Frequency and stability in language evolution

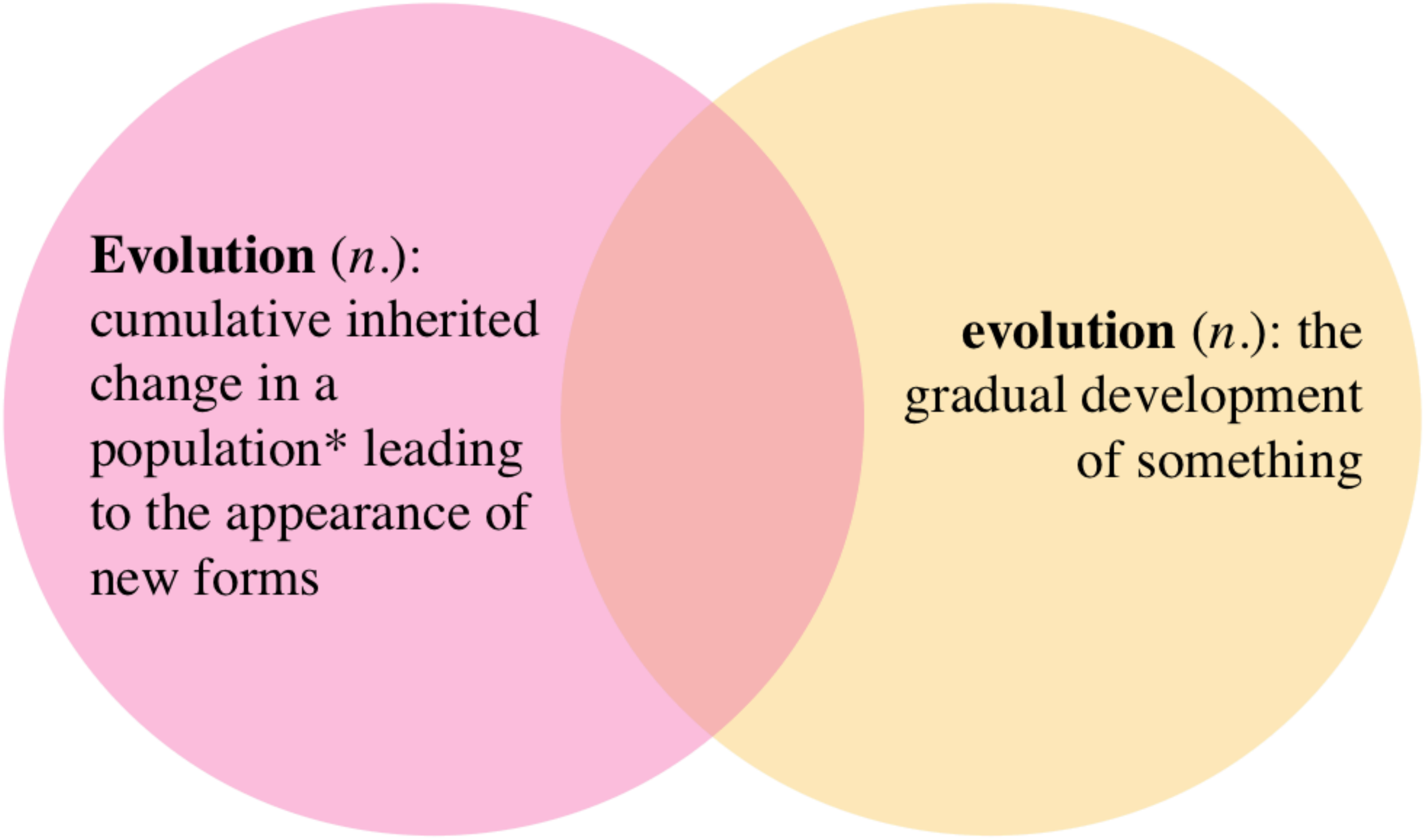
Christine Cuskley | ccuskley.github.io | @nerdpro

Reference list: bit.ly/freqrefs

What's evolution?

Pleyer & Hartmann (2019) identify different senses of language evolution a) origins, b) language change, or c) language universals

My framework: approaching language (and culture, and cognition) with (Darwinian, but not necessarily biological) Evolution in mind



*loosely defined, i.e., a population of variants (cultural or biological)

Frequency effects in language

Age of acquisition (e.g., Ambridge et al., 2015)

Processing (e.g., Ellis, 2002)

Word Length (e.g., Piantadosi et al., 2011)

Regularity (e.g., Cuskley et al., 2014; Fehér et al., 2016)

Regularity: An ambiguity

Sense 1: More likely to adhere to a type-dominant rule (e.g., Cuskley et al., 2014)

— *talk(ed)* takes the regular (type-dominant) past tense form, *be (was/were)* does not, *sneak* is somewhere in between

Sense 2: Absence of unpredictable variation, non-regularised means *free* variation in use between two or more variants (e.g., Fehér et al., 2016)

— *be (was/were)* is regularised (everyone uses the same form), but *sneak* has multiple past tense variants attested *(sneaked/snuck)*

Disambiguating regularity

Regularity: adherence to a type-dominiant rule (morphological, syntactic, phonological)

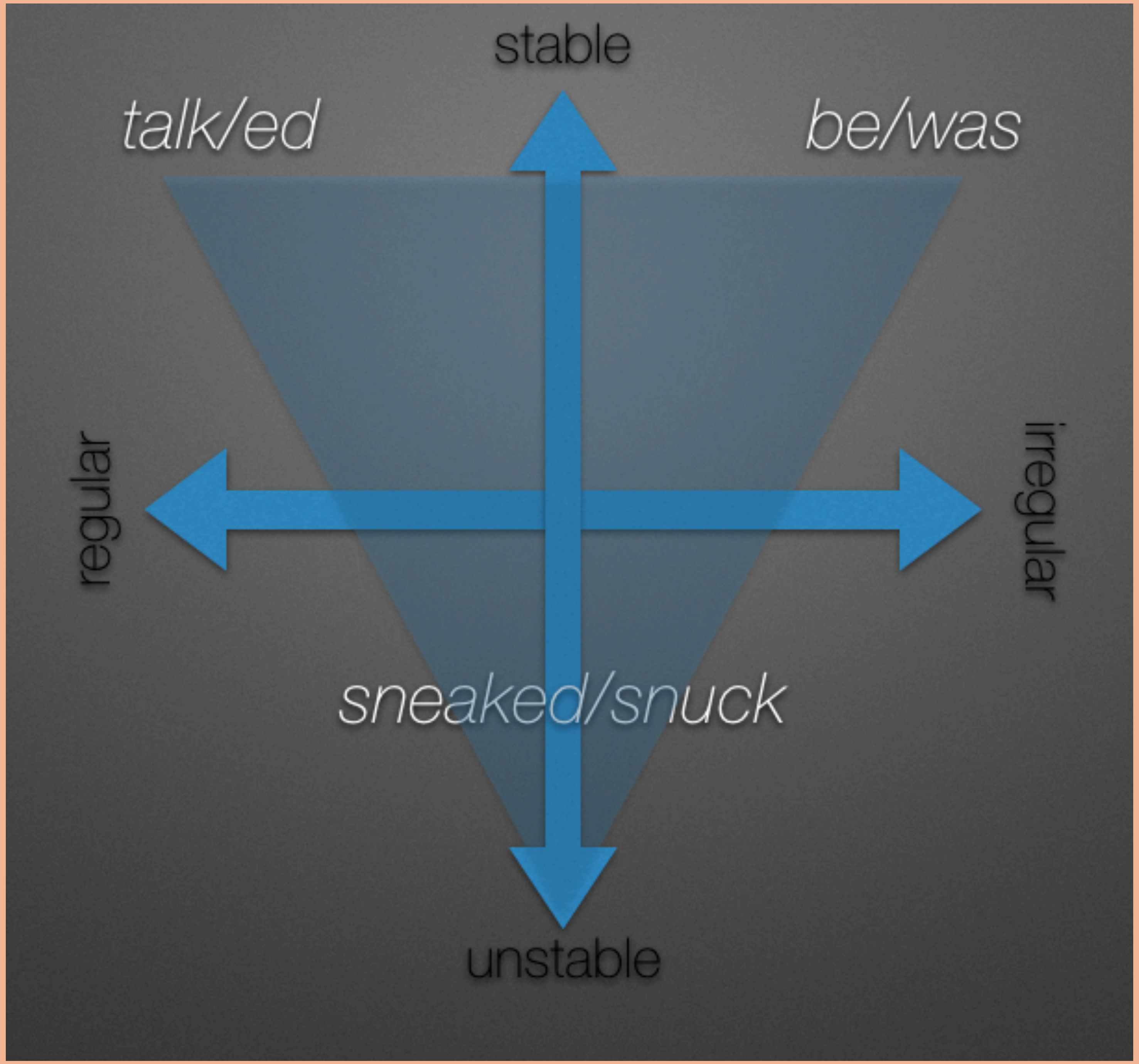
Stability: the convergence on a single variant, as opposed to unconditioned free variation[1]

[1] Note that once variation is conditioned, it is considered stable - i.e., there *is* a single variant for the context in question.

Regular: A more regular system has fewer conditioned variants (e.g., conditioned on stem), ***the rule is more predictable***

Stable: A more stable stem has a minimum of free variants (1), greater consensus has been reached across the population, ***the variant is more predictable***

i.e., both result in lower entropy



Measures

Regularity: Entropy of rule over all stems

$$H_r = \sum_{i=1}^{n} p(x_i) log_2 (\frac{1}{p(x_i)})$$

Stability: Conditional entropy of rule given stem

$$H_s = \sum_{i=1}^{n} p(x_i|L) log_2 (\frac{1}{p(x_i|L)})$$

Frequency and regularity

More frequent items are more likely to be irregular

Mechanism for regularisation: efficiency (i.e., lower entropy) of a dominant rule.

Mechanism for *irregularisation*: phonological similarity (but: requires existing irregulars, can't speak to *emergence* - but see Pijpops et al., 2015)

Lieberman et al., 2007, Cuskley et al., 2014, 2015; Newberry et al., 2017, Wu et al., 2019,

Frequency and stability

Frequent items tend to be more stable, exhibit less free variation

Lower likelihood of "decay" to rule (per Lieberman et al., 2007) for higher frequency irregulars

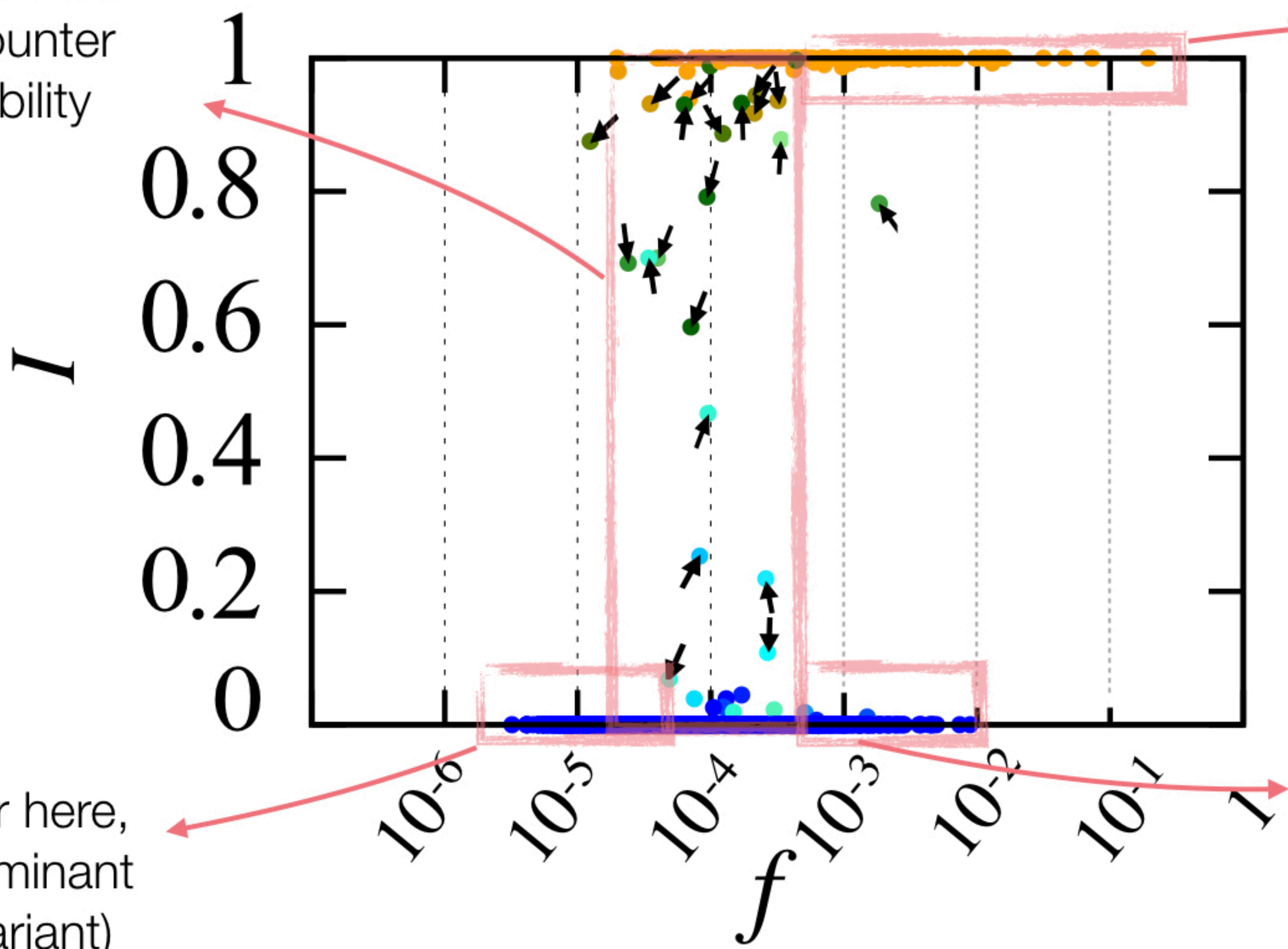
Mechanisms related to memory limitations -> **stabilizing selection for high frequency items**

Ferdinand et al., 2019; Morgan & Levy, 2016; Perfors, 2012

Newcastle University
The British Academy
CENTRE for LANGUAGE EVOLUTION

Forms below a certain frequency encounter potential instability

High frequency, high stability: entrenched forms which “violate” the rule



New forms enter here, take on type-dominant rule (i.e., one variant)

Forms which enter low, then increase in frequency keep regular rule*

Selection to keep the population at one stable, optimal value.

— Likely to occur on traits contributing significantly to function/fitness; Often involves purifying/negative selection *against* deleterious variants.

Zeng, J. et al. (2018). Signatures of negative selection in the genetic architecture of human complext traits. *Nature Genetics*, 50, 746-753

Stabilizing selection in language?

— Lemmas which contribute significantly to language's fitness for communication - i.e., **high frequency variants** - will experience stabilizing selection. This may involve selection *against* the regular variant if another variant has already become communicatively "fit"

Method: The regularity game

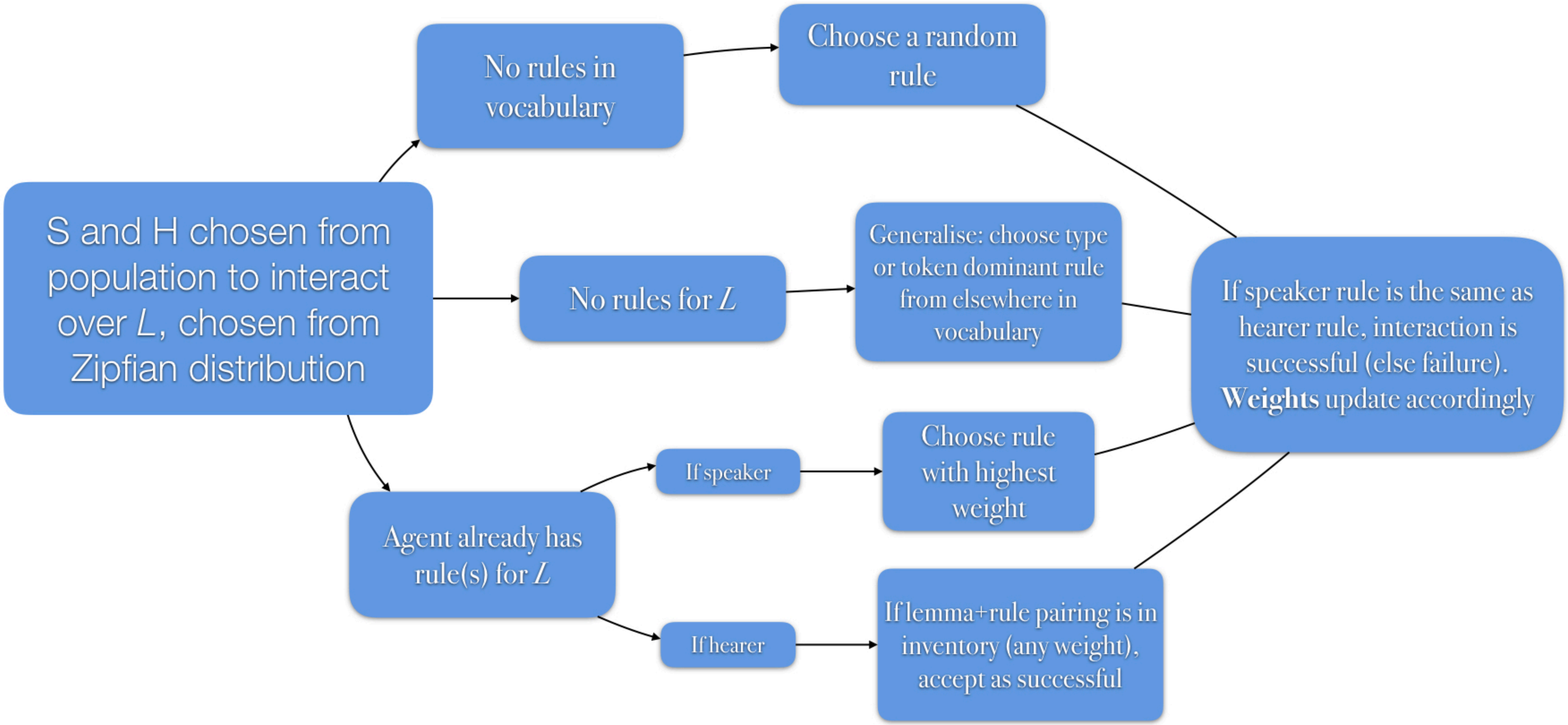
General ABM approach, well described by Ruland et al. (2019)

An adaptation of the Naming Game (e.g., Loreto & Steels, 2001), wherein agents converge on linguistic conventions.

Colaiori et al. (2015) and Cuskley et al., (2017): stability of irregularity at high frequencies

Cuskley et al. (2018): shows emergence of a regular rule and exceptions small and large populations - population growth in particular leads to expected patterns.

Cuskley, Loreto & Kirby (2018). A social approach to rule dynamics using an agent-based model. *Topics in Cognitive Science.*



Measures

We can measure both H_r across the population, and H_s across the population for each lemma, defined by its frequency

We measure "i-language" of agents, but makes little difference to dynamics

Measures

We can also measure selection of the type-dominant rule $s = 1 - w$, where w is relative fitness across variants

1 = selection *for* regular form

0 = selection *against* regular form

Detecting selection for regularity

Allows for detection of selection for or against the type-dominant, regular rule

High frequency lemmas might exhibit selection against the type-dominant regular, *if* another variant already has high fitness, i.e., communicative accuracy.

Contrast different population dynamics

$N = 20$ agents interacting on a complete network for 5000 timesteps, where $1\ t = N_{interactions}$. Agents have memory limitations.

Populations are either
a) closed, i.e., agents live for the entire simulation
b) there is **turnover** at a rate $r = 0.001$ per interaction,
c) there is **growth** at a rate $g = 0.001$ per interaction

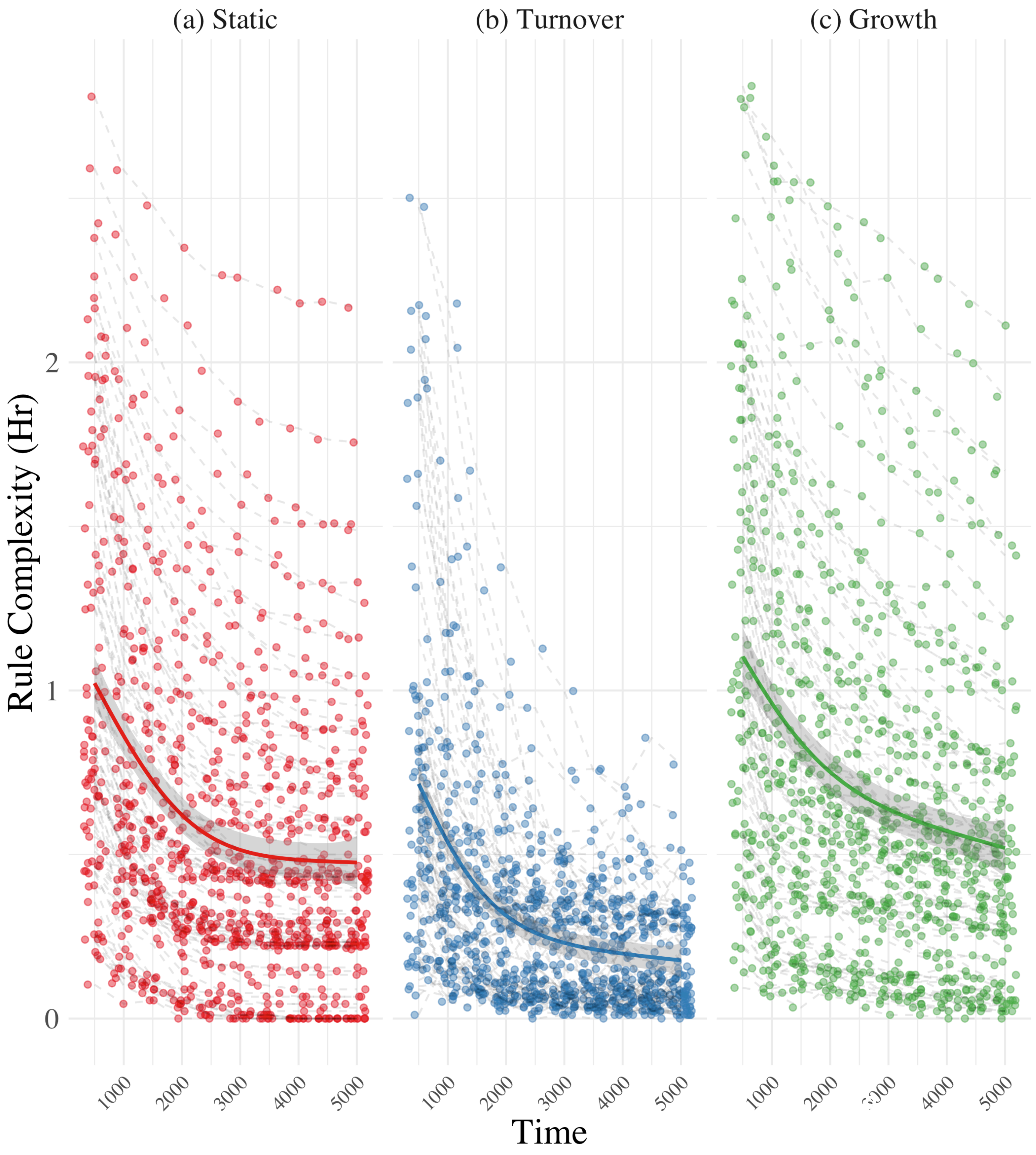
Newcastle University The British Academy CENTRE for LANGUAGE EVOLUTION

Results: Regularity, H_r

Static populations converge and plateau, sometimes on complex languages

Populations with turnover are more likely to converge on completely regular systems

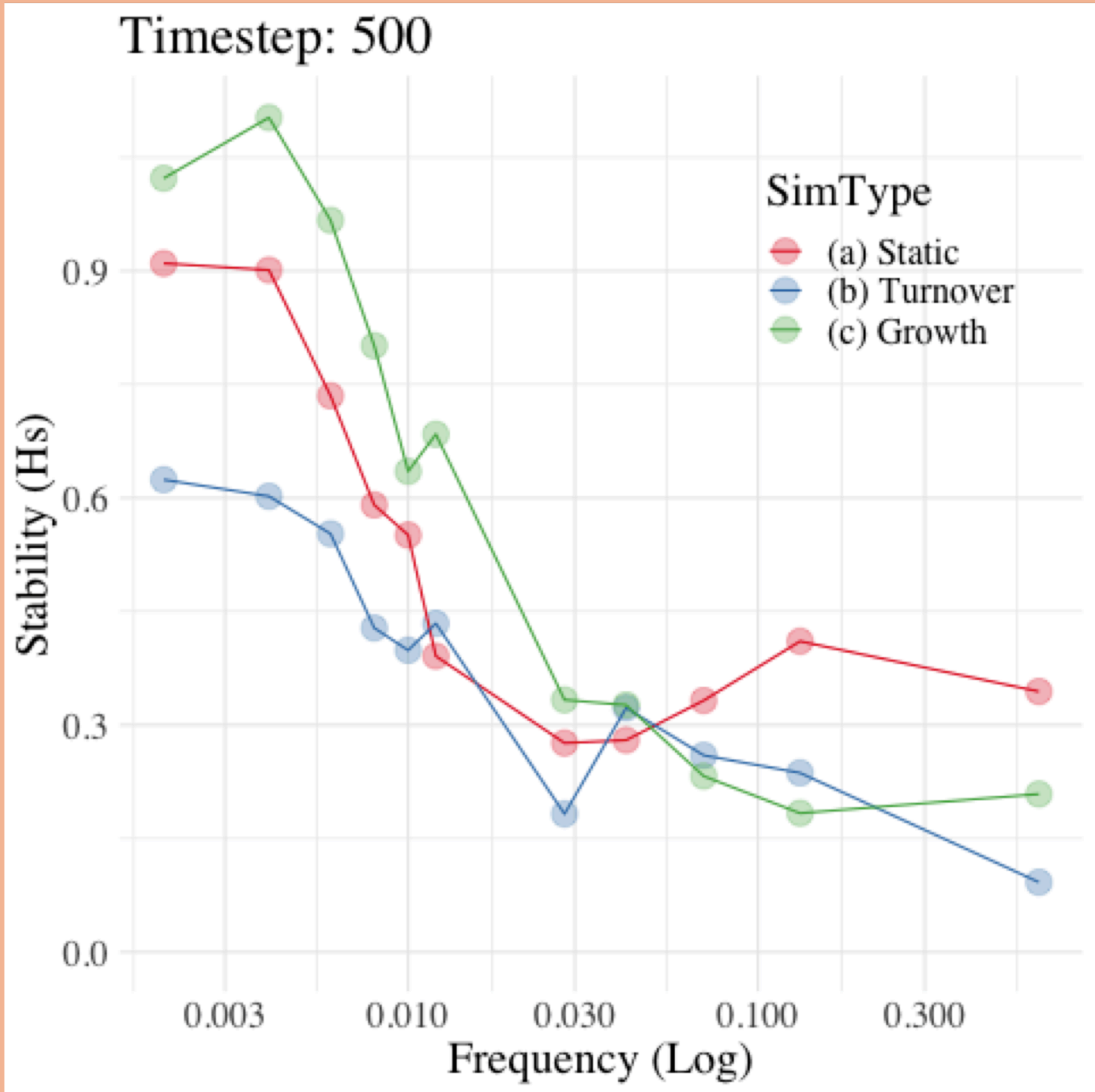
Growing populations show more variation



Results: Stability, H_s

Static population fixates on unexpected stability patterns that don't echo real language data (see Cuskley et al., 2018 for more on this) - conditioned variation via entrenchment in individual agents?

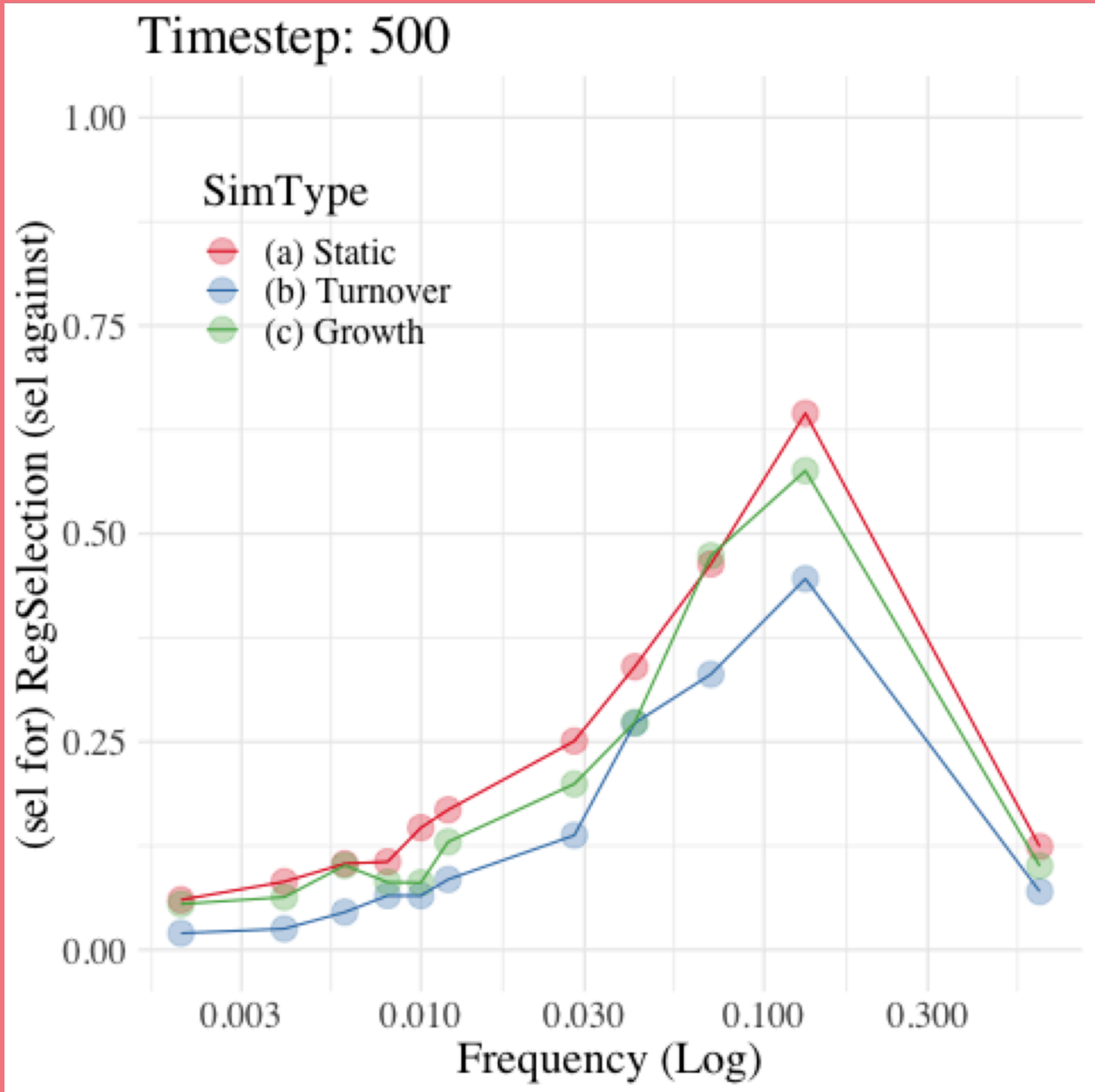
Generally, higher frequencies are more stable (lower H_s), particularly in the growth condition



Results: Selection for regular rule, s_{reg}

Selection against the regular rule is generally stronger in higher frequencies (though highest may *drive* type dominant rule)

In other words, **pressure for stability preserves irregular rules for higher frequences**



Takeaways (I)

Regularity and stability are distinct - terminology exhibits some confusing free variation

We can detect evidence of stabilizing selection for exceptions to rules, which seems to be stronger for higher frequency items

Selection for communicative fitness is a mechanism for both patterns in regularity (e.g., Wu et al., 2019) *and* stability (e.g., Morgan & Levy, 2016)

Takeaways (II)

Growth seems to show most "realistic" overall dynamics (echoing results from Cuskley et al., 2018): variable regularity, higher stability at higher frequencies, selection against regular variant at higher frequencies

Might this point to the importance of starting small (cf. Elman, 1993) on a larger scale?

Ongoing work

L1 and L2 regularity biases and audience design, with Stella Frank and Kenny Smith (Edinburgh)

Analogy and phonological "gangs" as attractors with Stella Frank

Social network structure and population size, with Michael Chimento (MPI Ornithology) and Simon Kirby (Edinburgh)

Future: variable r and g rates, integrating real-world data (e.g., from language contact, Jansson, Parkvall & Strimling, 2015), growing lexicons and rules

Thank you!

Christine Cuskley | ccuskley.github.io | @nerdpro

Reference list: bit.ly/freqrefs

Frequency and stability in language evolution

Christine Cuskley, ccuskley@gmail.com, ccuskley.github.io

Protolang6, September 2019

References

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of child language*, *42*(2), 239–273.

Cuskley, C., Colaiori, F., Castellano, C., Loreto, V., Pugliese, M., & Tria, F. (2015). The adoption of linguistic rules in native and non-native speakers: Evidence from a wug task. *Journal of Memory and Language*, *84*, 205–223.

Cuskley, C., Loreto, V., & Kirby, S. (2018). A social approach to rule dynamics using an agent-based model. *Topics in cognitive science*, *10*(4), 745–758.

Cuskley, C. F., Pugliese, M., Castellano, C., Colaiori, F., Loreto, V., & Tria, F. (2014). Internal and external dynamics in language: Evidence from verb regularity in a historical corpus of english. *PloS one*, *9*(8), e102882.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in second language acquisition*, *24*(2), 143–188.

Fehér, O., Wonnacott, E., & Smith, K. (2016). Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*, *91*, 158–180.

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. *Cognition*, *184*, 53–68.

Jansson, F., Parkvall, M., & Strimling, P. (2015). Modeling the evolution of creoles. *Language Dynamics and Change*, *5*(1), 1–51.

Lieberman, E., Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature*, *449*(7163), 713.

Loreto, V., & Steels, L. (2007). Social dynamics: Emergence of language. *Nature Physics*, *3*(11), 758.

Morgan, E., & Levy, R. (2016). Frequency-dependent regularization in iterated learning. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11).* Online at `http://evolang.org/neworleans/papers/193.html`.

Newberry, M. G., Ahern, C. A., Clark, R., & Plotkin, J. B. (2017). Detecting evolutionary forces in language change. *Nature*, *551*(7679), 223.

Perfors, A. (2012). When do memory limitations lead to regularization? an experimental and computational investigation. *Journal of Memory and Language*, *67*(4), 486–506.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Pijpops, D., Beuls, K., & Van de Velde, F. (2015). The rise of the verbal weak inflection in Germanic: An agent-based model. *Computational Linguistics in the Netherlands Journal*, *5*, 81–102.

Pleyer, M., & Hartmann, S. (2019, Sep). *Constructing a consensus on language evolution? Convergences and differences between biolinguistic and usage-based approaches.* PsyArXiv. Retrieved from `psyarxiv.com/fb2xp` doi: 10.31234/osf.io/fb2xp

Wu, S., Cotterell, R., & O'Donnell, T. J. (2019). Morphological irregularity correlates with frequency. *arXiv preprint arXiv:1906.11483*.

Zeng, J., De Vlaming, R., Wu, Y., Robinson, M. R., Lloyd-Jones, L. R., Yengo, L., ... others (2018). Signatures of negative selection in the genetic architecture of human complex traits. *Nature genetics*, *50*(5), 746.