Momentum in language change: a model of self-actuating s-shaped curves

June 13, 2016

Abstract

Like other socially transmitted traits, human languages undergo cultural evolution. While humans can replicate linguistic conventions to a high degree of fidelity, sometimes established conventions get replaced by new variants, with the gradual replacement following the trajectory of an *s-shaped curve.* Although previous modelling work suggests that only a bias favouring the replication of new linguistic variants can reliably reproduce the dynamics observed in language change, the source of this bias is still debated. In this paper we compare previous accounts with a *momentum-based selection* account of language change, a replicator-neutral model where the popularity of a variant is modulated by its *momentum*, i.e. its *change in frequency of use* in the recent past. We present results from a multi-agent model that are characteristic of language change, in particular by exhibiting spontaneously generated s-shaped transitions that do not require externally triggered actuation. We discuss several empirical questions raised by our model, pertaining to both momentum-based selection as well as other biases and pressures which have been suggested to influence language change.

keywords: language change; cultural evolution; momentum; age vectors; s-shaped curves

1 Introduction

Human languages are a prime example of a culturally evolving trait: they are made up of socially learned conventions which are constantly being replicated, and exhibit great diversity across the globe (Evans and Levinson, 2009). Important aspects of the dynamics of language change are well-understood. Firstly, language change is *sporadic* (de Saussure, 1959; Labov, 2001). Of all the conventions that make up a single language, at any given point most of them are not undergoing change, but are replicated faithfully, from basic word order patterns down to the pronounciation details of individual words. Languages are transmitted robustly over many generations, a necessary requirement for their use as a tool for communication (Lewis and Laland, 2012). Secondly, when a convention *does* change, individuals will gradually replace an established variant with a new variant. This gradual replacement exhibits directed transitions in the form of *s-shaped curves* such as the one shown in Fig. 1, akin to the patterns of logistic growth found in biological evolution (Bailey, 1973; Altmann et al., 1983; Kroch, 1989; Denison, 2003; Blythe and Croft, 2012)[1]. This similarity to the signature of adaptive selection in biology is puzzling (Labov, 2001, ch.1). Linguistic conventions are *arbitrary*, which means we should not expect an inherent advantage in particular linguistic variants, such as which basic word order is used by a language, or how exactly a distinctive phonemic segment is pronounced (as long as it maintains its contrastive function). How and why would an entire population of speakers go about replacing an existing convention with a different one "to say the same thing"?

[Figure 1 about here.]

1.1 Language-internal accounts

In order to explain *why* languages change, many studies have attempted to pin down the causes of individual changes by systematically comparing the states of the languages prior to and after a change (Hockett, 1965; McMahon, 1994). While many of the earliest such studies would attribute change to the gradual accumulation of performance

[1] While the notion of 's-shaped curves' is notoriously ill-defined, for the purposes of this paper it will suffice to use Blythe and Croft's definition as any directed trajectory that does not feature "large fluctuations and a tendency for an upward or downward trend to reverse one or more times before an innovative variant goes extinct or wins out" (2012, p.285).

and transmission errors alone (e.g. Jespersen, 1922; Hockett, 1958), the generativist paradigm with its focus on the language acquisition device shifted the attention to child-based language change. Studies of language change in the generative tradition trace changes back to the reordering or simplification of a language's grammatical rules (Kiparsky, 1968; Wang, 1969; Bailey, 1973; Lass, 1980; Vennemann, 1983), typically assumed to be due to children's reanalysis of linguistic parameters based on their limited linguistic input (see Kroch (2001) and Foulkes and Vihman (2013) for reviews concerning syntactic and phonological change, respectively). Rather than characterising change as the result of imperfect transmission, a more recent strand of research regards language as a *complex adaptive system* which evolves to fulfill the communicative needs of its speakers, while at the same time adapting to the constraints imposed by their learning mechanisms (Kirby, 1999; Steels, 2000; Griffiths and Kalish, 2007; Beckner et al., 2009).

What unites these *language-internal* accounts is that they all rely on a qualitative difference between the language states prior to and after the change. This difference can be based on a variety of factors, such as the languages' expressivity, processing efficiency, or simply their stability with respect to error-prone language acquisition. Within historical and variationist linguistics such explanations of language change have long been criticised on the basis that they *overpredict* change (de Saussure, 1959; Greenberg, 1959; Weinreich et al., 1968; Lass, 1980; Ohala, 1989; Croft, 2000; Labov, 2001; Winter-Froemel, 2008). In their seminal paper, Weinreich et al. succinctly summarised the issue and coined it the *actuation problem*: "Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?" (Weinreich et al., 1968, p.102).

In other words, language-internal pressures by themselves do not account for the *sporadicity* of language change: many non-adaptive or suboptimal structures that are claimed to have been selected against in one language will happily persist in other languages – and when they finally do change, language-internal accounts often offer no explanation of what triggered the *actuation* of the change (de Saussure, 1959; Postal, 1968; Ohala, 1993). While language-internal factors offer good predictions of *which* changes are more likely to occur than others (Jaeger and Tily, 2010; Wedel et al., 2013), they do not explain *when* or *why* the stable transmission of language suddenly caves under functional pressures when it does. To account for the sporadic nature of language change, many have argued that it is not enough to rely on intra-linguistic factors alone.

1.2 Social accounts

Sociolinguistic research of the past five decades has shown that innovations do not spread uniformly across a given speech community, but that the progression of change is stratified based on factors such as a speaker's age, ethnicity, or socio-economic status (Foulkes and Docherty, 2006; Tagliamonte, 2012). *Social accounts* hold that social features of linguistic variants, rather than their inherent linguistic character, are responsible for driving language change (Sturtevant, 1947; Croft, 2000; Labov, 2001; Croft, 2006). Social accounts of language change are *evolutionary* in nature: they decouple the generation of *variation* from the process of *selection* which leads to the diffusion of variants through a speech community. The underlying mechanisms, however, are very different from biological evolution. While the generation of new variants is assumed to be driven by linguistic or functional factors, social accounts attribute the ultimate *selection* of variants to extra-linguistic social factors (Ohala, 1989; Croft, 2000; Labov, 2001; Stevens and Harrington, 2013). The 'division of labour' between language-internal and social pressures in this approach can simultaneously account for the arbitrary adoption of one linguistic convention from the pool of variants over another, while at the same time explaining the crosslinguistic distribution of linguistic features which reflect functional pressures.

Recent work on a mathematical model of language change suggests that only the presence of a bias which favours the replication of a newly incoming variant can reliably reproduce the s-shaped transitions observed in language change (Blythe and Croft, 2012). While this mechanism, known as *replicator selection*, is in principle also compatible with language-internal biases, the authors eschew this conclusion. In line with social accounts of language change they conclude instead that it is the *social prestige* of a new variant that is responsible for its preferential replication. Importantly, the sociolinguistic use of the term *prestige* actually refers to a *content bias*: rather than preferentially copying variants used by prestigious individuals, *prestige* is simply another name for a bias that, while social in origin, is actually inherent to the linguistic variant (Sturtevant, 1947; Labov, 2001). Crucially, social accounts do not solve the underlying logical problem of how a population would agree on the selection of a new variant if there is no objective advantage to that variant. The choice of the population to attach preferential prestige to some variant is as arbitrary and requires just as much explanation as the population's increased use of one linguistic variant over another. Because variant prestige is not accounted for within the theory (Meillet, 1921; Labov, 2001) and can only be attributed post-hoc (Sankoff, 1988; Trudgill, 2004; Maegaard et al., 2013), social accounts are typically unable to make a priori predictions about whether particular changes are likely to happen or not. If we saw competing variants as completely identical in terms of both their linguistic *and* social value, how could directed transitions come about? To address this question, it is useful to consider ideas from the wider domain of cultural evolution.

1.3 Replicator-neutral accounts

The evolutionary approach to language variation and change outlined above has also been adopted widely to study processes of cultural change more generally (Boyd and Richerson, 1985; Mesoudi, 2011). Interestingly, even though replicator-neutral accounts – where individuals have no inherent preference for any of the competing variants – have been studied extensively in the context of cultural evolution (Bentley et al., 2004, 2007), such models have received relatively little attention in the study of linguistic change (e.g. Trudgill, 2008; Baxter et al., 2009).

One of the attempts to build a bridge between general models of cultural evolution and the dynamics of language change is Reali and Griffiths (2010). Starting from a model of pure neutral evolution by random copying – where individuals replicate the different variants proportionally to their current frequency – the authors add what they frame as an inferential bias for *regularisation.* They show that the trajectories produced by this 'regularising' neutral model exhibit s-shaped growth, as long as only those trajectories initiated at 0% use of a novel variant and terminating at 100% use are considered. Crucially, however, their mathematical model captures all possible trajectories between those two points, and their result holds only for the *average of all possible trajectories.* This idealised trajectory is highly unlike the 'typical' transitions produced by their neutral evolution model, which are characterised by a noisy trajectory, often with many reversals. The strict symmetry of their Markov model also predicts that there should be as many completed language changes as there are actuated changes that went to the 50% mark before being interrupted, a situation that does not seem to be the case. These considerations call into question whether neutral evolution by random copying can provide an adequate model of the dynamics of language change (Blythe, 2012).

While in pure neutral evolution models the likelihood of replicating a variant is assumed to be dependent on that variant's current frequency alone, another class of replicator-neutral models that has received increased attention recently considers the effects of *temporal information* and *memory* on the diffusion of cultural traits. Labov (2001) for example suggested that the systematic incrementation of sound changes across generations could be explained by the notion of *age vectors*. He hypothesises that, following an initial stage where learners acquire the average community usage of linguistic variants, adolescents advance their productions in line with the age stratification of variable usage that can be observed in the population – in other words, it presumes that youngsters have a bias against sounding *outdated*. This idea was taken up by Mitchener (2011), who framed it in terms of *prediction-driven instability*: in his mathematical model, individuals are able to observe the usage levels of a categorical sociolinguistic variable among the 'older' and 'younger' individuals in the population. New individuals entering the population then adopt a usage rate according to the predicted future use of the variants, by extrapolating from the usage levels of the two groups along an idealised logistic curve. While the model exhibits spontaneous transitions between the two (or more) competing language states, it produces trajectories that exhibit rapid growth from the onset of the change, unlike the gradual uptake observed in empirical data such as shown in Fig. 1. The individuals' usage rates also remain fixed after they are initally acquired, leaving open the question of whether the same mechanism could also give rise to directed changes when individuals adjust their usage rates throughout their lifetime, as has been observed in linguistic changes (Sankoff and Blondeau, 2007).

Another general model of cultural evolution based on a similar principle is *momentum-based selection* (Gureckis and Goldstone, 2009), which we will study more closely in the remainder of the current analysis. In this model, an individual's choice between competing cultural variants is influenced by the variants' *momentum*, i.e. by *changes to the variants' frequency of use* in the recent past. Individuals are assumed to be biased towards variants which have recently seen an increase in their usage rate, and conversely biased against variants that have been adopted relatively less frequently in the recent past.

Gureckis and Goldstone test their model on a dataset of the frequency of names given to children in the US over 127 years. Their prediction for the popularity of a name in a given year, which is based on the name's long-term popularity modulated by its short-term momentum, leads to a significantly better fit of the empirical data than the prediction made by pure random copying accounts which do not incorporate momentum. Importantly, their model was primarily intended to be fit to empirical data, but not meant as a generative model of individual behaviour. The authors rule this out, noting that "if rising names are preferred, which in turn causes them to rise, then a momentum bias might quickly lead to convergence on a single token" (p.668). They regard this as a negative property of the model, as they are interested in mechanisms that exhibit *cycles* in the popularity of traits, such as found in the realm of fashion (Kroeber, 1919; Berger and Le Mens, 2009; Acerbi et al., 2012). In language, on the other hand, convergence on a single convention is the rule rather than the exception, suggesting that momentum-based selection may be an appropriate model for language change.

2 Momentum-based selection

Our main contribution in this work is to investigate the dynamics of momentum-based selection by integrating it into an existing framework of language change, and evaluating it with respect to the characteristics of language change we identified above: the *sporadic* nature of changes which, once actuated, proceed in an orderly, *directed* manner. We begin by reviewing the original formulation of momentum-based selection in Gureckis and Goldstone (2009). The model is built around tracking exponentially weighted moving averages (EWMAs) of the relative frequencies of competing cultural traits over time. Given a time series of relative frequencies $\vec{n} = \langle n_1, n_2, n_3, \ldots \rangle$, the weight of each data point towards the moving average, which we denote $\hat{n}_\alpha(t)$, decreases exponentially over time (hence the name). Given a new datum n_t received at time t, the moving average is updated iteratively using

$$\hat{n}_\alpha(t) = \alpha \cdot n_t + (1 - \alpha) \cdot \hat{n}_\alpha(t-1) \; . \tag{1}$$

The parameter $\alpha \in [0, 1]$ is a smoothing factor that determines both the weight given to the newest data point, as well as how quickly the data points' weight decreases over time. At time t, the relative weight of datum n_{t-i} to the current average is $\alpha \cdot (1-\alpha)^i$. The higher α, the more weight is given to more recent data points. Based on this, the momentum of a variant at time t, $m(t)$, is determined by calculating two EWMAs $\hat{n}_\alpha(t), \hat{n}_\gamma(t)$ of the variant's attested frequencies $\langle n_1 \cdots n_t \rangle$ with two distinct smoothing factors $\gamma > \alpha$, and taking their difference,

$$m(t) = \hat{n}_\gamma(t) - \hat{n}_\alpha(t). \tag{2}$$

Because the higher γ gives more weight to recent data points, the moving average $\hat{n}_\gamma(t)$ corresponds to the more recent popularity of a trait while $\hat{n}_\alpha(t)$ captures its long-running popularity. The momentum term $m(t)$ will consequently be positive if a variant has been more popular in the recent past compared to its long-term popularity, and negative if the variant has been adopted relatively less frequently in the recent past.

2.1 Mathematical properties of momentum

To understand just what is captured by the momentum term $m(t)$, we can investigate the general dynamics of the difference between two EWMAs $\hat{n}_\alpha(t), \hat{n}_\gamma(t)$ based on their parameters $\gamma > \alpha$. The strongest possible trend in changes to relative variant frequency can be achieved by initialising both EWMAs so that they indicate categorical usage of, say, the outgoing variant (i.e. $\hat{n}_\alpha(0) = \hat{n}_\gamma(0) = 0$), and then continuously updating both EWMAs with input data suggesting that, actually, everyone is using the novel, incoming variant categorically (i.e. $\vec{n} = \langle 1, 1, 1, \ldots \rangle$). Even in this simple case, the dynamics of the momentum term are complex, as can be seen in Fig. 2.

For the underlying EWMAs themselves, the higher the smoothing factor, the faster they approach the input values (Fig. 2a.i), and the more quickly they reflect changes in the distribution too (Fig. 2a.ii). The corresponding momentum terms that are derived by subtracting an EWMA with a high parameter γ from a more slowly changing one with a lower parameter α are shown directly underneath (Fig. 2b). What is of interest to us are the different *shapes* of these momentum curves: a parameter combination which exhibits a rapidly rising curve will cause an individual to posit a trend based on just a few suggestive input data points, while a curve that slopes off slowly means that a momentum bias will persist for a longer time after the initial detection of a trend.

Both the number of data points it takes to reach their maximum value as well as the amplitude of this highest possible momentum value depend on both smoothing factors in complex ways. The short-term memory parameter γ is of particular importance, as it controls the time depth at which the momentum term is most sensitive to underlying trends in the data: a high γ causes the momentum term to immediately reflect short-term variation in the input, while settings of γ closer to α lead to more conservative trend estimates which smooth over the noise present in individual input data points.

The sudden change in trend after 60 data points shown in Fig. 2b.ii illustrates this point: a momentum term based on high $\gamma = 0.15$ (dotted line), while very quick to reflect sudden changes in the input, is very unstable. After receiving only five data points of the new input value $n_t = 0$, the previous sustained upward trend is 'forgotten', with the momentum term first quickly returning to 0, then going negative to reflect the new, short-term downwards trend from the series of 1s back to 0s.

[Figure 2 about here.]

Generally, assuming an abrupt change in the input values such as above, the number of iterations that both EWMAs have to be updated with the same constant input value before the maximum possible difference between the two is reached is

$$t_{\text{mmax}}(\alpha, \gamma) = \frac{\ln \frac{\alpha}{\gamma}}{\alpha - \gamma} . \tag{3}$$

The maximum possible amplitude of the momentum term at that point is

$$m_{\max}(\alpha, \gamma) = e^{-\gamma t_{\text{mmax}}(\alpha,\gamma)} - e^{-\alpha t_{\text{mmax}}(\alpha,\gamma)} . \tag{4}$$

Knowing the mathematical boundaries of the momentum term $m(t)$ we can now go on to incorporate the momentum bias into a model of language change.

2.2 The Utterance Selection Model of language change

To investigate the dynamics of momentum-based selection as a model of individual behaviour, we implemented the momentum-based selection bias in the *utterance selection model* of language change (USM) (Baxter et al., 2006; Blythe and Croft, 2012). Derived from Croft's evolutionary theory of language change (2000), the USM provides a well-studied multi-agent framework to study the dynamics of the competition and diffusion of *discrete* linguistic replicators, be they lexical items, constructions, or different categorical variants of a speech sound[2].

[2]For an account of how age vectors can drive change in a continuous dimension such as vowel productions, see Swarup and McCarthy (2012).

Two fundamental principles underlie the design of the USM: firstly, the individual agents use the competing variants *proportionally*, rather than categorically. In the minimal case with only two competing variants studied here, an agent's usage rates can be fully described by a single number, call it x, in the range $[0, 1]$. While this value can be interpreted as reflecting some cognitive state of the speaker, it also has a more direct behavioural correspondent: when an agent is selected to participate in an interaction, their probability of producing the novel variant is equal to x, while the probability of producing the competing variant is $1 - x$. This aspect of the USM is in line with linguistic evidence which shows that human language use is inherently variable and probabilistic (Kroch, 1994; Labov, 1994; Bybee, 2007; Nardy et al., 2013).

Secondly, agents continuously tune their own variable usage rate towards the production rates they observe in interactions with other agents, thus mimicking the human tendency to *align* linguistic behaviour with that of interlocutors (Giles et al., 1991; Branigan et al., 2000; Jaeger and Snider, 2013). This aspect of the USM is in line with the finding that many aspects of linguistic behaviour do not remain fixed, instead remaining malleable across an individual's life span (Kerswill, 1996; Sankoff and Blondeau, 2007; Beckner et al., 2009; Bowie and Yaeger-Dror, 2013; Stanford, 2014). According to the formal definition of the USM (Baxter et al., 2006), an agent's current proportion of use of a variant $x_\alpha(t)$, is simply an exponentially weighted moving average (EWMA) of the frequencies of the incoming variant that the agent has observed in their input over time[3]. The rate of alignment is controlled by the smoothing factor α of this EWMA, which can be understood as a *learning rate*. This learning rate is typically held small (in the range of 0.01): there is alignment, but the individual frequency adjustments after an interaction are very small and it takes many interactions for an agent to change their preferred variant.

On top of this basic update rule, a USM agent's alignment behaviour can be altered by applying biases to their input data before it gets incorporated into the EWMA. This is where momentum-based selection comes into play.

[3]For simplicity of notation we will henceforth omit the ˆ above the variables denoting EWMAs.

2.3 Momentum-based selection in the USM

We now explain how to minimally incorporate momentum-based selection into the USM. Assuming an agent using learning rate α has just engaged in its t-th interaction and observed another agent use the incoming variant with a relative frequency of y, then their own frequency of use x_α is updated to be

$$x_\alpha(t) = \alpha \cdot f(y) + (1 - \alpha) \cdot x_\alpha(t-1) , \tag{5}$$

where $f(y)$ is a function from $[0, 1]$ to $[0, 1]$ which transforms the *objective* observed frequency y of the variant into a subjective *perceived frequency* which the agent then aligns to. Similar to Gureckis and Goldstone (2009) we can now simply define the perceived frequency $f(y)$ of an agent in the momentum-based USM as the objective frequency y of a variant observed in an interaction offset by that variant's momentum,

$$f(y) = y + b \cdot m'(t) \tag{6}$$

with the exception of

$$f(0) = 0 \quad \text{and} \quad f(1) = 1 . \tag{7}$$

We impose the latter since our focus lies on modelling the diffusion of existing linguistic variants, independent of how those variants were introduced into the population to begin with. It simply stops our momentum-biased selection function $f(y)$ from introducing novel, unattested variants, a constraint that is typical of models of selection generally (see e.g. Boyd and Richerson, 1985). The positive bias parameter b in equation 6 controls the strength with which the normalised momentum term $m'(t)$ as defined below in Equation 8 influences the perceived frequency. Should the momentum bias cause $f(y)$ to go below 0 or above 1, it is simply truncated at 0 and 1, respectively[4]. Crucially, because the momentum term can be positive or negative (depending on the direction of the trend), this perceived frequency function is *symmetric*, which makes it *replicator-neutral*: no matter which bias strength b is used, the function does not a priori favour one of the variants over the other.

Since the effect of different strengths of this bias parameter b on the model dynamics is relevant to our analysis, we have to make sure

[4]The exact form of the bias function $f(x)$ matters much less than its monotonicity and the fact that $f(x) > x$ when the momentum term is positive (i.e. when the agent perceives an upward trend) and $f(x) < x$ when it is negative (indicating a downward trend).

that its settings are comparable across settings of the other parameters. This isn't as straightforward as it might seem, because the range of values that the original momentum term definition $m(t)$ in Equation 2 can take on depends on both smoothing factors α and γ, as could be seen in Fig. 2. The absolute amplitude of the momentum curves is of little interest to us; on the contrary, the differences in maximum possible amplitude distort the effect of the bias parameter b which is supposed to control the strength with which momentum is applied. To counteract this, we normalise the momentum term $m(t)$ based on the α, γ used in a given simulation condition. For any given pair of smoothing factors α, γ, we can scale the momentum term to the $[-1, 1]$ range by defining the normalised momentum

$$m'(t) = \frac{x_\gamma(t) - x_\alpha(t)}{m_{\max}(\alpha, \gamma)} \ . \tag{8}$$

To calculate the momentum component in the numerator, the difference between two EWMAs, we simply re-use the agent's own usage frequency, which according to the USM definition is also an EWMA. To augment the basic USM with momentum-based selection, every agent simply has to keep track of another x_γ on top of the long-term estimate x_α it already maintains.

3 Results

3.1 Analytical approximation

Before proceeding to a full population-based simulation we can establish the general dynamics of the model by investigating the behaviour of an individual agent set in an idealised, deterministic production-perception loop (Wedel, 2006). We initialise a single agent to use the incoming variant at some low level and repeatedly update their two EWMAs $x_\alpha(t), x_\gamma(t)$ by having them align to their own proportion of use $x_\alpha(t)$ for 100 iterations. As can be seen in Fig. 3, nothing happens: an agent aligning to their own usage rate simply remains at that proportion and, in the absence of any changes in the input sequence, the momentum term stays 0. To test how the model reacts to fluctuations in the input we alter the agent's input by fabricating a data point which suggests that their interlocutors are actually categorically using the incoming variant (see Fig. 3a). When the agent aligns to this input it leads to a small punctual increase in their variant use, but the sudden change in the input data also makes the momentum term take on a positive value (dashed grey line). Following the fabricated data point, the agent again receives their own samples as input data. But the bias exerted by the momentum term, which makes the agent's *perceived* usage rate higher than their actual usage rate, causes further increases in their use of the incoming variant. However, the lack of further perturbations causes the momentum to decay back towards 0, and the agent becomes stationary again at a usage level not far from their initial setting. If we introduce a second fabricated data point shortly after the first one, the model's behaviour changes dramatically: the system enters a regime where the momentum bias generated by the two fabricated data points affects the perceived frequency of the agent's input so much that it causes the momentum term to increase even further, leading to self-reinforcing runaway change (Fig. 3b).

[Figure 3 about here.]

This preliminary analysis shows that the momentum-based selection model exhibits two different regimes, accounting for both periods of stability and of directed change. Capturing the dynamics of the transition between the two regimes is however not trivial: particularly the switch from stability to a directed transition depends crucially on both the strength of the momentum bias as well as random fluctuations in the agents' input as they sample data from their interlocutors. We therefore turn to numerical simulations, where the data production and agent interactions will be driven by stochastic processes.

3.2 Numerical simulation

In order to get a fuller picture of the momentum-based selection dynamics we ran simulations with a total of 2,520 parameter combinations[5]. The six parameters of the momentum-based USM are summarised below. Only one, the learning rate α, was held constant across all simulation runs, the other five parameters were varied at the levels given in parentheses:

- α: the agents' learning rate (0.01)
- γ: the agents' short-term memory smoothing factor $(0.015, 0.02, 0.025, 0.03, 0.35, 0.4)$
- T: the Binomial sample size determining the resolution at which agents can observe each other's relative usage frequencies $(2, 3, 4, 5)$

[5]The source code for running the simulations as well as the analytical approximation are available at `http://github.com/kevinstadler/momentum`

- b: the bias strength with which agents apply the normalised momentum to yield their *perceived* frequency of usage $(0.5, 1.0, 1.5, 2.0, 2.5)$

- N: number of agents in the population $(2, 5, 10, 20, 30, 50, 100)$

- x_0: initial proportion of the incoming variant used by all agents $(0.01, 0.02, 0.03)$

Combining all these possible parameter combinations and running the 2,520 conditions for 48 trials each resulted in a total of 120,960 simulation runs. On top of the conditions listed above, we also produced simulation runs where we set the bias strength $b = 0$, which is equivalent to pure neutral evolution. 24,192 runs from this additional condition provide a baseline that the dynamics of our momentum-based selection model can be compared against. Every simulation run proceeds as follows:

Firstly, initialise N agents, setting both their x_α and x_γ to x_0. Then, carry out interactions between agents by repeating the following steps:

1. randomly select two agents i, j from the pool of N agents – we assume that all pairs of agents have the same probability of interacting with each other.

2. let both agents produce T tokens of the variable by taking a random sample n_i, n_j for each agent from the Binomial distribution $B(T, x_\alpha)$, using the two agents' respective usage rates x_α at the time of the interaction.

3. calculate the perceived frequencies that the agents will align to, using equation 6. For agent i, who will align to j's productions, calculate $f(\frac{n_j}{T})$ using agent i's current normalised momentum term $m'(t)$; for agent j, calculate $f(\frac{n_i}{T})$ using j's $m'(t)$.

4. update both agents' x_α as well as x_γ by incorporating their perceived frequency according to equation 5.

The simulations were run until every individual in the population had converged to within a ten-thousandth of a percent of using only one of the two competing variants, or for a maximum of 200,000 interactions per agent[6].

[6]More than 99% of simulation runs had terminated before this time limit was reached.

3.3 Simulation results

For the sake of our analysis we use a simple definition of what a 'transition' is. Taking a fixed threshold (say 5%), we can define the two extreme areas where the mean population usage level of the minority variant is below this threshold as the two regions of 'near-categorical use' of either variant. A transition, then, is the period in which the mean usage levels of the population crosses from near-categorical use of one to near-categorical use of the other variant. A first striking finding when analysing the simulation results is that changes are rare: of the 120,960 simulation runs using the momentum bias, only 18,040 (around 15%) ever exhibit a transition, while the majority of runs simply converge on categorical use of the majority variant. This result is in line with the observation that the actuation of language change is *sporadic*: even when a novel variant is known to the entire population, this alone is not likely to lead to a community-wide language change.

[Figure 4 about here.]

When we investigate the distribution of transitions across the different parameter settings, we find that the bias strength b carves the space into two regions with distinct dynamics: while simulation runs with $b \geq 1$ exhibit directed transitions at comparable time scales, the neutral evolution condition with $b = 0$ as well as the weak momentum bias setting at $b = 0.5$ yield both fewer and temporally less consistent transitions, as shown in Fig. 4. The difference between those two regimes is exacerbated as population sizes become larger, making transitions in the neutral evolution conditions even rarer and slower.

Beyond this qualitative difference in successful transitions, our earlier prediction regarding the general directedness of trajectories in the neutral evolution condition are also borne out by the results: of all simulation runs where the incoming variant ever reaches the half-way mark (i.e. average 50% usage of both variants across the population), only 55% of trajectories in conditions with $b \leq 0.5$ actually result in the diffusion of the incoming variant. The remaining half-completed transitions are interrupted and revert back to majority usage of the established variant. In contrast, in conditions with $b \geq 1.0$, 97% of the trajectories that reach the half-way mark also lead to the population-wide adoption of the incoming variant.

In contrast to the low-bias conditions which exhibit the dynamics of neutral evolution, conditions with a sufficiently high momentum bias b reliably produce s-shaped transitions between the two regions of near-categorical use at irregular intervals. before eventually converging on categorical use of either of the variants. The dynamics are robust under many different parameter settings which give rise to highly similar transition dynamics (see Fig. 4; the parameters' much greater influence on the likelihood of transitions occurring is beyond the scope of this paper). While similar transitions are also found in models driven by replicator selection, an important difference is that our model has no a priori preference for any of the variants built in. Instead of having a constant bias applied from outwith the model, the momentum term provides the opportunity for a bias to emerge dynamically from within the system, as can be seen from the temporal development of the momentum term in Figs. 5. Crucially, rather than relying on an external trigger, the s-shaped transitions are *self-actuating*: agents constantly read weak trends into the random fluctuations in their input but these temporary individual biases will vary across the population, and more often than not cancel each other out. There is, however, always the possibility for these weak biases to overlap, which could cause a subset of agents to slowly shift their variant use in parallel. When this shift is detected by other agents they will themselves start to amplify it, leading to a self-reinforcing feedback loop. The directed transitions in a momentum-based model of language change are triggered *spontaneously* and, while it is the most likely outcome, changes are not guaranteed to succeed either: even if a change is actuated, its propagation is not completely inevitable, as can be seen in interrupted changes such as the one shown in Fig. 5b. The dynamics of momentum-based selection provide an intriguing account of the unpredictability of the actuation of linguistic changes without the need for an external bias or trigger.

[Figure 5 about here.]

The trajectories shown in Figs. 5 are exemplary of the dynamics of momentum-based selection across the full range of parameter settings we explored. Only for settings of the momentum bias b close to 0 as well as for short-term smoothing factors γ very close to the learning rate α do the momentum-based selection dynamics break down, and the model reverts to pure neutral evolution-like behaviour. In comparison to the prediction-driven model of Mitchener (2011), the momentum-based selection model shows that it is not necessary for learners to engage in active prediction of the population's *future* state along a particular trajectory. Rather, having a simple bias based on variant history is sufficient to drive orderly directed changes, and the transitions generated by our model appear to exhibit a more gradual uptake than the trajectories reported by Mitchener. We also find that having a bias for *regularisation* is not absolutely necessary to guarantee an orderly progression of the changes. In a population of agents who are continuously updating their usage rates, the momentum bias presented here is robust enough to drive changes to near-completion.

4 Discussion

We have shown that the momentum-based selection model fulfills two defining requirements of a model of language change: the spontaneous, sporadic actuation of changes, and their progression in the form of a directed, s-shaped curve. However, other accounts of language change which posit a selection bias in favour of the incoming variant also predict s-shaped trajectories, so how can we know which account best describes the empirical data? While the progression of every instance of language change will be influenced by several factors concurrently or at different times (see e.g. Ghanbarnejad et al., 2014; Stanford, 2014), it is still interesting to investigate which (if any) of the mechanisms of language change discussed in the introduction can be identified as the main driving force behind language change. Here, we want to highlight some of the more subtle differences in the predictions made by different accounts of language change which would allow us to tease apart the momentum-based, language-internal and social accounts of language change based on cross-linguistic data.

4.1 The two rates of linguistic change

An interesting (and to our knowledge novel) way to evaluate competing theories of language change is to look at the predictions they make regarding the *rates* of linguistic change. It is important to note that 'rate' can refer to two different things in the context of language change: one interpretation of 'rate' refers to the probability of a particular change occurring, such as when talking about different English past tense forms becoming regularised over time (Lieberman et al., 2007) or the rate of lexical replacement more generally (Monaghan, 2014). Rather than referring to the time frame within which a specific change takes place, this really describes the *likelihood of a (type of) change*, or an *actuation probability.* The other use of 'rate' refers to the *speed* of the transition of one particular change, i.e. it is a measure of the time span from the introduction of a new variant to its completely replacing an established one. Under the assumption that language change follows an s-shaped pattern, this second rate of change is often taken to be the growth rate parameter of the logistic function (Pintzuk, 2003), and it is this 'rate' that is referred to by the 'Constant Rate Effect' observed in syntactic change (Kroch, 1989).

What is interesting about these two rates of change is that different accounts of language change make implicit predictions regarding the relationship between them, in particular whether the likelihood of a change occurring is correlated with the rate at which the change proceeds once it has been actuated. Under the assumption that the same pressures that lead to the introduction of more functional or 'adaptive' variants are also responsible for their preferred selection once they have been innovated, language-internal accounts would predict that changes which occur more often cross-linguistically should also be selected for more strongly in individual languages. This would translate into faster changes so that, controlling for other factors such as frequency and size of the speech community, the two rates of change should be positively correlated. This differs from the prediction made by the momentum-based account: while the probability of a new variant appearing, and consequently its random actuation from the pool of variants, is dependent on linguistic factors, these factors are not what drives the diffusion of the variant. Assuming that individuals apply similar momentum biases to all linguistic variables, a momentum-based account would therefore predict the speed of individual transitions and the changes' actuation probability to be uncorrelated.

The situation with social accounts is trickier: the fact that many different social factors have been posited to influence the selection of linguistic variants, both positively and negatively, makes it difficult to derive a general prediction regarding the speed of individual changes. What determines the probability of actuation is an equally open question: it has been proposed that the occurrence of changes might be driven by the need to create distinct social identities within a community (Labov, 2002; Matthews et al., 2012; Roberts, 2013), implying that we should not expect actuation probabilities to be constant cross-linguistically.

While it is difficult to derive specific predictions regarding the correlation between the two rates of change from social accounts of language change, many insights into the respective roles of the different pressures could be gleaned from studying cross-linguistic datasets of changes (see also Bickel, 2015). The crucial issue is that the three qualitatively very different accounts discussed here might predict quantitatively similar selection pressures for particular language changes, making it impossible to distinguish the contribution of the different types of pressures on a *per-change* basis. Our understanding of the issue could therefore profit immensely from investigating the empirical distribution of *both* rates of change as well as their relationship based on cross-linguistic data.

4.2 Momentum-sensitivity in the individual

While momentum-based selection successfully reproduces the macro-level s-shaped curves that are characteristic of linguistic change, this raises the question of whether the model makes valid assumptions about individuals' micro-level behaviour (Mesoudi and Lycett, 2009). Firstly, it is clear that both linguistic knowledge and performance are embedded in diachrony – language users are sensitive to changes in the frequencies of variants (Jaeger and Snider, 2013) and well aware of diachronic connotations (Labov, 2001; Guy, 2003; Tagliamonte, 2012), both types of information that could drive momentum-based selection. In the general cultural evolution literature it is well-established that frequency-dependent biases are a natural strategy for social learning tasks, since frequency can be an indicator of the *social value* of a variant (Boyd and Richerson, 1985). Similarly, *changes* in frequency can be a good indicator of the *future* social value of a cultural variant (Gureckis and Goldstone, 2009). Laboratory experiments on cultural evolution in humans have provided empirical evidence for the self-perpetuating nature of trends, where people will amplify trends even against their own personal preferences (Salganik and Watts, 2008; Willer et al., 2009), suggesting that individuals might also have an incentive to use metalinguistic information about the history of linguistic variants. While there is plenty of qualitative and anecdotal evidence on speakers' explicit evaluation of language changes (see e.g. Trudgill 1972; Labov 2001; Guy 2003; Tagliamonte 2012), quantitative research on the extent of people's explicit or implicit knowledge about the direction of ongoing changes is just starting. Experimental evidence shows that listeners employ their implicit knowledge about ongoing sound changes during speech perception (Hay et al., 2006; Drager, 2011), and there is evidence of explicit knowledge both in the area of phonetic (Carrera-Sabaté, 2014) and syntactic change (Stadler et al., 2016).

5 Conclusion

In this paper we investigated the *momentum-based selection* model and studied its evolutionary dynamics. Our analysis shows that this model, where individuals are biased towards variants which have recently seen an increase in their frequency of use, exhibits two features characteristic of language change: the spontaneous, sporadic actuation of changes, and their progression in the form of directed, s-shaped curves.

Crucially, the momentum-based selection mechanism demonstrates that the apparent selection of a particular cultural variant in a population is not sufficient evidence for any inherent asymmetry between the variants in competition. Instead, selection biases can be an emergent property of the system, particularly in the case of social learning where individuals possess meta-level knowledge about the variants. The human capacity to acquire meta-linguistic knowledge about ongoing language changes, for example by tracking changes in the variants' frequencies of use over time, therefore deserves further study.

Finally, we highlighted the importance of collecting and studying cross-linguistic data sets of comparable historical changes to test the general predictions made by different accounts of language change. This strand of research in particular needs to be expanded further in order to help us gain deeper insights into the respective roles of the myriad pressures involved in language change.

Acknowledgments

Kevin Stadler was funded by a PhD scholarship from the College of Humanities & Social Sciences at The University of Edinburgh. We would like to thank the anonymous reviewers as well as Graeme Trousdale, Joachim De Beule, Peeter Tinits, W. Garrett Mitchener and all members of the Centre for Language Evolution for useful comments on earlier drafts of this paper.

References

Acerbi, Alberto, Stefano Ghirlanda, and Magnus Enquist. 2012. The logic of fashion cycles. *PloS one* 7: e32541. doi:10.1371/journal.pone.0032541.

Altmann, Gabriel H., Haro von Buttlar, Walter Rott, and Udo Strauß.

1983. A law of change in language. In Barron Brainerd (ed.), *Historical linguistics*, 104–115. Bochum: Studienverlag Dr. N. Brockmeyer.

Bailey, Charles-James N. 1973. *Variation and linguistic theory.* Arlington, VA: Center for Applied Linguistics.

Baxter, Gareth J., Richard A. Blythe, William Croft, and Alan J. McKane. 2006. Utterance selection model of language change. *Physical Review E* 73: 046118. doi:10.1103/PhysRevE.73.046118.

Baxter, Gareth J., Richard A. Blythe, William Croft, and Alan J. McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change* 21: 257. doi:10.1017/S095439450999010X.

Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. 2009. Language is a complex adaptive system: position paper. *Language Learning* 59: 1–26. doi:10.1111/j.1467-9922.2009.00533.x.

Bentley, R. Alexander, Matthew W. Hahn, and Stephen J. Shennan. 2004. Random drift and culture change. *Proceedings of the Royal Society B: Biological Sciences* 271: 1443–1450. doi:10.1098/rspb.2004.2746.

Bentley, R. Alexander, Carl P. Lipo, Harold A. Herzog, and Matthew W. Hahn. 2007. Regular rates of popular culture change reflect random copying. *Evolution and Human Behavior* 28: 151–158. doi:10.1016/j.evolhumbehav.2006.10.002.

Berger, Jonah and Gaël Le Mens. 2009. How adoption speed affects the abandonment of cultural tastes. *Proceedings of the National Academy of Sciences of the United States of America* 106: 8146–8150.

Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine and Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 901–923. Oxford: Oxford University Press. doi:10.5167/uzh-109110.

Blythe, Richard A. 2012. Neutral evolution: a null model for language dynamics. *Advances in Complex Systems* 15: 1150015. doi:10.1142/S0219525911003414.

Blythe, Richard A. and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88: 269–304. doi:10.1353/lan.2012.0027.

Bowie, David and Malcah Yaeger-Dror. 2013. Phonological change in real time. In Joseph Salmons and Patrick Honeybone (eds.), *The Oxford handbook of historical phonology*, chap. 34. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199232819.013.004.

Boyd, Robert and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press. ISBN 9780226069333.

Branigan, Holly P., Martin J. Pickering, and Alexandra A. Cleland. 2000. Syntactic co-ordination in dialogue. *Cognition* 75: B13–B25. doi:10.1016/S0010-0277(99)00081-5.

Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford University Press. ISBN 9780195301564.

Carrera-Sabaté, Josefina. 2014. Does meta-linguistic awareness play any role at the beginning of an ongoing sound change? The case of some vowel-ended verbs in Catalan. *Sociolinguistic Studies* 8: 193–221. doi:10.1558/sols.v8i2.193.

Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow: Pearson Education Limited.

Croft, William A. 2006. The relevance of an evolutionary model to historical linguistics. In Ole Nedergård Thomsen (ed.), *Competing models of linguistic change*, Current Issues in Linguistic Theory 279, 91–132. John Benjamins Publishing Company. doi:10.1075/cilt.279.08cro.

de Saussure, Ferdinand. 1959. *Course in general linguistics*. New York: The Philosophical Library, Inc.

Denison, David. 2003. Log(ist)ic and simplistic S-curves. In Raymond Hickey (ed.), *Motives for language change*, 54–70. Cambridge: Cambridge University Press. ISBN 978-0-521-79303-2.

Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech* 54: 99–121. doi:10.1177/0023830910388017.

Ellegård, Alvar. 1953. *The auxiliary do: the establishment and regulation of its use in English.* Gothenburg studies in English. Stockholm: Almqvist & Wiksell.

Evans, Nicholas and Stephen C. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *The Behavioral and brain sciences* 32: 429–48; discussion 448–494. doi:10.1017/S0140525X0999094X.

Foulkes, Paul and Gerard Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34: 409–438.

Foulkes, Paul and Marilyn Vihman. 2013. First language acquisition and phonological change. In Patrick Honeybone and Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, chap. 18. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199232819.013.001.

Ghanbarnejad, Fakhteh, Martin Gerlach, José M. Miotto, and Eduardo G. Altmann. 2014. Extracting information from S-curves of language change. *Journal of the Royal Society, Interface* 11: 20141044. doi:10.1098/rsif.2014.1044.

Giles, Howard, Justine Coupland, and Nikolas Coupland (eds.). 1991. *Contexts of accommodation.* Studies in Emotion & Social Interaction. Cambridge: Cambridge University Press. ISBN 9780521361514.

Greenberg, Joseph H. 1959. Language and evolution. In Betty J. Meggers (ed.), *Evolution and anthropology: a centennial appraisal.* Washington DC: The Anthropological Society of Washington.

Griffiths, Thomas L. and Michael L. Kalish. 2007. Language evolution by Iterated Learning with Bayesian agents. *Cognitive Science* 31: 441–480. doi:10.1080/15326900701326576.

Gureckis, Todd M. and Robert L. Goldstone. 2009. How you named your child: understanding the relationship between individual decision making and collective outcomes. *Topics in Cognitive Science* 1: 651–674. doi:10.1111/j.1756-8765.2009.01046.x.

Guy, Gregory R. 2003. Variationist approaches to phonological change. In Brian D. Joseph and Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, chap. 8, 369–400. Oxford, UK: Blackwell Publishing Ltd. ISBN 9780470756393. doi:10.1002/9780470756393.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34: 458–484. doi:10.1016/j.wocn.2005.10.001.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: The Macmillan Company.

Hockett, Charles F. 1965. Sound change. *Language* 41: 185–204.

Jaeger, T. Florian and Neal E. Snider. 2013. Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* 127: 57–83. doi:10.1016/j.cognition.2012.10.013.

Jaeger, T. Florian and Harry Tily. 2010. On language utility': processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science* 2: 323–335. doi:10.1002/wcs.126.

Jespersen, Otto. 1922. *Language: Its Nature, Development and Origin*. London: George Allen & Unwin Ltd.

Kerswill, Paul. 1996. Children, adolescents, and language change. *Language Variation and Change* 8: 177–202. doi:10.1017/S0954394500001137.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In Emmon Bach and Robert T. Harms (eds.), *Universals in linguistic theory*, 170–202. New York: Holt, Rinehart, and Winston. ISBN 0039100774.

Kirby, Simon. 1999. *Function, selection, and innateness*. Oxford: Oxford University Press. ISBN 978-0-19-823811-9.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199–244. doi:10.1017/S0954394500000168.

Kroch, Anthony S. 1994. Morphosyntactic variation. In Katherine Beals (ed.), *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*.

Kroch, Anthony S. 2001. Syntactic change. In Mark Baltin and Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, Blackwell handbooks in linguistics, chap. 22, 698–729. Malden, Massachusetts: Blackwell Publishers. ISBN 978-0-631-20507-4.

Kroeber, Alfred Louis. 1919. On the principle of order in civilization as exemplified by changes of fashion. *American Anthropologist* 21: 235–263. doi:10.1525/aa.1919.21.3.02a00010.

Labov, William. 1994. *Principles of linguistic change. Internal factors*, *Language in Society*, vol. 1. Oxford: Blackwell.

Labov, William. 2001. *Principles of linguistic change. Social factors*, *Language in Society 29*, vol. 2. Malden, Massachusetts: Blackwell Publishers Inc.

Labov, William. 2002. Driving forces in linguistic change. In *2002 International Conference on Korean Linguistics*. Seoul: Seoul National University.

Lass, Roger. 1980. *On explaining language change*. Cambridge: Cambridge University Press. ISBN 0521228360.

Lewis, Hannah M. and Kevin N. Laland. 2012. Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367: 2171–2180. doi:10.1098/rstb.2012.0119.

Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A. Nowak. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449: 713–716. doi:10.1038/nature06137.

Maegaard, Marie, Torben Juel Jensen, Tore Kristiansen, and Jens Normann Jørgensen. 2013. Diffusion of language change: Accommodation to a moving target. *Journal of Sociolinguistics* 17: 3–36. doi:10.1111/josl.12002.

Matthews, Cristina, Gareth Roberts, and Christine A. Caldwell. 2012. Opportunity to assimilate and pressure to discriminate can generate cultural divergence in the laboratory. *Evolution and Human Behavior* 33: 759–770. doi:10.1016/j.evolhumbehav.2012.06.004.

McMahon, April M. S. 1994. *Understanding language change*. Cambridge: Cambridge University Press. ISBN 9781139166591. doi:10.1017/CBO9781139166591.

Meillet, Antoine. 1921. *Linguistique historique et linguistique générale*. Paris: La Société de Linguistique de Paris.

Mesoudi, Alex. 2011. *Cultural Evolution*. Chicago: University of Chicago Press. ISBN 9780226520445.

Mesoudi, Alex and Stephen J. Lycett. 2009. Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior* 30: 41–48. doi:10.1016/j.evolhumbehav.2008.07.005.

Mitchener, W. Garrett. 2011. A mathematical model of prediction-driven instability: how social structure can drive language change. *Journal of Logic, Language and Information* 20: 385–396. doi:10.1007/s10849-011-9136-y.

Monaghan, Padraic. 2014. Age of acquisition predicts rate of lexical evolution. *Cognition* 133: 530–534. doi:10.1016/j.cognition.2014.08.007.

Nardy, Aurélie, Jean-Pierre Chevrot, and Stéphanie Barbu. 2013. The acquisition of sociolinguistic variation: Looking back and thinking ahead. *Linguistics* 51: 255–284. doi:10.1515/ling-2013-0011.

Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In Leiv Egil Breivik and Ernst Håkon Jahr (eds.), *Language change: contributions to the study of its causes*, 173–198. Berlin: Mouton de Gruyter.

Ohala, John J. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: problems and perspectives*, 237–278. London: Longman.

Pintzuk, Susan. 2003. Variationist approaches to syntactic change. In Brian D Joseph and Richard D Janda (eds.), *The Handbook of Historical Linguistics*, chap. 15, 509–528. Oxford: Blackwell Publishing Ltd. doi:10.1002/9780470756393.ch15.

Postal, Paul Martin. 1968. *Aspects of phonological theory.* New York: Harper & Row.

Reali, Florencia and Thomas L. Griffiths. 2010. Words as alleles: connecting language evolution with Bayesian learners to models of genetic drift. *Proceedings of the Royal Society B: Biological Sciences* 277: 429–436. doi:10.1098/rspb.2009.1513.

Roberts, Gareth. 2013. Perspectives on language as a source of social markers. *Language and Linguistics Compass* 7: 619–632. doi:10.1111/lnc3.12052.

Salganik, Matthew J. and Duncan J. Watts. 2008. Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly* 71: 338–355. doi:10.1177/019027250807100404.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. In Frederick J Newmeyer (ed.), *Linguistics: The Cambridge Survey*, chap. 8, 140–161. Cambridge: Cambridge University Press. ISBN 9780511620577. doi:10.1017/CBO9780511620577.

Sankoff, Gillian and Hélèn Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83: 560–588. doi:10.1353/lan.2007.0106.

Stadler, Kevin, Elyse Jamieson, Kenny Smith, and Simon Kirby. 2016. Metalinguistic awareness of trends as a driving force in linguistic evolution: an empirical study. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*. New Orleans, LA: World Scientific.

Stanford, James N. 2014. Language acquisition and language change. In Claire Bowern and Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*, chap. 21, 466–483. London: Routledge.

Steels, Luc. 2000. Language as a Complex Adaptive System. In Marc Schoenauer, Kalyanmoy Deb, Günter Rudolph, Xin Yao, Evelyne Lutton, Juan Julian Merelo, and Hans-Paul Schwefel (eds.), *Parallel Problem Solving from Nature PPSN VI*, 17–26. Springer. doi:10.1007/3-540-45356-3_2.

Stevens, Mary and Jonathan Harrington. 2013. The individual and the actuation of sound change. *Loquens* 1: e003. doi:10.3989/loquens.2014.003.

Sturtevant, Edgar Howard. 1947. *An introduction to linguistic science.* New Haven: Yale University Press.

Swarup, Samarth and Corrine McCarthy. 2012. Representational momentum may explain aspects of vowel shifts. In *Artificial Life 13*, 267–274. MIT Press. ISBN 9780262310505. doi:10.7551/978-0-262-31050-5-ch036.

Tagliamonte, Sali A. 2012. *Variationist sociolinguistics. Change, observation, interpretation.* Language in Society. Wiley-Blackwell.

Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1: 179–195. doi:10.1017/S0047404500000488.

Trudgill, Peter. 2004. *New-dialect formation: the inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society* 37: 241–254. doi:10.1017/S0047404508080287.

Vennemann, Theo. 1983. Causality in language change: theories of linguistic preferences as a basis for linguistic explanations. *Folia Linguistica Historica* VI: 5–26. doi:10.1515/flih.1983.4.1.5.

Wang, William Shi-Yuan. 1969. Competing changes as a cause of residue. *Language* 45: 9–25. doi:10.2307/411748.

Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: a corpus study. *Cognition* 128: 179–186. doi:10.1016/j.cognition.2013.03.002.

Wedel, Andrew B. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23: 247–274. doi:10.1515/TLR.2006.010.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann and Yakov Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*, 95–188. University of Texas Press.

Willer, Robb, Ko Kuwabara, and Michael W. Macy. 2009. The false enforcement of unpopular norms. *American Journal of Sociology* 115: 451–490. doi:10.1086/599250.

Winter-Froemel, Esme. 2008. Towards a comprehensive view of language change: Three recent evolutionary approaches. In Ulrich Detges and Richard Waltereit (eds.), *The Paradox of Grammatical Change: Perspectives from Romance*, *Current Issues in Linguistic Theory*, vol. 293, 215–250. Amsterdam: John Benjamins Publishing Company.

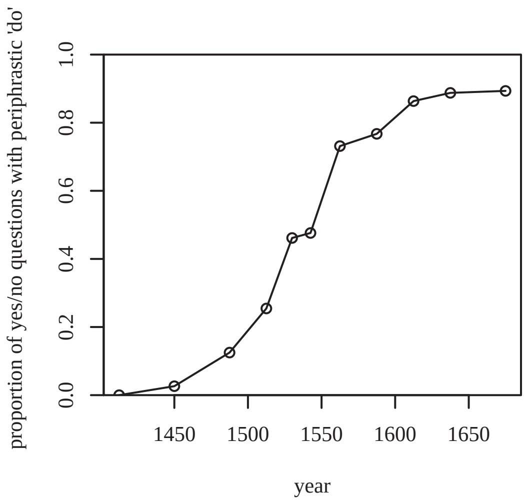

Figure 1: Competition between two syntactic patterns of *yes/no questions*, as observed in a corpus of Middle English writing (Ellegård, 1953). The established question syntax (e.g. “Went he?”) was gradually replaced by its modern variant (e.g. “Did he go?”) along an s-shaped trajectory.

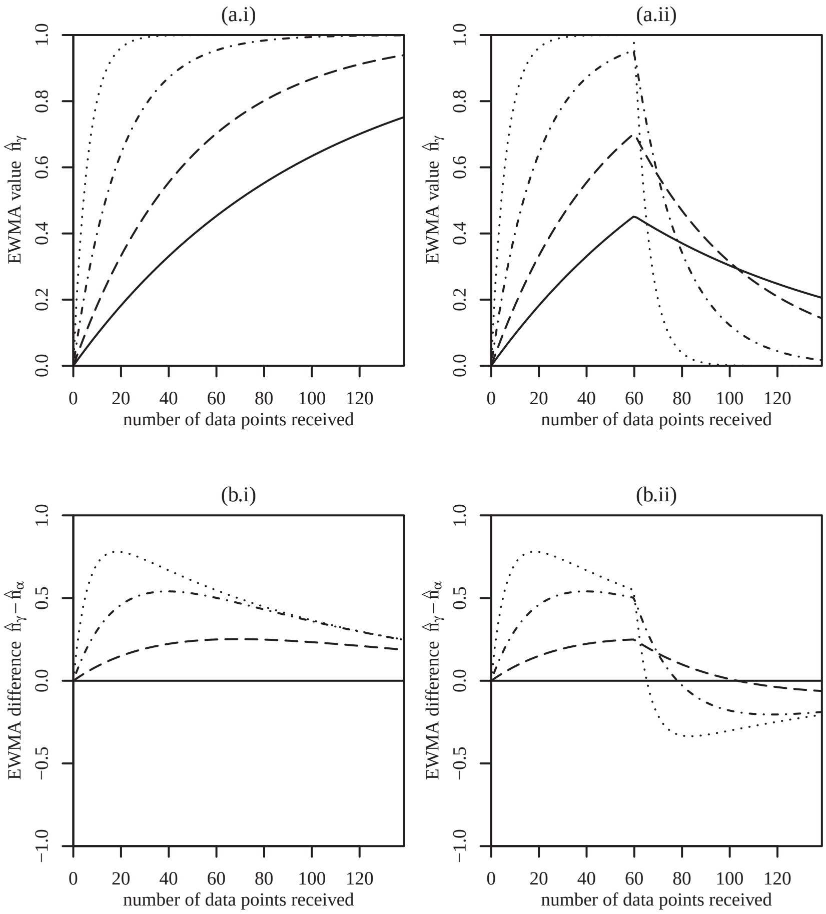

Figure 2: Exponentially weighted moving averages (EWMAs) of the same input data but with different smoothing factors, as well as their corresponding momentum terms. *(a.i)* Four EWMAs with smoothing factors $\gamma = 0.15, 0.05, 0.02, 0.01$ (from top to bottom) are initialised at $\hat{n}_\gamma(0) = 0$ and repeatedly updated using the same constant input data series $\vec{n} = \langle 1, 1, 1 \ldots \rangle$. *(a.ii)* same as (a.i), but with the input data series $\vec{n}$ switching from all 1s to all 0s after 60 data points. *(b)* Corresponding momentum terms $m(t) = \hat{n}_\gamma(t) - \hat{n}_\alpha(t)$ derived from the trajectories above, by taking each EWMA and subtracting the value of the EWMA with the lowest smoothing factor from above ($\alpha = 0.01$). Line styles correspond to those in (a).

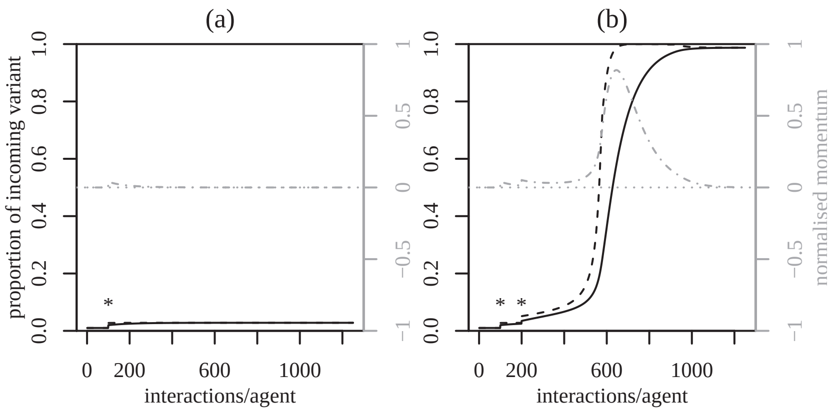

Figure 3: Momentum-based selection dynamics of a single agent's variable usage rate in a deterministic production-perception loop, with learning rates $\alpha = 0.01, \gamma = 0.02$ and momentum bias $b = 2$. At every time step the agent updates their own usage rate (solid black line) by aligning to their own average momentum-biased production with a sample resolution of $T = 5$ (indicated by the dashed black line). This stable loop is perturbed by administering fabricated input data suggesting 100% usage of the incoming variant at the time points marked by asterisks, demonstrating the two regimes of momentum-based selection: *(a)* stability: a single fabricated data point after 100 interactions causes a sudden increase in the agent's usage rate (solid black line) as well as the momentum term (dot-dashed grey line, right axis), but the feedback loop stabilises again. *(b)* directed transitions: adding another fabricated data point after 200 interactions raises the momentum term high enough to trigger self-reinforcing runaway change, giving rise to an s-shaped transition.

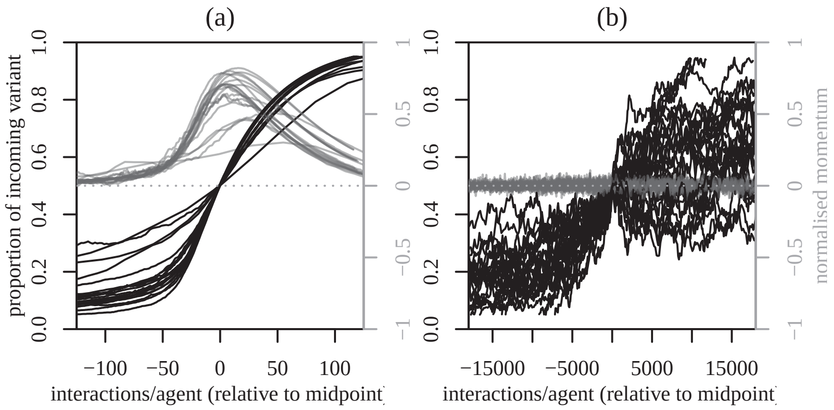

Figure 4: Successful transitions generated by simulation runs in conditions with and without momentum-based selection. The graphs show the development of the average proportion of use of the incoming variant across the population (black line, left axis) from the point where it crosses the 5% mark until it reaches 95%, alongside the average momentum term during that period (grey line, right axis). Transitions are aligned at the point where the trajectory first crosses the 50% mark of incoming variant usage. *(a)* 20 trajectories randomly drawn from the 21,909 successful transitions generated by momentum-based selection with momentum bias $b \geq 1$, population sizes $N \geq 5$ and various settings of γ, T, x_0. *(b)* all 28 transitions generated in 17,280 simulation runs with $b = 0$, equivalent to neutral evolution, with various settings of γ, T, x_0 and population sizes $N \geq 5$. Note the different time scales. The momentum term, ineffective when $b = 0$, is shown for reference.

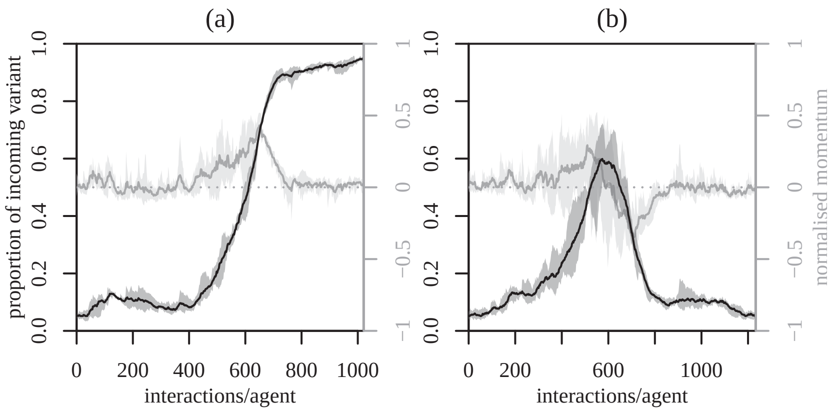

Figure 5: Transitions generated by two simulation runs using identical parameter settings ($N = 5, b = 2.0, T = 2, \alpha = .01, \gamma = .04$). The graphs show the development of the average proportion of use of the incoming variant across the population (black line, left axis) as well as the average momentum term influencing the agents' perception (grey line, right axis). Shaded intervals indicate the range (minimum and maximum values) attested in the population. *(a)* A successful, s-shaped transition typical of momentum-based selection: an initially noisy momentum value rises high enough to trigger self-reinforcement of the momentum bias (at around 450 interactions) until it saturates and tails off again *(b)* Example of a rare, interrupted transition: despite the onset of a directed shift, the wide range of momentum biases across the population destabilises the feedback loop, causing the average momentum to break down and invert, returning the usage frequency of the incoming variant back towards its initial low level.