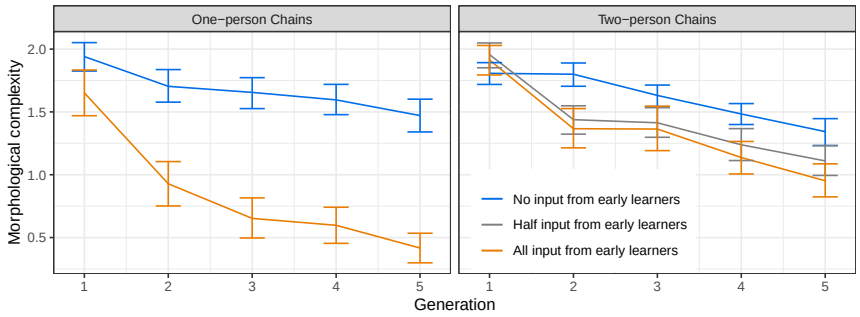

Figure 1. Experiment 3 results, showing morphological complexity (measured using the method from Atkinson et al., 2018) against generation and proportion of input that learners at generation N+1 received from early learners at generation N. Error bars indicate 95% CIs.

ence of multiple generations of transmission in amplifying weak biases in learning (e.g. Reali & Griffiths, 2009). Berdicevskis and Semenuks (2022) therefore use a multi-generation iterated learning paradigm to show that languages transmitted in chains featuring early learners simplify more rapidly than languages from chains featuring no early learners. This result provides support for the hypothesis that simplifications by adult learners could ultimately lead to language simplification. However, Berdicevskis and Semenuks only run chains where each learner is exposed to the output of a single individual at the previous generation, meaning that the mixing mechanism identified in Atkinson et al. is not at play. In Experiment 3 I therefore combine the Atkinson et al. learning paradigm with an iterated learning procedure similar to that used by Berdicevskis and Semenuks, running 5-generation iterated learning chains manipulating (1) the proportion of simplified input each learner receives (i.e. coming from early learners at the previous generation) and (2) the number of individuals in each generation, running both one-person versus two-person chains (N=400, 50 chains). Experiment 3 therefore manipulates proportion of simplified input and allows for mixing effects, as per Atkinson et al., but allows for cumulative effects as per Berdicevskis and Semenuks. Experiment 3 languages gradually simplify (see Figure 1), with simplification being more rapid when learners receive at least some input from early learners at the previous generation. This effect is seen in both one-person and two-person chains, but there is some evidence for mixing effects, with simplification being slower when each learner receives input from multiple individuals.

These results therefore reconcile the apparent mismatch in the experimental literature, being consistent with both Atkinson et al. (2018) and Berdicevskis and Semenuks (2022), and strengthen the experimental evidence for simplification during adult learning as a mechanism which could account for negative correlations between adult learners and linguistic complexity in natural languages.

Acknowledgements

This research received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant 681942).

References

Atkinson, M., Smith, K., & Kirby, S. (2018). Adult learning and language simplification. *Cognitive Science*, *42*(8), 2818–2854.

Bentz, C., & Winter, B. (2013). Languages with More Second Language Learners Tend to Lose Nominal Case. *Language Dynamics and Change*, *3*, 1–27.

Berdicevskis, A., & Semenuks, A. (2022). Imperfect language learning reduces morphological overspecification: Experimental evidence. *PLOS ONE*, *17*(1), e0262876.

Dale, R., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15*, 1150017.

Kauhanen, H., Einhaus, S., & Walkden, G. (2023). Language structure is influenced by the proportion of non-native speakers: A reply to Koplenig (2019). *Journal of Language Evolution*, *8*(1), 90–101.

Koplenig, A. (2019). Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society Open Science*, *6*(2), 181274.

Koplenig, A. (2023). Still no evidence for an effect of the proportion of non-native speakers on language complexity – A response to Kauhanen, Einhaus & Walkden (2023).

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLOS ONE*, *5*, e8559.

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, *111*, 317–328.

Sinnemäki, K. (2020). Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach. *Journal of Historical Sociolinguistics*, *6*(2), 20191010.

Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. (2017). Language learning, language use, and the evolution of linguistic variation. *Philosophical Transactions of the Royal Society B*, *372*, 20160051.

Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford University Press.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*, 543–578.