This article is part of the topic "Sociolinguistic Variation and Cognitive Science," Jean-Pierre Chevrot, Katie Drager and Paul Foulkes (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

This is a preprint, please cite the published version: Cuskley, C., Loreto, V. and Kirby, S. (2018), A Social Approach to Rule Dynamics Using an Agent-Based Model. Topics in Cognitive Science. doi: 10.1111/tops.12327

A social approach to rule dynamics using an agent-based model

Christine Cuskley*,[1], Vittorio Loreto†,[2,3], Simon Kirby‡,[1]

[1] Centre for Language Evolution, University of Edinburgh, United Kingdom
[2] Department of Physics, University of Rome, La Sapienza
[3] Institute for Scientific Interchange Foundation, Turin, Italy
* Corresponding author: ccuskley@gmail.com
† vittorio.loreto@roma1.infn.it
‡ simon@ling.ed.ac.uk

WORD COUNT: 5,645

Abstract

A well-trod debate at the nexus of cognitive science and linguistics, the so-called "past tense debate", has examined how rules and exceptions are individually acquired (Pinker & Ullman, 2002; McClelland & Patterson, 2002). However, this debate focuses primarily on individual mechanisms in learning, saying little about how rules and exceptions function from a sociolinguistic perspective. To remedy this, we use agent-based models to examine how rules and exceptions function across populations. We expand on earlier work by considering how repeated interaction and cultural transmission across speakers affects the dynamics of rules and exceptions in language, measuring linguistic outcomes within a social system rather than focusing individual learning outcomes. We consider how population turnover and growth effect linguistic rule dynamics in large and small populations, showing that this method has considerable potential particularly in probing the mechanisms underlying the linguistic niche hypothesis (Lupyan & Dale, 2010).

1 Introduction

Natural languages generally exhibit rule-based structure: for example, to form the past tense form of a verb in English, you suffix the root with *-ed* (e.g., walk-walked). However, even in natural language, rules have exceptions - the past tense of *teach* is *taught*, not **teached*. A large body of work going back at least three decades examines rules and exceptions primarily from the perspective of individual learning or acquisition: how can a learner accurately acquire both the past tense rule and its exceptions given some minimal data (see e.g., Rumelhart & McClelland, 1986; Pinker & Prince, 1994; Marcus, 1996; Pinker & Ullman, 2002; McClelland & Patterson, 2002)? However, this lively debate says little about why these exceptions persist in a system in the first place (although see O'Donnell, Snedeker, Goodman, & Tenenbaum, 2011).

A smaller, more recent body of work has taken a more social approach to rule dynamics (Colaiori et al., 2015; Cuskley et al., 2017; Pijpops, Beuls, & Van de Velde, 2015; Dale & Lupyan, 2012). Rather than examining how individuals infer or induce rules from a confined set of data, this approach aims broadly to assess how social processes such as interaction and transmission affect rules (and exceptions) shared across a population. This population-focused approach is particularly well-suited to answering sociolinguistic questions, as it can probe phenomena over large timescales and examine different population-wide properties, helping to form new theories about how different kinds of populations affect *linguistic outcomes* (Sankoff, 2008): the characteristics of a language, in this case, its set of rules.

One recent theory about how rules function across populations is the Linguistic Niche Hypothesis (LNH; Lupyan & Dale, 2010). Drawing on data from natural language as well as experiments and models (Lupyan & Dale, 2010; Dale & Lupyan, 2012), the LNH suggests a contrast between *exoteric* and *esoteric* linguistic niches (see also, Wray & Grace, 2007), tying this primarily to population size. Languages with large speaker population sizes occupy a more *exoteric* niche, and tend to have more non-native speakers (e.g., English, Spanish), while languages with smaller population sizes occupy a more *esoteric* niche, and tend to be comprised predominantly of native speakers (e.g., Algonquin). Using data from a wide range of languages, Lupyan and Dale (2010) showed that languages occupying an exoteric niche tend to be less morphologically complex, while those occupying an esoteric niche tend to be more morphologically complex.

The differences between exoteric and esoteric languages are fundamentally sociolinguistic nature, arising from population contact (and often also involving language contact; (Tria, Servedio, Mufwene, & Loreto, 2015)). Lupyan and Dale (2010) suggest that one of the driving mechanisms behind the exoteric/esoteric contrast the presence of more non-native, adult learners in exoteric niches. In other words, adult learners have biases which favour simpler rule systems, thereby driving languages over time towards less morphological complexity . The exoteric/esoteric difference seems to have been confirmed broadly in historical corpus work (Bentz, Verkerk, Kiela, Hill, & Buttery, 2015; Bentz & Winter, 2013): languages with more speakers have simpler morphology. Other historical perspectives in sociolinguistics have also suggested that language contact, in changing the fundamental makeup of speaker populations, results in changes to language structure (Tria et al., 2015; Trudgill, 2011).

However, the precise mechanisms underlying this effect are still not well understood. In other words, while a larger population size (and thus, proportion of adult learners) is one feature of an exoteric language, other features may also act as important mechanisms. For example, experimental work has shown that at least for English, while native speakers in areas with high concentrations of non-native speakers prefer *-ed* forms to irregular forms (Dale & Lupyan, 2012), non-native speakers themselves are significantly more likely to produce irregular forms for novel non-verbs than native speakers (Cuskley et al., 2015). Cuskley et al. (2015) suggest this may mean that while non-native speakers do prefer simpler rule systems, this preference might play out in unexpected ways. For example, lacking exposure to the 'long tail' of low frequency regular verbs, non-native speakers may infer rules from similarly formed high frequency irregulars (e.g., *sting/stung, swing/swung*) and attempt to reproduce these in their output.

While the kinds of learners in a population (which has previously been operationalised as population size) likely plays an important role, exoteric and esoteric niches also have other contrasting characteristics, such as different rates of growth. In other words, the process of new learners entering and growing a speaker population may be as important a factor in the exoteric niche as sheer population size itself. Agent-based models are ideally situated to make a first examination of growth in particular, as this method allows for a controlled contrast between turnover, growth, and relative population size. The resulting findings will provide more targeted hypotheses to explore experimentally and query in natural corpus data.

1.1 Earlier models

While no previous work has examined growth specifically, previous agent-based models have examined rule dynamics in transmission chains of individual learners (see e.g., Kirby, 2001; Reali & Griffiths, 2009). However only recently has a body of work emerged which considers outcomes in larger populations. These approaches are based broadly on repeated language games, in which a speaker-receiver pair chosen from a wider population engages in an interaction over a topic, and this process is repeated, usually at least hundreds if not thousands of times (Baronchelli, Loreto, & Steels, 2008). This gives a time trajectory of a language evolving across a population, which may eventually reach a stable end state. Averaged data over independent runs of such simulations can give a fuller picture of the likelihood of a language reaching a particular state under certain conditions, providing insights that are difficult to arrive at from experimental or natural data alone. This approach can shed new light on sociolinguistic dynamics, allowing for a detailed simulation of both the individual architecture and the processes of social interaction which contribute to the cultural evolution of language (Labov, 2011).

Pijpops et al. (2015) modelled repeated language games in populations (N=1000) of agents examining how the weak (e.g.,-*ed* in English) past tense form might have arisen to become the regular form in Germanic languages. Starting with a predominantly irregular (i.e., *strong*) verb inventory, Pijpops et al. (2015) showed that over time, functional advantages to the weak form gave it sufficient frequency advantages over strong forms to become the type-dominant, regular inflection.

Colaiori et al. (2015; see also, Cuskley et al., 2017) used a model of three-state dynamics wherein agents have either a regular, irregular, or mixed (optionally regular or irregular) rule for a given lemma. This approach examined both gradual population turnover with biased child learners and memory limitations (aka, information loss; Spike, Stadler, Kirby, & Smith, 2016). These factors, combined with a skewed frequency distribution of lemmas, recovered patterns in regularity reminiscent of natural language data (Lieberman, Michel, Jackson, Tang, & Nowak, 2007; Carrol, Svare, & Salmons, 2012; Cuskley et al., 2014), wherein highly frequent forms are more likely to be irregular given certain starting conditions.

Most relevant to questions posed by the LNH, Dale and Lupyan (2012) presented a model which contrasted population size ($N = 5$ vs $N = 20$) and learner type, building in 'child' agents who had a complex preference for morphological marking and 'adult' agents who preferred unmarked forms. Agents engaged in repeated interactions and updated their rule states based on these interactions over time, for 500 steps (modelled after Chater, Reali, & Christiansen, 2009). In closed populations of child agents, they found that particularly larger population sizes gravitated towards simpler unmarked morphology, while smaller populations showed a wider distribution of language types. However, when population turnover was introduced - some agents were replaced with naive "child" agents - marked morphology became an attractor for small populations, while large populations still tended towards unmarked forms. When a small core population of 5 child learners was hit with an abrupt influx of 15 adult learners halfway through the simulation, the language of the core 5 learners started to prefer unmarked forms.

In the current paper, we aim to examine the relative contributions of population size, turnover, and growth to rule dynamics in populations of interacting agents. Like the models reviewed above, these models are not meant to provide a detailed replica of any particular population or natural language. Rather, the goal is to use the model to create a controlled "petri dish" in which we can examine the effects of population size, turnover, and growth on morphological systems.

2 Current Model

To accomplish this, we extend the "regularity game" from Cuskley et al., 2017 to include multiple inflections, different population sizes, and population growth. The inclusion of multiple inflections allows agents agents make generalizations across their vocabulary, extending inflection strategies from one lemma to another. Most earlier models often only considered two possible rule states, regular and irregular(Colaiori et al., 2015; Cuskley et al., 2017) or marked and unmarked (Dale & Lupyan, 2012), which are pre-defined (although see Pijpops et al., 2015). However, in natural languages, the set of rule states is rarely so straightforward. Thus, in the current model, agents have 12 rules to choose from, meaning a vocabulary could in theory emerge which has several competing rules, each potentially as 'regular' (in that it applies to the same number of verb types) as the next.

How an agent chooses a rule to apply to a given lemma is motivated by recent experimental data. Cuskley et al. (2015) showed that native and non-native speakers inflect novel verbs irregularly at different rates, with non-natives being significantly more likely to provide irregular forms. However, for both native and non-native speakers, the irregular forms they produced were not random: they were largely identifiable as instances of irregular classes (e.g., change an internal vowel, as in *spit-spat*). The tendency to inflect irregularly was predicted significantly not only by binary nativeness. Among non-natives, age of acquisition and self-reported proficiency predicted lower irregluarization rates: participants who acquired English earlier, and were more proficient in the language, were less like to provide irregular forms. Comparison with corpus data from Cuskley et al. (2014) suggests that less proficient speakers may be inferring and extending rules from highly token-frequent irregular verbs, given that they lack the same exposure to the long tail of low frequency regulars available to more proficient speakers.

To model how changes in proficiency mirror changes in inflection strategies, individual agents in the current model mature over time as they learn the language, adjusting their generalisation strategy accordingly. Less proficient agents *token generalize* by looking across their vocabulary and extending the rule which has been used most frequently *across all tokens of any type*. More proficient agents *type generalize* across their vocabulary by extending the rule which applies to the most types in their vocabulary.

We contrast populations under two different demographic conditions, considering small (N=20) and large (N=100) population sizes for each. First, we examine populations with gradual turnover: population size remains stable, but new learners replace existing agents with a certain rate, r. We contrast this with population growth, wherein the population expands over time, with new learners entering and integrating with existing agents at a certain rate, g.

To avoid confusion surrounding the terms 'regular' and 'regularization', these terms will be reserved for rule regularity, where regular means a rule which applies to a majority of types (i.e., is type-dominant), and regularization refers to the movement of a particular item towards such a rule. Regularization in the sense of reduction of variation (e.g., Feher, Wonnacott, & Smith, 2016) will be framed in terms of *stability*: if a lemma has virtually no interspeaker variation, it can be said to be highly stable, whereas a lemma with high interspeaker variation is unstable.

Both regularity and stability are relevant measures. The LNH would predict that exoteric languages move towards rule regularity, eliminating idiosyncratic exceptions (and thus, morphological complexity), while esoteric languages are more likely to sustain irregularity (and maintain morphological complexity). On the other hand, stability of individual lemmas is also an important measure to detecting when and why dominant rules might begin to destabilize and shift, as was the case with the weak form becoming regular in Germanic languages (Pijpops et al., 2015). Additionally, there is good data regarding how both regularity and stability function in natural language. In terms of regularity, highly frequent items tend to be irregular, while low frequency items tend to be regular (Cuskley et al., 2014, Lieberman et al., 2007). In terms of stability, frequent items tend to be highly stable, while rare items are more unstable (Morgan & Levy, 2016).

2.1 Model details

2.1.1 Vocabulary and rules

The vocabulary in the current model is fixed in type size and frequency; in other words, there is a fixed number of root lemmas (28), each with a frequency that remains fixed over the course of the simulation. Frequencies were generated using Python in a Zipfian distribution (using the Numpy random.zipf function, with an exponent of 2, roughly characteristic of English texts;

Corral, Boleda, & Ferrer-i Cancho, 2015). Most earlier variants on the language game approach have demonstrated mechanisms by which populations converge on shared naming conventions (Baronchelli et al., 2008). In the current model, it is assumed that agents share conventions for root lemmas, but must negotiate and learn conventions for inflecting these lemmas.

In terms of inflectional rules which can be applied to each lemma, the system consists of 12 (functionally equivalent) potential rules. The number of rules available to agents is more than the simple regular and irregular rules implemented in Cuskley et al. (2017) and Colaiori et al. (2015), but the rule set is still finite.

2.1.2 Agent architecture

Each population starts with N agents, and each agent starts with no interaction history. Each agent has an entry for each lemma including a counter, f, which tracks the total number of interactions with the lemma over the agent's lifetime. Each interaction with a lemma will involve inflectional rules. Thus, for each lemma, each agent stores a set of inflections. For each lemma-inflection pairing, the agent stores:

1. f_i, the total number of interactions with that inflection and lemma;
2. w, a weight for the lemma-inflection pair defined as the number of successful interactions with the lemma-inflection pair divided by the total number of interactions with the lemma-inflection pair, f_i;
3. t_{last}, the time step of the last interaction with that lemma-inflection combination

Agents have temporally constrained memory for each lemma-inflection combination. Memory limitations are crucial to recovering realistic frequency-dependent rule dynamics found in natural language (e.g., in corpora; Cuskley et al., 2017), and have been identified as an essential factor in the emergence of learned signalling in agent-based models more generally (Spike et al., 2016). This memory constraint is implemented as a deterministic loss of a lemma-inflection pairing after a set period of time: if a pairing has not been encountered within a specific time window, $d = 100$ (regardless of the pairing's weight, w), the pairing will be forgotten. Prior to an agent engaging in an interaction, the time window elapsed since the last encounter with the lemma inflection pairing is calculated as the current t minus t_{last}.

Each run of a simulation goes for $t = 10000$ time steps, where a single timestep is defined by N interactions, for a total of $t \times N$ interactions in each run, and 100 independent runs.

2.1.3 Interaction rules

At each interaction, two random agents are chosen: one as the speaker, S, and one as the receiver, R. A lemma is chosen as the topic of the interaction; the likelihood of choosing a given lemma is defined by its frequency. If S has inflections stored for the topic lemma, they choose the inflection with the highest weight (figure 1, A). If S has no inflections stored for the topic lemma, their first step is to look across the rest of their vocabulary. If they are a high proficiency speaker (i.e., have encountered $k \geq 1500$ tokens), they will engage in type generalisation (figure 1, D). However, if they are a new learner (i.e., have encouktered $k < 1500$ tokens), they will engage in token generalisation (figure 1, C). If the agent has no inflections in their vocabulary, they will choose a random inflection from the pre-defined set of 12 (figure 1,B).

If R's inventory of lemma-inflection pairings contains the pairing used by S, the interaction is successful. Otherwise, R also generates a lemma-inflection pairing using processes B, C or D from Figure 1. If the inflection matches the one produced by S, the interaction is successful, otherwise it is a failure. This asymmetry - wherein the R finds any inflection they have encountered before acceptable - mirrors real-world usage to some extent. For example, a speaker may use the past tense form *sneaked* uniformly, but still has no problem processing the form *snuck* as a listener. Communication is successful if the listener can, by some route, interpret the correct content of what the speaker says, even if it is not identical to how the listener would have phrased it.

At the end of the interaction, both S and R update the f_i, w, and t_{last} for the lemma-inflection pairing based on the interaction; if the interaction was successful, the w will increase for both agents, while failure will decrease w for both. Figure 1 outlines the details of each interaction graphically.

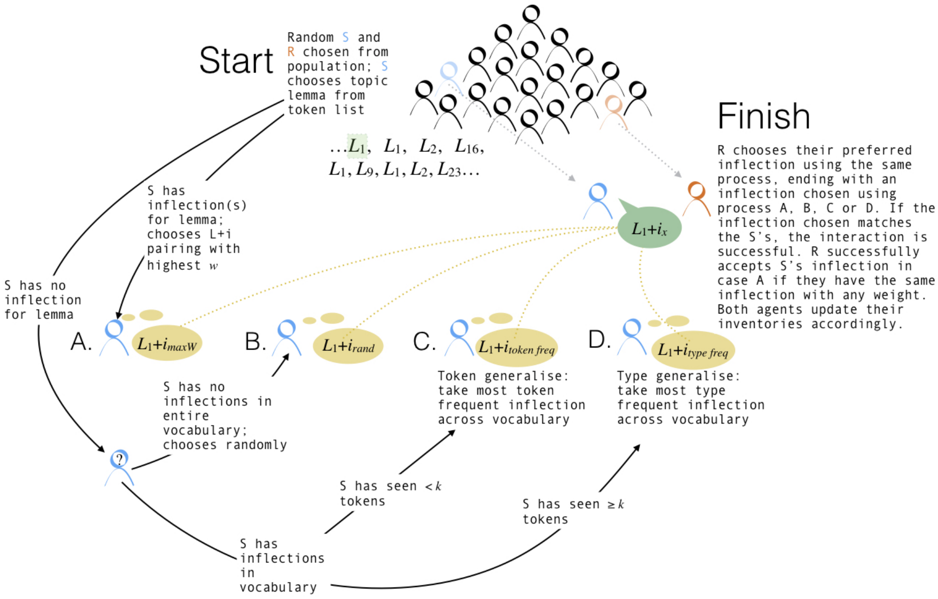

Figure 1: The process for a single interaction (shown for $N = 20$).

2.1.4 Measuring the system

At each timestep, the entropy of the system is measured in two different ways to quantify regularity and stability. First, the regularity of the system is measured using overall entropy of across the vocabulary, H_v, is measured, shown in equation 1 below. $p(x_i)$ represents the probability of a particular inflection i, when the highest weighted inflection for each lemma is chosen across the population:

$$H_v = \sum_{i=1}^{12} p(x_i) log_2(\frac{1}{p(x_i)}) \tag{1}$$

H_v is low where the rule for any given type is predictable because the language is highly regular, and the inflection varies little across vocabulary items in the population.

Second, the stability of each vocabulary item is measured as the conditional entropy of each lemma in the vocabulary, H_l:

$$H_l = \sum_{i=1}^{12} p(x_i|L) log_2(\frac{1}{p(x_i|L)}) \tag{2}$$

Given a particular lemma L, $p(x_i|L)$ is the probability of inflection i when the highest weighted inflection for L is chosen across each agent in the population. If H_l is low for a given lemma, there is little variation between speakers in the rule used for that lemma.

3 Analysis and Results

3.1 Turnover

At each interaction, a random agent is chosen and replaced with a new learner agent with a probability of $r = 0.001$, meaning that every ten time steps, the population will have experienced roughly 1% turnover. Conventional rules quickly emerge, with each run having a dominant "regular" rule which holds sway over most types. The regular form holds sway over slightly more types in the large population (on average covering $\approx 95\%$ of types), while the regular form covers $\approx 87\%$ of types in the small population. Figure 2a shows that smaller populations have a much broader distribution of end state H_v values than large populations, indicating that rules are less predictable in smaller populations. In contrast, languages of large populations show a sharp peak at low values of H_v, indicating more predictable rule systems.

However, paradoxically, small populations actually have fewer rules on average than large populations (Figure 2c). The constant influx of new learners also means a constant, low-level influx of new inflections; since turnover is proportional to population size, this influx is greater in larger populations. Yet, although larger populations may have more inflections, each minority inflection is highly irregular in that it holds sway over very few types (e.g., *do* $\rightarrow$ *did*), leaving more types overall for the dominant regular rule. However, among smaller populations, each minority inflection is more likely to be influential over several types (e.g., *breed* $\rightarrow$ *bred*, *feed* $\rightarrow$ *fed*, etc.). This results in higher inflectional unpredictability in smaller populations, while in large populations any given type is more likely to be governed by the regular rule.

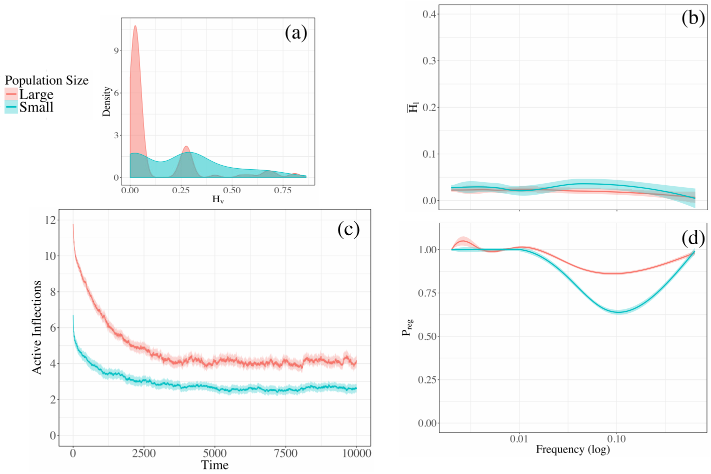

Figure 2: (a) Distribution of end-state values of H_v over 100 runs. (b) Mean end-state values of H_l by lemma frequency. (c) Number of active inflections across the population over time for large and small population sizes (averaged over 100 runs). (d) Probability, at $t = 10000$, of a given lemma (as defined by its frequency) of having its most common inflection measured across the population be the type-dominant "regular" inflection.

In terms of individual lemmas, the highest frequency tends towards the regular rule (Figure 2d), while mid range frequencies are least likely to adhere to the dominant rule. The value of H_l is close to 0 for all frequencies in both population sizes (Figure 2b), indicating that despite new learners constantly replacing more proficient speakers in the population, there is low interspeaker variation. This may be because turnover combines with memory constraints to create greater overall information loss, which can bolster the emergence of shared conventions (Spike et al., 2016).

3.2 Growth

In the case of growth, the core population remains, but at each interaction, a new learner is added with a probability $g = 0.001$. This means that by $t = 10000$, the population has grown by approximately ten fold, meaning the small population grows to $N \approx 220$, while the large population grows to $N \approx 1100$.

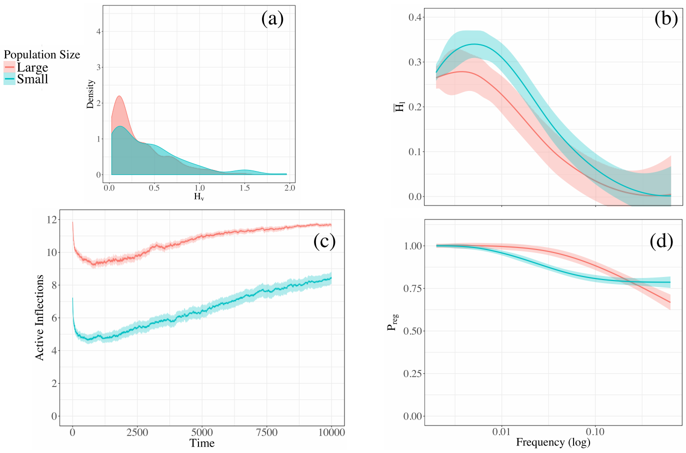

Figure 3: (a) Mean values of H_v (over 100 runs) over time for large and small population sizes. (b) Mean number of active inflections across the population over time for large and small population sizes. (c) Proportion of types using a given inflection over time for a small population (top), and a large population (bottom), averaged over all 100 runs.

The growth case shows a distribution of end state H_v values reminiscent of the turnover case, although less extreme. Larger populations skew toward lower values of H_v (Figure 3a), but the divide between large and small populations is less marked. This indicates that a stable core population of proficient learners can help to support more predictable rule systems in general, regardless of relative population size.

Unlike with turnover, the number of active inflections in the population over time increases (after an initial decrease) for both population sizes (Figure 3c). An inflection is considered active if it is the highest weighted inflection for at least one lemma in at least one agent in the population. This is likely due to the constant influx of new learners, who in the early stages of acquiring rules are constantly introducing new inflections, particularly for low frequency items. As with turnover, absolute number of active inflections is considerably higher for the large population. Since growth is proportional to population size, at any given moment there are more new learners in the large population than in the small population, making the total number of active inflections is higher. Although this may seem trivial, this does mean that the potential for rule variation to enter the language is higher in the larger population. As with turnover, although large populations have more rules in absolute terms, the regular form holds sway over more types in the large population (on average covering $\approx 87\%$ of types), while the regular form covers only $\approx 81\%$ of types in the small population.

Growth shows a pattern more familiar from natural language than turnover in terms of the relationship between regularity, stability, and frequency. Low frequency items are less stable (i.e., show more variation across the population, with a higher H_l) than high frequency items (Figure 3b), and high frequency items are less likely to be regular, while low frequencies adopt the type-dominant rule (Figure 3d).

4 Discussion

In all populations, a type-dominant "regular" rule quickly emerges, but there are some notable differences between how dominant this rule is across different population sizes, and how realistic other frequency related properties of the system are. This confirms what earlier work on the LNH has found: that relative population size is an important mechanism underlying variation in rule dynamics.

Both turnover and growth fit the predictions of the LNH in terms of population size: larger populations have more predictable rule systems than small populations. However, crucially, large populations actually have more rules in absolute terms. In other words, while the likelihood of any given lemma adhering to the regular rule is higher in large populations, the total number of rules used across the system is also higher. This fits with some extent to the picture which emerges from experimental results from Cuskley et al. (2015). In an attempt to infer and generalise rules where there may be none, less proficient learners tend to generalise over minority (i.e., irregular) rules, collapsing the rule system and increasing overall rule predictability. As a large body of empirical work in sociolinguistics has suggested, this indicates that the effects of changes in speaker populations are complex (Labov, 2011; Trudgill, 2011).

Turnover shows strange by-lemma patterns for stability and regularity: unlike in real languages (Cuskley et al., 2014), there is low stability and high regularity for the highest frequency lemmas. In terms of regularity, this pattern is likely caused by agents fixating early on an inflection for the highest frequency; this becomes entrenched and quickly transmits to new learners since the population size remains fixed. However, in the growth case, as new learners overwhelm the original core population, inflections of lower frequency items destabilise to the type-dominant form as type generalization becomes the dominant strategy. A similar explanation underlies the stability results: turnover makes for convergence on stable, shared inflections for all lemmas (i.e., low H_v regardless of frequency). Although the core populations are constantly churning in new learners in the case of turnover, the fixed size of the population means that convergence on a stable form is more likely. Unlike turnover, growth recovers the frequency-regularity and frequency-stability patterns found in natural language (Cuskley et al., 2014) in addition to relative population size differences: high frequency lemmas are the most stable across the population, but also the least likely to adhere to the type-dominant rule (i.e., most likely to be irregular). This indicates that growth plays an important role in how morphological complexity differs between exoteric and esoteric niches, providing a specific hypothesis which can now be tested with more traditional experimental or corpus linguistic methodologies.

This shows that relative population size is only one of many contributing factors to different rule dynamics: rates of turnover, and particularly growth, are likely candidates for other important mechanisms underlying the exoteric/esoteric divide. These models can help to direct future efforts to explore mechanisms underlying exoteric and esoteric niches using both experiments and corpora. For example, artificial language learning experiments which take a group approach might consider how the process of group growth effects the structure of languages in more specific detail, and corpus investigations might target particular bursts of speaker population growth in languages where this is well documented (e.g., relatively recent growth of English and Spanish). The current results show that growth and turnover are important forces in rule dynamics; future work should focus on how different rates of growth and turnover - and rates which vary over time - may result in different linguistic outcomes.

Our models confirm that new learners have an important role in shaping the dynamics of rule systems in language, and that large scale social shifts in a population such as growth and turnover play an important role. Overall, the model highlights that the specific mechanisms by which large scale social shifts affect morphological are still under-explored. Furthermore, they show agent-based models are an increasingly valuable tool for taking a large-scale, controlled, perspective on key questions in sociolinguistics. In addition to giving a better picture of exactly how different rates of growth and turnover effect rule systems, this approach has the potential to query how more complex nuances of social network structure effect rule dynamics. These models provide a perspective is impossible to take with experimental or corpus methods alone (Conte & Paolucci, 2014), making population based models an ideal complement to a better understanding of language as a complex social system.

References

Baronchelli, A., Loreto, V., & Steels, L. (2008). In-depth analysis of the naming game dynamics: the homogenous mixing case. *International Journal of Modern Physics C*, *19*(5), 785-812.

Bentz, C., Verkerk, A., Kiela, D., Hill, F., & Buttery, P. (2015). Adaptive communication: Languages with more non-native speakers tend to have fewer word forms. *PLoS ONE*, *10*(6), e0128254.

Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, *3*, 1-27.

Carrol, R., Svare, R., & Salmons, J. (2012). Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics*, *2*(2), 153-172.

Chater, N., Reali, F., & Christiansen, M. (2009). Restriction son biological adapataiton in language evolution. *Proceedings of the National Academy of Sciences*, *106*, 1015.

Colaiori, F., Castellano, C., Cuskley, C., Loreto, V., Pugliese, M., & Tria, F. (2015, Jan). General three-state model with biased population replacement: Analytical solution and application to language dynamics. *Phys. Rev. E*, *91*, 012808.

Conte, R., & Paolucci, M. (2014). On agent-based modeling and computational social science. *Frontiers in Psychology*, *5*, 668.

Corral, A., Boleda, G., & Ferrer-i Cancho, R. (2015). Zipf's law for word frequencies: Word forms versus lemmas in long texts. *PLoS ONE*, *10*(7), e0129031.

Cuskley, C., Castellano, C., Colaiori, F., Loreto, V., Pugliese, M., & Tria, F. (2017). The regularity game: Investigating linguistic rule dynamics in a population of interacting agents. *Cognition*, *159*, 25–32.

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., Pugliese, M., Castellano, C., Colaiori, F., Loreto, V., & Tria, F. (2014, 08). Internal and external dynamics in language: Evidence from verb regularity in a historical corpus of english. *PLoS ONE*, *9*(8), e102882.

Dale, R., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15*(03n04), 1150017.

Feher, 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.

Kirby, S. (2001). Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, *5*(2), 102-110.

Labov, W. (2011). *Principles of linguistic change, cognitive and cultural factors (vol. 3).* John Wiley and Sons.

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

Lupyan, G., & Dale, R. (2010, 01). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559.

Marcus, G. F. (1996). Why do children say “breaked”? *Current Directions in Psychological Science*, *5*, 81-85.

McClelland, J. L., & Patterson, K. (2002). ’words *or* rules’ cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences*, *6*(11), 464-645.

Morgan, E., & Levy, R. (2016). Frequency-dependent regularization in iterated learning. In S. Roberts, C. Cuskley, L. McCrohon, L. Barcelo-Coblijn, O. Feher, & T. Verhoef (Eds.), *The evolution of language: proceedings of the 11th international conference (evolang11).*

O’Donnell, T. J., Snedeker, J., Goodman, N., & Tenenbaum, J. (2011). Productivity and reuse in language. In *Proceedings of the thirty-third annual conference of the Cognitive Science Society.*

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.

Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima & R. L. (Eds.), *The reality of linguistic rules* (p. 321-352). Philadelphia: John Benjamins.

Pinker, S., & Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456-463.

Reali, F., & Griffiths, T. (2009). The evolution of frequency distributions: Relating regularisation to inductive biases through iterated learning. *Cognition*, *111*(3), 217-328.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of english verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing (vol 2): Psychological and biological models* (p. 216-271). Cambridge: MIT Press.

Sankoff, G. (2008). Linguistic outcomes of language contact. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (chap. 25). Blackwell.

Spike, M., Stadler, K., Kirby, S., & Smith, K. (2016). Minimal requirements for the emergence of learned signalling. *Cogntive Science*, doi:10.1111/cogs.12351.

Tria, F., Servedio, V. D., Mufwene, S. S., & Loreto, V. (2015). Modeling the emergence of contact languages. *PLoS ONE*, *10*(4), p.e0120771.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford, UK: Oxford University Press.

Wray, A., & Grace, G. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543-578.