Modeling language change across the lifespan: individual trajectories in community change

Gareth Baxter, *University of Aveiro*

William Croft, *University of New Mexico*

Corresponding author:

William Croft

MSC03 2103 Linguistics

1 University of New Mexico

Albuquerque NM 87131-0001

505-277-6353

wcroft@unm.edu

Modeling language change across the lifespan: individual trajectories in community change

ABSTRACT

We use a mathematical model to examine three phenomena involving language change across the lifespan: the apparent time construct, the adolescent peak, and two different patterns of individual change. The apparent time construct is attributed to a decline in flexibility towards language change over one's lifetime; this explanation is borne out in our model. The adolescent peak has been explained by social networks: children interact more with caregivers a generation older until later childhood and adolescence. We find that the peak also occurs with many other network structures. The two patterns of language change are one in which most individuals change gradually, following the mean of community change, and one in which most individuals have more categorical behavior, and change rapidly if they change at all. Our model suggests that they represent different balances between the differential weighting of competing variants and degree of accommodation to other speakers.

1. Introduction

From Weinreich, Labov and Herzog onward (1968:188), it has often been observed that the grammar of a speech community is more regular than the grammar of individuals (cf. e.g. Ashby 2001:13). It appears that the same is true of patterns of language change, based on more recent studies of individual speaker behavior in real time. That is, the overall trajectory of change of a linguistic variant in a speech community appears to be more regular than the trajectories of change for individual speakers in the community. Again, from at least Weinreich, Labov and Herzog onward (1968:113), it has been observed that the trajectory of community change is an S-curve: 'the overall changes…display an S-shaped curve despite the variation in the behavior of individual words, speakers, texts, geographical regions, or social classes over the trajectory of the change' (Blythe and Croft 2012:281; see §2). But individual speaker trajectories are quite different, even if the cumulative outcome for the community is an S-curve.

In this paper, we examine certain types of variation in the behavior of individual speakers and its relationship to community change, and present a mathematical model that accounts for the patterns, based on the model presented in Baxter et al. (2006, 2009; see also Blythe and Croft 2012). Although the observed variation in individual trajectories is not as regular as the recurrent S-curve of community change, there are certain patterns related to an individual's lifespan from childhood to maturity that appear to be relatively robust and have been discussed to a greater or lesser extent in the sociolinguistics literature. Our model demonstrates that the variation in individual trajectories implied by these explanations nevertheless can have a cumulative effect of regular, S-shaped community change.

Our model suggests explanations for certain patterns of language change across the lifespan and their relationship to community change. The value of models of complex systems such as speech communities is that it is extremely difficult if not impossible to collect empirical data of the linguistic behavior of all or even most speakers in a speech community, in their natural interactions with their interlocutors in various social contexts. Language, that is, the linguistic behavior of a speech community, is a complex adaptive system (Beckner et al. 2009): there is a large set of agents that interact, and the outcome of their interactions influences the future behavior of the individual agents. Hence it is difficult to evaluate proposed explanations for the samples of linguistic behavior that we do have access to. Models allow us to test different hypotheses as to their plausibility, as was done by Baxter et al. (2009) for Trudgill's theory of new-dialect formation, and by Blythe and Croft (2012) for various mechanisms proposed for the S-curve of propagation of a linguistic variant in a speech community.

Models are not a substitute for reality, of course, and one can only model limited aspects of the behavior of a complex adaptive system such as a speech community. For example, Baxter et al. (2009) and Blythe and Croft (2012) model only community change, that is, overall patterns of language change that are the cumulative effect of individual linguistic acts as speakers interact with each other. Nevertheless, the models imply certain general properties of community change, described in §2, that are required in order to account for general properties of language change, such as the rate of change and the S-curve trajectory of community change.

Also of course, overall community change is not all that sociohistorical linguists are interested in. At the very least, sociohistorical linguists are also interested in histories of

language use by individual speakers in the community as they age, and the role in language change of social groups that exist within a community, defined by such features as gender, socio-professional status, high school gang membership, and so on. In this paper, we model patterns of individual speakers as they age, that is, language change across the lifespan (Sankoff and Blondeau 2007). Our goal is to explore the effects of a plausible model of a speaker's linguistic development through her lifetime. The result is a model of community change as the cumulative effect of the linguistic behavior of individual speakers who are born, mature, die and are replaced by new generations of speakers. Our basic goal in this paper is to modify the model developed by Baxter et al. to include individual change across the lifespan, in such a way that the cumulative effect is observed patterns of community change. We focus our attention on three patterns that have been reliably observed, and certain explanations offered for those patterns. One explanation is essentially physiological/cognitive: changes in a speaker's flexibility in their linguistic behavior as they age. The other explanations are essentially social, in terms of who a speaker interacts with, and how much she accommodates to the linguistic behavior of her interlocutors, and how she weights different variants of a sociolinguistic variable.

We model three patterns here. The first is the apparent time construct, widely used to extrapolate real-time changes from a sample of behavior of speakers of different ages collected at a single time (Labov 1963; Bailey 2002, inter alia). The apparent time construct is based on a particular assumption about individual trajectories in a language change, namely that an individual speaker changes her linguistic behavior reflecting a community change through adolescence, but then more or less ceases to change her

behavior afterwards. To the extent that this observation is correct, it is attributed to a physiological/cognitive loss of linguistic flexibility post-adolescence (the critical period hypothesis). In fact, it has been documented that adult speakers also may adjust their linguistic behavior with respect to an ongoing community change, although on the whole the apparent time construct remains a reasonably accurate gauge of an ongoing change (Bailey 2002:329-30; Wagner 2012a:377). We discuss and model the apparent time construct, including the possibility of adult language change, in §3.

The second pattern is the adolescent peak (Cedergren 1988; Labov 2001:446-65; Tagliamonte and d'Arcy 2009). The adolescent peak is an anomaly in the otherwise S-curved trajectory of community change found in apparent time studies. It is assumed to result from an individual trajectory of change reflecting the child's primary exposure to different speakers at different stages of childhood (first caregivers, then older peers). We discuss and model the adolescent peak in §4.

The third pattern is one that has been only recently remarked upon. There appear to be two contrasting ways in which adults adjust their linguistic behavior with respect to an ongoing community change (Sankoff and Blondeau 2007:580; Nevalainen et al. 2011). The first is by gradual change in variable use of the incoming and outgoing variants over time. The second is more categorical behavior on the part of individual speakers, with individual speaker change happening rapidly (if a speaker changes at all). In §5, we discuss and model these two ways in which change over the lifespan takes place, and argue that it reflects in part differences in the degree to which speakers accommodate to their interlocutors in the speech community.

2. Modeling language change in the speech community

2.1. Four mechanisms of language change

Human society and social behavior, including language, is a good example of a complex adaptive system. A complex adaptive system can be characterized by the following traits (Beckner et al. 2009:1-2). The system consists of multiple entities—speakers, in the case of language—interacting with one another. The behavior of the entities (speakers) evolves adaptively on the basis of past and present interactions, and future behavior is determined by past and present interactions. The system is complex in that a range of competing factors influence the behavior of the interacting individuals, and hence the system as a whole. In the case of language, a wide range of physiological, cognitive, and social factors interact to produce the behavior of individual speakers and hence of the speech community as a whole.

The variationist approach to language change can be interpreted as a theory of language as a complex adaptive system. Speakers interact with each other in a speech community. Their linguistic behavior is variable, affected by the interaction of many different factors, social, language-internal and otherwise. A speaker's language behavior changes over time in response to the patterns of variation of linguistic behavior to which she is exposed. In this respect, the variationist approach to language change is an examples of a usage-based model (Bybee 2001, 2007, 2010), in which speaker knowledge about her language is variable and responds to interactions with other speakers (that is, language use). Language change at the community level results from the cumulative effect of language behavior at the level of linguistic and social interactions among individuals.

The mathematical model proposed by Baxter et al. (2006, 2009) is based on an evolutionary model of language change proposed by Croft (2000) that integrates usage-based and variationist approaches to language change. The central hypothesis of the model is that language change emerges from the replication of linguistic structures in utterances produced by speakers. Language use is inherently variable: replication generates variation in both phonological and morphosyntactic structure. Language change in a speech community is the cumulative effect of the evolution of the tokens of linguistic structures (called linguemes by Croft) as they are replicated by speakers and they interact with each other over time.

Croft's model is based in turn on a generalized theory of evolution by Hull (1988, 2001). In Hull's theory, evolution is change by replication, as opposed to inherent change of an entity. Hull distinguishes two distinct roles of entities involved in an evolutionary process. One role is the replicator: the entity that is replicated in the process, and which evolves through differential replication. The other role is the interactor: an entity whose interaction with the environment causes differential replication. In language change, the interactor is the speaker: the speaker interacts with her environment, in particular other speakers in the speech community, and that interaction causes the speaker to replicate certain variants instead of others.

Hull's distinction between the replicator and the interactor roles in an evolutionary process fits language change and other cultural evolution processes, in which the human agent plays the role of interactor, and the cultural product or behavior, such as the production of linguistic structures in utterances, plays the role of replicator. Hull's model also leads to a more complex set of selection mechanisms than is usually presented in

accounts of biological evolution. These selection mechanisms, described briefly below (see Baxter et al. 2009:269072; Blythe and Croft 2012:272-77), correspond broadly to different sorts of mechanisms proposed by sociohistorical linguists to account for language change.

The mechanism that corresponds to fitness in population genetics models is called replicator selection by Baxter et al. (2009:269-70): differential, that is, asymmetric weighting of variant replicators by speakers leads to differential replication of the replicators, so that one replicator (variant) is propagated and the other falls out of use, ultimately disappearing. Although the model (or the mathematics) does not say anything about what brings about the differential weighting of linguistic variants, we hypothesize that it represents differential social valuation of variants by speakers. This mechanism is associated with the classic Labovian sociohistorical model, though Labov himself notes that it was proposed by linguists before him (Labov 2001:24).

In addition to replicator selection, there are also two mechanisms for selection based on properties of the interactor, not the replicator. Differential replication of a linguistic variant by a speaker may occur as a consequence of different rates of interaction with other speakers, even if the variants produced by them are equally weighted (that is, no replicator selection is operating). This mechanism is called neutral interactor selection by Baxter et al. (2009:270-71). Another way of describing this mechanism is that how I speak depends on who I talk to. Neutral interactor selection models social network structure, an important factor in many theories of language change (e.g. Milroy 1987, although she incorporates replicator selection into her theory as well).

Neutral interactor selection is symmetric: since it is simply how often a pair of speakers interact with each other, its effect is the same on both speakers. There is also an asymmetric mechanism of interactor selection, in which a speaker accommodates (Giles 1973; Giles and Smith 1979) more to her interlocutor in her linguistic behavior than the other way around. More generally, a speaker may accommodate more to one interlocutor than another, even if she interacts equally often with both of them. This mechanism is called weighted interactor selection by Baxter et al. (2009:271). Weighted interactor selection is a plausible model of the leader-follower or adopter theories of diffusion of innovations (Rogers 1995), which have also been proposed to play a role in language change (e.g. Labov 2001:356-60, Milroy and Milroy 1985; Sankoff and Blondeau 2007; Nevalainen et al. 2011).

Finally, change can happen in finite populations of replicators by the stochastic nature of the replication process (whereby variants are produced). This process is also recognized in population genetics models in biology. It is called neutral evolution or genetic drift; we call it neutral evolution to avoid confusion with linguistic drift (Sapir 1921), an entirely different concept (cf. Baxter et al. 2009:270). One important feature of neutral evolution is its sensitivity to the frequency of the variants: a more frequent variant is more likely to propagate through the population of replicators. As a consequence, neutral evolution is a plausible model of the frequency effects documented in usage-based approaches to language behavior (e.g. Bybee 2001, 2007, 2010).

2.2. A model of speaker interaction and change

Baxter et al. (2009) and Blythe and Croft (2012) use a model incorporating these different mechanisms to examine theories of new-dialect formation and the S-curve trajectory of community change, respectively; the description of the model that follows is a summary of Baxter et al. (2009:272-77). The model assumes that linguemes are independent, that is, it does not model interactions between linguemes such as chain-shifts. Linguemes occur in variants; that is, a lingueme is a sociolinguistic variable. The speech community is made up of N speakers. Each speaker's knowledge about her language (her grammar) includes the frequency of use of each variant.

Speakers interact (that is, talk to) one another, and replicate variants in the process. The likelihood of interaction of speakers is described by a matrix G_{ij} for speaker i interacting with interlocutor j; this matrix represents the social network structure of the speech community, and hence neutral interactor selection. Speakers replicate (that is, produce) a lingueme a certain number of times in the interaction, and the variant(s) of the lingueme that they replicate is the result of a probabilistic function of the representation of the frequency of the variants in the speaker's mental grammar.

Not included in the Baxter et al. model is differential weighting of the variants, that is, replicator selection. This differential weighting plays a role in the speaker's selection of which variant to replicate. In the simplest case, one variant is selected by all speakers with an increased probability, controlled by a parameter which we will call b.

After the interaction, the speakers' grammars are updated; this corresponds to the feedback effect in the complex adaptive system. The updating process involves two variables. The first variable, λ, represents the weight assigned to the heard variants

relative to the current grammar. That is, λ represents a speaker's receptiveness to changing her grammar; it corresponds to how flexible a speaker is in adjusting her linguistic behavior to the language she hears around her. As a small fraction λ of the grammar is replaced after each interaction, the influence of previously heard tokens is reduced. We can therefore also think of λ as controlling how long tokens are remembered. The amount of the grammar occupied by a token decays as $\exp(-\lambda n)$ where n is the number of subsequent interactions taken part in by the speaker.

The second variable governing the updating of the speakers' grammars is the weight that a speaker assigns to her interlocutor's utterances compared to her own. The weight assigned by speaker i to interlocutor j's utterances is described by the matrix H_{ij}. This matrix represents the degree of accommodation that the speaker makes to her interlocutor, and hence weighted interactor selection. The speaker's grammar is thus updated until the next interaction.

Baxter et al. (2009) use this model to evaluate Trudgill's theory of new-dialect formation, using data from the Origins of New Zealand English project. Trudgill's theory advances two hypotheses about new-dialect formation as a result of the coming together of speakers from different source dialects (in this case, different parts of the United Kingdom) in a new speech community. The first is a 'majority wins' rule: the variant that is the most frequent is the one most likely to be propagated in the new dialect. The second is that no differential social valuation of variants or of speakers plays a role in the process of new-dialect formation. That is, only neutral evolution and neutral interactor selection operate in new-dialect formation (Baxter et al. 2009:271-72).

Baxter et al.'s model confirms the first hypothesis, that the majority variant is most likely to be propagated in new-dialect formation without any other social factors influencing propagation. As noted above, one trait of neutral evolution is that the most frequent variant in the population is most likely to propagate. However, neutral evolution and neutral interactor selection alone are highly unlikely to lead to the fixation of the New Zealand English dialect in the time interval that the New Zealand English dialect actually formed, given other (empirically determined, but generous) values of relevant variables in the model. One reason for this is that the time for neutral interactor selection increases linearly with the population size.

Blythe and Croft (2012) use the same model to examine the temporal trajectory, rather than the time scale, of community language change. When one variant successfully competes against another variant in being propagated across a speech community, the trajectory of the change is an S-curve (the full length of the S-curve may not be documented, and changes may cease before they have gone to completion; Blythe and Croft 2012:278-81). Blythe and Croft argue that the only selection mechanism that consistently produces an S-curve trajectory is replicator selection. Neutral evolution and neutral interactor selection (social network structure) produce highly fluctuating trajectories. Weighted interactor selection normally produces strong fluctuations; it can produce an S-like trajectory, but only under very specific assumptions representing social structures that are not characteristic of known speech communities (Blythe and Croft 2012:287-91). Finally, the time length for replicator selection does not have the same sensitivity to population size that interactor selection does.

Blythe and Croft's result implies that whatever else is going on in the propagation of a competing variant in a speech community, it must include differential weighting of the variants in order to produce the ubiquitous S-curves that are repeatedly observed. However, Blythe and Croft model only community change, that is, the population-level pattern of language change (the S-curve). Their modeling does not include generational replacement or changes in language behavior across an individual speaker's lifespan. Nor do they model the structured heterogeneity that is characteristic of all speech communities and in many cases represents the orderly propagation of an incoming variant through the community. In this paper, we attempt to remedy the first of these omissions in Blythe and Croft's analysis: the relationship between individual and community change.

3. The apparent time construct

3.1. Evidence and explanations for the apparent time construct

The apparent time construct has long been used to allow sociolinguists to make inferences about language change in progress from a single synchronic sample of speaker behavior. If there is a difference in linguistic behavior across speakers of different ages, then it is possible to infer that there is an ongoing change, with older speakers representing the earlier stage of the language and younger speakers representing the later stage. The result is that if one plots language behavior of different age cohorts, with the older speakers to the left on the graph, one should obtain the same S-curve (or part of an S-curve) of a community change that is in progress in the lifetimes of the speakers in the synchronic sample.

The inference of community change from age-related synchronic variation is based on two assumptions. The first is a hypothesis about the nature of language change across the individual lifespan: speakers are flexible in their linguistic behavior up through adolescence—the so-called critical period for language acquisition—but after adolescence their linguistic behavior is more or less fixed. This is an essentially cognitive/biological phenomenon (e.g. Lenneberg 1967). Hence a speaker's adult linguistic behavior, even decades later, reflects her linguistic behavior, and the linguistic behavior of the speech community, at the time of her adolescence. The second assumption is that the difference in linguistic behavior across different age cohorts at a single time reflects a language change in progress.

It has long been recognized that the second assumption does not apply to all age-related synchronic variation. Some age-related differences are due to changes in linguistic behavior as a speaker passes through different life stages, and the age-related differences are the result of stable age-determined variation. This phenomenon is described as age-grading, and is contrasted with the apparent time construct. Thus, one must obtain other evidence that in fact there is language change in progress and not simply stable age-determined variation.

Labov gives four logical possibilities for the relationship between individual and communal language behavior:

[TABLE 1 ABOUT HERE]

The first case represents absence of any variability in linguistic behavior, at either the individual or the community level. The second case, age-grading: individuals change their linguistic behavior as they age, but all individuals change their behavior in more or

less the same way across their lifespans, and so the community exhibits stable variation. The third case is the situation that the apparent time construct assumes: individuals of a particular generation are stable in their linguistic behavior, but across generations a language change is taking place. This characterization is actually a simplification, because individual behavior does change from infancy through adolescence, pushing the change forward; this entry in Labov's table describes adult linguistic behavior only. We will discuss linguistic behavior from childhood through adolescence in §4. Finally, community change is one in which 'all members of the community alter their frequencies together, or acquire new forms simultaneously' (Labov 1994:84).

Labov's description of 'community change' assumes that the combination of individual and community instability only occurs with rapid changes, that is, changes that happen in much less than a generation's time and therefore do not display an apparent time curve. In fact, however, there is evidence that the first assumption presented above, the critical period assumption, is also problematic: speakers do change their behavior to some degree after adolescence, even as the community is changing. This observation has led to some variation in the use of the term 'age-grading' and 'community change' (see, e.g., the discussion in Wagner 2012a). We will use the term 'community change' to describe temporally directional changes in community linguistic behavior, independent of variation in the behavior of individuals in the community; and the term 'age-grading' in the narrow sense in the table, describing age-related individual variation in the context of community-level stability (i.e. stable variation).

Changes in linguistic behavior over an individual's lifespan—that is, beyond adolescence—has been documented by numerous real-time studies, in which

sociolinguistic methods have been used to collect data for the same community and/or the same individuals at different points in time. This is possible due to the time elapsed from earlier sociolinguistic studies, and also due to the use of archives of recorded material (Wagner 2012a:377). Another source of evidence of post-adolescent change are studies of speakers adjusting to a new dialect as adults (e.g. Shockey 1984; Prince 1987).

Considerable attention has been given to changes in individual linguistic behavior connected to the fact that an individual's social status changes over his or her lifespan. Not only does an individual age, a biological/cognitive phenomenon that affects linguistic behavior, but s/he also undergoes changes in social role and in the social networks that s/he is engaged in (Eckert 1997), and hence must construct and re-construct social identities as s/he ages. In particular, significant social changes occur to individuals in American society in the transitions from high school to college to the workforce. Sociolinguistic variants index the changes in social identity (Eckert 2000), and, unsurprisingly, individuals who undergo these transitions change their linguistic behavior to some degree (see, for example, Bailey 2002, referring to an unpublished 2000 study by Cukor-Avila; Cukor-Avila 2002; De Decker 2006; Pritchard and Tamminga 2012; Wagner 2012b). However, these changes in linguistic behavior appear to involve the manipulation of linguistic variants for social purposes even when the variation is stable at the community level (age-grading in the narrow sense; Wagner 2012b), or the status of the variation at community level is unclear (Cukor-Avila 2002), as well as cases of variants undergoing community change. That is, it appears that these age-related changes reflect linguistic effects of variation in social identity and social network across a

speaker's lifespan, whether or not the variation is part of a community change in progress.

The relationship between community change and individual change across the lifespan, particularly post-adolescence, is addressed more directly by studies that combine real time and apparent time. These are studies in which the linguistic behavior of a sample of speakers from a community are analyzed by age cohort (i.e., the input to an apparent time analysis) and across at least two different time points (i.e., a study in real time). These studies are generally of two types: a panel study, in which the same speakers are tracked down and interviewed at a later time (or times); or a trend study, in which a new set of speakers is sampled at a later time, with a similar social profile to the set of speakers sampled at the first time. Panel studies are challenging to conduct, because tracking down the original speakers is not easy; so panel studies are often small (in fact, sometimes just one individual, such as Queen Elizabeth II [Harrington et al. 2000]). Trend studies are easier to conduct, particularly with a large sample, and hence more trend studies combining apparent time and real time have been published (see Table 2 below). The studies summarized in Table 2 also sample speakers across the full lifespan, including middle age and old age.

Trend studies do not record directly changes in individual behavior (Wagner 2012a:377). One can only assume that, for instance, the population behavior of an age cohort of speakers born between 1940 and 1960 that are recorded in 2010 reflects the population behavior of the same age cohort of a similar sample recorded in 1990. Panel studies provide direct evidence of changes in individual behavior, but they are often

much smaller. In one important larger-scale panel study (Nahkola and Saanilahti 2004), the panel data is given in aggregate, rather than by individual.

Both trend and panel studies of change in real time also have potential methodological problems (Bailey 2002; Tillery and Bailey 2003). It is very difficult to exactly replicate the interview context and protocols of a prior sociolinguistic study, and to exactly replicate the coding practices of the prior study (e.g. in impressionistic transcription of continuous phonetic variation into discrete variants). Hence the variation in linguistic behavior from one time point to another must be taken with a grain of salt: significant changes in behavior between the two time points may be an artifact of the data collection and analysis process. Moreover, almost all studies of apparent time and real time sample only two different time points, and usually only 2-3 age cohorts are sampled in both time points (see Table 1; usually the oldest age cohort from the first time is gone, and new young age cohort from the second time is introduced that was not born at the first time).

Finally, in surveying published studies of apparent time and real time, it turns out that most studies do not indicate whether differences in speaker behavior from one time point to the next are significant (in fact, almost all studies give only percentage data). As consequence, in presenting the results of our survey of studies of apparent and real time, we use a difference in linguistic behavior of 10% (that is, a speaker changes her use of a variant from X% to X+10% or more) as significant, following in part the significance results in Sankoff and Blondeau 2007.

Table 2 summarizes the results of several surveys of linguistic variables in apparent and real time. In these studies, data is presented that allows us to compare the behavior of

the same age cohorts (or in the case of some panel studies, the same individuals) across at least two different time intervals (in many of the studies, the authors do not make this direct comparison).

[TABLE 2 AROUND HERE]

The first observation is that in more than half of reported cases of individual changes of variants undergoing community change, adults do not change their behavior by more than 10% after adolescence. Although Trudgill does not give numerical or percentage data for his restudy of Norwich, he states that changes in adult linguistic behavior are 'in most cases rather small', and adults did not participate in more recent changes diffusing through the Norwich speech community (Trudgill 1988:37). These observations imply that on the whole, the apparent time construct is supported (Bailey et al. 1991; Bailey 2002:324; Wagner 2012a:377). When adults are advancing, three patterns appear (leaving aside the methodological caveats given above): all adults are advancing by a similar degree; older adults are advancing by a lesser degree than younger adults; and there are even cases of adults retreating from a community change. There are no reported instances to our knowledge of older adults advancing to a greater degree than younger adults in a community change. Finally, there is a more complex pattern revealed by the panel study of Montréal French /r/ that will be discussed and modeled in §5.

One must therefore weaken the assumption that adult language use is completely fixed for physiological/cognitive reasons connected to biological aging. Nevertheless, it appears that linguistic malleability is considerably reduced as speakers age, probably at least in part due to physiological and/or cognitive reasons. The cases in which older adults advance a community change but not as much as younger adults might suggest that

an individual's ability to adjust their linguistic behavior gradually decreases through adulthood. But the fact that there is at least one well-documented case of adults retreating from a community change (the Montréal French future described in Wagner and Sankoff 2011, Sankoff et al. 2012) demonstrates that the changes in adult linguistic behavior cannot be purely a consequence of decreasing ability to change as speakers age. If that were the case, adult speakers might not advance at all, but they would not retreat. Instead, a retreat would represent a change in the valuation of the linguistic variant, as suggested by Sankoff, Wagner and Jensen: the more formal nature of the outgoing inflectional future variant gives it a higher value for the older speakers. The fact that use of the inflectional variant is highly sensitive to socioprofessional status of the speaker confirms that the sort of changes observed in adult speaker response to a community change in progress is at least partly social, not biological.

3.2. Modeling apparent time

To model apparent time, we modified Baxter et al.'s model (Baxter et al. 2006, 2009) to allow the parameter λ, which controls receptiveness, to change over a speaker's lifetime. For the apparent time construct to be possible, this change should happen in such a way that a speaker's ability to change is greatest in childhood and adolescence, and is considerably reduced in adulthood. For simplicity, the way that λ changes as a function of a speaker's age was made the same for every speaker. Because speakers in the population have a variety of ages at any given time, a range of λ values are present in the population. As we have already discussed, in real populations, some speakers change more than

others, even among those having the same age. However, this approximation is sufficient to capture the aggregate behavior across the whole population.

There are many possible choices for this λ function. We could, for example, set a certain value for young speakers, then make a step change to a lower value when they reach a certain age. Instead, we chose to model the receptiveness λ as a function that decays smoothly with age. In this way, we do not artificially impose a change in cognitive behavior or abilities of speakers at a specific age. We seek a function that decays sufficiently quickly that speakers' linguistic malleability becomes significantly reduced after a certain age (though not necessarily completely eliminated), but slowly enough that speakers remain adaptable into late adolescence. One suitable function is an exponentially decaying function

$$\lambda(a) \propto \exp(-\beta a) \tag{1}$$

where a is the age of the speaker concerned, and β controls the speed of the decay.

[FIGURE 1 ABOUT HERE]

To choose a sensible value of the parameter β, we simulated a speaker of a certain age, who initially uses a conventional variant almost exclusively, and who applies a replicator selection boost b to a second variant, while also interacting with a large population who exculsively use the new variant. These conditions allow us to estimate the maximum change a speaker of a given change is able to make in their remaining lifetime. An example is shown in Figure 1a. A speaker below a certain threshold age is able to reach categorical usage of the new variant. Older speakers, on the other hand,

initially move towards the new variant, but are not able to complete the change. We see that speakers older than about twenty-five years do not change at all. The threshold age depends strongly on the parameter β and weakly on the strength b of the replicator selection, and on other details such as how many interactions a speaker participates in per year. The results for this decay function depend only weakly on the parameter h. We chose the parameter $\beta = 0.4$, as this gives fall-off at approximately 15-20 years of age, using reasonable values for the other parameters.

The disadvantage of using function (1) is that the fall-off may be too dramatic: speakers are able to change only a very little during adulthood. We can instead consider a powerlaw decay function

$$\lambda(a) \propto a^{-\gamma}. \qquad (2)$$

Choosing $\gamma = 2$ for this function again gives a fall-off in speaker response at around 15-25 years of age (though the precise age is more dependent in the replicator selection parameter b). Now, however, adult speakers remain able to continue to change, albeit much more slowly as shown in Figure 1b. Even adult speakers are able to change their usage by a few percentage points. Unlike the exponential decay, the amount of change adults are able to make depends on the accommodation parameter H. In the example shown, we used a moderate value of $H = 0.02$. Larger values would allow adults to change even more. As we will see, using either function (1) or function (2) gives qualitatively similar results, indicating that the exact choice of the decay function does not unduly influence our results.

[FIGURE 2 ABOUT HERE]

The full population is fixed at a certain size N. Periodically one speaker is chosen to be removed, and is replaced by a new speaker. (New speakers are given an age of 1, corresponding to roughly the age at which they may begin participating in the speech community.) The probability that a speaker of age a will be removed is given by an exponentially increasing hazard function $\exp(\omega a)$. With $\omega = 0.085$ this function results in a good approximation of the age distribution found in a modern developed country (Gompertz, 1825); see Figure 2. The mean longevity is 80 years. When a new speaker is created, she is essentially a 'blank slate', without any predefined grammar knowledge. In her first conversation, she will take whatever tokens she hears from her interlocutor to set her initial grammar value. Subsequently she will interact as normal, adapting to her own utterances and those of their interlocutors. For this initial investigation, we chose the most simple network of social interactions possible: every speaker is equally likely to speak to every other, that is, the non-diagonal entries in the matrix G_{ij} are all equal. Similarly, the accommodation weights H_{ij} were also set to equal the same value H. In order to choose an appropriate order-of-magnitude estimate for the frequency of interaction, we refer to estimates made in Baxter et al. (2009:282) that a speaker may hear a million or more tokens of a given linguistic variable in her lifetime, corresponding to somewhere on the order of ten thousand tokens heard per year. If around 10 tokens are produced in a typical conversation, this corresponds to around 1000 such conversations per year.

All speakers in the population are presumed to have learned the same preference for the incoming variant, and so apply replicator selection with the same parameter b. As we

have already mentioned, this is the most likely explanation for the frequently observed S-curve trajectory of language change. The choice of b controls how quickly the change happens. A very small value may lead to a change that occurs over centuries, while a larger value means the change occurs more quickly. The apparent time effect can only be observed in changes make significant progress within a single lifetime. We set b= 0.01, which produces a change whose transition takes between 80 and 150 years for the other parameter settings chosen.

[FIGURE 3 ABOUT HERE]

The typical behavior of this model is represented in Figure 3. The new, preferred, variant comes to dominate the population, its usage rising in a typical S-shaped curve. This is shown by the solid black line in the left-hand side of the Figure. The population can be subdivided into cohorts of speakers born within time intervals of a certain length. In the left-hand side of the figure, each fine (magenta) line shows the average grammar of each such cohort, here using 10-year windows, over time. We see that each cohort initially moves rapidly towards usage of the incoming variant, but after reaching about 20 years of age, slows down markedly and settles at a certain mixed usage of the two variants. This is the effect of the decaying λ function. If, at a specific moment, we plot the mean grammars of each cohort as a function of their age, we recover an apparent time curve similar to that we would expect to find in a survey carried out at that specific time. An example is shown in the right-hand side of Figure 3, which represents the mean grammars of each cohort at the time indicated by the vertical dashed line in the left-hand plot.

As the time scale of a change is limited by the lifetime of speakers, an apparent time curve will not represent the full extent of the change. Even over the range of the change that the apparent time curve is able to represent, the shapes of the real-time and apparent time S-curves are different. The advantage of numerical simulation is that we can try many different parameter combinations, to quantify this difference. In Figure 4 we plot the time taken for the apparent time curve to rise from 20% to 80%, as a function of the equivalent time interval for the real time curve, for a large variety of parameter settings. We see that the apparent time interval is always shorter than the real time interval. Each cohort is leading the change at approximately the same time as they start to slow down in their change. Their adult usage therefore represents this leading value. If we compare the real time taken for the leading cohort to go from 20% to 80% usage against the apparent time estimate, as shown in the right-hand side of Figure 4, the agreement is much better.

[FIGURE 4 ABOUT HERE]

We repeated the experiment using the powerlaw decay function (2) which does not decay as quickly as the exponential function. An example of the result of using this function is shown in Figure 5, using $\gamma = 2$. Notice in the left-hand figure that now each cohort's trajectory, after a fast initial rise, slows down significantly, but continues to rise slowly throughout adulthood. In the middle period of the change, adults cohorts increase their usage of the incoming variant by 10-15%. Nevertheless, when we take a sample at a specific time, we still see an apparent time S-curve that closely matches the real time usage of the leading cohorts.

[FIGURE 5 ABOUT HERE]

4. The adolescent peak

4.1. Evidence and explanations for the adolescent peak

The apparent time construct described in §3 has one consistent anomaly in apparent time data collected in various studies. For those studies that include children (preadolescents), it has been observed that children as a cohort have a lower proportion of the incoming variant in a community change than do adolescents. That is, children are not as progressive in the community as one might expect: as the youngest cohort, one would expect children to be the most advanced users of the incoming variant. One of the earlier observations of this pattern is Cedergren's (1988) study of CH lenition in Panamanian Spanish. The anomaly occurred in both her initial data from 1969 and the data she collected in 1982-84 (Cedergren 1988:53-54). In fact, for the Panamanian data, the peak occurs in early adulthood; the adolescent cohort is the one that is not as progressive. (Cedergren does not report data for preadolescents.) Cedergren suggests that the peak is due to the response of the young adult cohort to the linguistic marketplace (see §3).

Labov (2001:446-65) also documents the adolescent peak in sound changes in progress in Philadelphia English. In the Philadelphia English data collected by Laov and his colleagues, the peak is in adolescence, and the trough is in the preadolescent cohort. Labov assumes the explanation that the preadolescent trough is due to the fact that children acquire language primarily from their caregivers; since the caregivers are a full generation older than the children, their use of the incoming variant is not as advanced as the use of adolescents (Labov 2001:447). As the child grows older, she is exposed to the wider speech community, including older peers, and begins incrementation of the change, that is, increased use of the incoming variant. The change continues to advance because

each new generation of children starts from a somewhat higher base level (the level of use of their caregivers) and has more time to increment the incoming variant higher than their older peers until they reach adolescence and their language use stabilizes (Labov 2001:455; he acknowledges further change beyond adolescence but in discussing the adolescent peak uses the simplifying assumption of stabilization at adolescence; ibid., 454).

Labov focuses on an asymmetry between language change across the lifespan between females and males. In the Philadelphia English data, female preadolescents clearly exhibit the adolescent peak in apparent time for changes led by females, whereas males exhibit a more confusing pattern for the same changes in preadolescent years (compare Figs. 14-9 and 14-10 in Labov 2001:458-59). Labov argues that males do not participate in the incrementation process, and therefore lag behind females at about a generation's length—that is, they retain the system they acquired from their caregivers, who are a generation behind them (Labov 2001:457).

Tagliamonte and d'Arcy (2007, 2009) investigate the adolescent peak through apparent time studies of morphosyntactic changes in Toronto English that include preadolescents. They find that the adolescent peak occurs in morphosyntactic changes in progress as well as phonological changes. They also find in their data that male speakers as well as female speakers exhibit an adolescent peak, albeit not as prominent as the peak in female speakers' behavior, even for changes dominated by females (Tagliamonte and d'Arcy 2009:98). They do not offer an explanation for this difference between Labov's results and theirs, suggesting that a finer-grained analysis of social structure and behavior might offer an explanation (ibid.).

Tagliamonte and d'Arcy also note that the adolescent peak appears to be less prominent in the early stages of a change, when the frequency of the incoming variant is low, in their data (Tagliamonte and d'Arcy 2009:96). Their data also suggests that the peak appears to be less prominent in slower changes (ibid., 96, 99 also citing Labov 2001:446). Finally, they also suggest that in rapid changes, a peak may be prominent even at a late stage in the change, when it is nearing completion (ibid., 96).

The explanation for the adolescent peak offered by Labov and echoed by Tagliamonte and d'Arcy assumes that a young child acquires the system of their caregiver, but it does not indicate how the child acquires that system. We hypothesize that the child acquires the system of their caregiver because the child is overwhelmingly if not exclusively exposed to the caregiver's system, and not the language behavior of other members of the speech community, at first. Evidence from the acquisition of phonological variables in the new town of Milton Keynes and in the city of Philadelphia supports the view that the youngest children most closely follow the linguistic behavior of their caregivers, and later shift towards their peers (Kerswill and Williams 2000; Roberts 1997a; cf. Labov 2001:423-29; Tagliamonte and d'Arcy 2009:64-65).

The child may also acquire the differential weighting of linguistic variants from her caregiver via the caregiver's style shifting (Labov 2001:437) and other cues of the social conditioning of variable linguistic behavior. Vihman's study of the differentiation of English and Estonian in simultaneous bilingual acquisition supports the hypothesis that a child develops social awareness, including sociolinguistic awareness, in the second half of her second year (Vihman 1985:313-14; Kagan 1981). On the other hand, in Roberts' analysis of *-t/-d* deletion in Philadelphia English, 3- and 4-year-olds had not yet mastered

the social constraints on the variable (Roberts 1997b:365). However, boys and girls were learning culturally-based gender roles by this age (ibid., 368-69), indicating that social awareness is developing at that time. At any rate, the child will acquire the differential weighting of linguistic variants as she begins to interact more extensively with other members of the speech community as she matures, and has a wider range of social cues to linguistic behavior to draw inferences from.

4.2. Modeling the adolescent peak

To investigate this hypothesis, we further modified the model of the previous section. We introduced a more complex network of social interactions, the matrix G_{ij}, which changes over time. As before, we consider a population of speakers in which periodically an older speaker is removed and a new young speaker is introduced. But now each new speaker is associated with a "caregiver" speaker, chosen at random from among existing speakers within a certain age window (e.g. 20 to 35 years old). The age window is chosen to be after the typical time at which a speaker's grammar ceases to change rapidly. The new speaker initially only interacts with this caregiver. As the child grows older, she steadily adds new connections one by one, representing the broadening sphere of influence they encounter as they progress through life. The interaction probabilities G_{ij} are rebalanced with each change so that each speaker interacts roughly equally often. The overall effect is that the youngest speakers interact in a very focused way with their caregiver, while the oldest speakers merely interact with a relatively random sampling of speakers from throughout the population.

[FIGURE 6 ABOUT HERE]

When we ran simulations of this model, we indeed found that an adolescent peak frequently occurred in apparent time curves generated from the simulations (though the random fluctuations inherent in the process mean that it does not always occur – see for example Figure 7). An example is shown in Figure 6. If we look closely at the trajectory of each cohort, we see that they begin with a usage below that of the leading cohort, before catching up, overtaking to become the leaders, and then falling behind each subsequent cohort. This progression produces the adolescent peak in the apparent time curve.

[FIGURE 7 ABOUT HERE]

In Figure 7 we plot apparent time curves at multiple real-time points through the change. The peak is less prominent at the beginning and end of the change (lowest and highest curves in the Figure) and most prominent in the middle of the change, when the rate of change is fastest. This is consistent with the observations of Tagliamonte and D'Arcy (2009) that we noted above. They also suggest that more rapid changes should exhibit a stronger peak, and this is again observed in our simulations. We measured the difference in usage of the incoming variant between the youngest and second-youngest 10-year cohorts at each time during the change, and recorded the largest difference observed. We repeated this for simulations using several different replicator selection strengths *b*, to control the speed of the change.

[FIGURE 8 ABOUT HERE]

In Figure 8 we plot the maximum peak height (the maximum difference between the mean usage of the first two cohorts) against the time of the change (time taken for the

population usage to go from 20% to 80%). Measurements were averaged over 10 repetitions.

[FIGURE 9 ABOUT HERE]

To understand how frequently and consistently the adolescent peak appears, we repeated the simulation one hundred times with the same parameter settings. For each realization, we measured the difference in usage between the youngest two cohorts at multiple times throughout the change. These results were filtered to only include times when the second youngest cohort had a usage between 30% and 70%, so as to avoid times when the usage was very small or very large, as the adolescent peak is much less prominent in these circumstances, because most speakers, young and old, have simular usage patterns. A histogram of the resulting measurements is shown in Figure 9. Positive values correspond to an adolescent peak (second cohort leading first cohort). We see that an adolescent peak appears very frequently, in about 70% of cases, with the average difference being 5.1%. For comparison, we also plot the distributions of differences between the second and third and the fourth and third cohorts. These results confirm that our hypothesized mechanism of the child being at first overwhelmingly exposed to the caregiver's system, and not the language behavior of other members of the speech community is sufficient to produce the observed adolescent peak.

These experiments were all carried out with an exponentially decaying λ. If instead the power law function is used, an adolescent peak is still observed, as can be see in Figure 10.

[FIGURE 10 ABOUT HERE]

We now examine the robustness of these results, by considering variations of several aspects of this model, in order to see whether the absence of any of them also destroys the adolescent peak effect. Keeping the network development the same (i.e. new speakers initially speak exclusively to a single "caregiver", and gradually add new contacts through their lifetime), we tried choosing the first primary contact in different ways.

First, we tried the case in which a speaker's first primary influence was a "peer" - a speaker chosen with an age in a much younger window, 1 to 15 years old. In this case the adolescent peak effect is essentially removed, and instead the youngest cohort is almost always leading. In the same statistical analysis as described above, the youngest cohort led 63% of the time. The distribution of peak sizes is shown in Figure 11. If instead we choose the primary influence to be even older than in our first model, the adolescent peak becomes larger. With primary influence chosen in the age window 30-45 years, the second cohort led 81% of the time, with an average difference in usage of 8.2%. We next tried choosing the caregiver at random from the whole population, rather than from a specific age window. The majority of speakers in the population are adults whose usage lags behind the leading teenagers at any given time, so adopting the system of a randomly chosen contact will generally cause the child to adopt an "old" level of new variant usage. The resulting distribution of peak sizes is also shown in Figure 11. An adolescent peak still occurs, and is on average even larger than any of the previous models. Because a cohort trajectory averages over multiple speakers, the overall effect is very similar to that if all children chose a middle aged caregiver. This indicates that the exact age of the caregiver is not important to produce an adolescent peak, so long as they are an adult that is no longer in a leading cohort.

[FIGURE 11 ABOUT HERE]

Next, we relaxed the condition that a child speaker should interact primarily with a single speaker in the first years of life. We ran the model again, with the same overall pattern of connections in the population as a whole, but now this contact network was fixed, and speakers were randomly assigned a position in the network. This means that new speakers generally have multiple interlocutors, and some adults only one. Surprisingly, even this randomized model frequently produced an adolescent peak, although somewhat more variably than in the previous cases. The second cohort led 65% of the time, with an average difference between the first two cohorts of 3.7%. Finally, we compared the statistics for the original model, with every speaker speaking equally frequently to every other speaker. This model has no special network features designed to produce an adolescent peak. We found that even in this case, an adolescent peak appears, and very frequently. For parameters chosen to give a change time similar rate of change to that seen in the other model variations, the adolescent peak appeared 88% of the time. As for the randomly chosen caregiver model, in these last two cases, a young speaker receives input from a broad range of speakers of differing ages. Together, these provide a usage pattern similar to the population average, which naturally lags behind the leading cohorts. These results show that this average influence is sufficient to produce an adolescent peak. The main statistics of the adolescent peak for these model variations are given in Table 3:

[TABLE 3 ABOUT HERE]

We conclude that the adolescent peak appears robustly whenever a child is exposed exclusively or even simply on average to the speech patterns of adults (older than the

critical period). Exclusive exposure to peers is sufficient to remove the adolescent peak effect, and in fact appears to be the only means to do so.

5. Two patterns of language change across the lifespan

5.1. Evidence and explanations for the two patterns

In the preceding sections, we described well-known, empirically observed patterns of language change across the lifespan that are directly connected to the course of language change in the community: the apparent time construct and the adolescent peak. We showed that these patterns can be modeled with quite simple assumptions based on the results of prior work on the mechanisms of community change (Baxter et al. 2009; Blythe and Croft 2012). In this section, we describe a pattern of language change across the lifespan that has been observed but not (yet) attracted much attention. We show that this pattern also can be accounted for in a simple fashion in our model, but using another variable, one that represents the degree of accommodation of a speaker to her interlocutors in the community.

In §3, we discussed the observation that speakers may continue to change their variable linguistic behavior in adulthood. As we observed there, usually (though not always), this change is in the direction of the community change. In most cases, the change in linguistic behavior in adulthood is relatively gradual. However, results from the largest panel study which describes each individual's trajectory of change, Sankoff and Blondeau's (2007) analysis of Montréal /r/ in 1971 and 1984, show a different pattern: 'this change differs from others described in the literature in one important way: the relative lack of stable variation. More speakers tended to be categorical than variable, and

those who changed did so very rapidly' (Sankoff and Blondeau 2007:580). Specifically, ten speakers were near categorical (≥ 83%) in their use of the outgoing variant [r], and ten speakers were near categorical (≥ 85%) in their use of the incoming variant [R] (Sankoff and Blondeau 2007:571); of the twelve remaining speakers in the panel, nine increased their use of [R] between 1971 and 1984 to a significant degree (the increase in use ranging from 17% to 74%; the other three increased or decreased their use to a nonsignificant degree).

Sankoff and Blondeau contrasts this pattern of little change or categorical change in Montréal /r/ across the lifespan to gradual changes across the lifespan for vowels in the *Atlas of North American English* (Labov et al. 2006). They suggest that the sharp changes across the individual lifespan may be due to the fact that a change in the pronunciation of /r/ does not have as great an effect on the phonological system as the North American vowel shifts, which are often chain shifts involving multiple vowel phonemes (Sankoff and Blondeau 2007:580-81). However, as we observed in §3, there are many examples of adult change in a variety of morphosyntactic as well as phonological variables whose systemic effects appear to be variable; and there seems to be no pattern to the magnitude of changes. Admittedly, these other cases of changes in adult lifespans are reported as averages over panels, or averages over speakers in a trend study. For these, an average of gradual change by adult cohorts might hide major shifts by individuals, but this would only be true if some speakers are changing strongly away from the incoming variant, which we saw was rare (see §3.1). And an average of major change by adult cohorts would mean that at least some individual speakers are substantially changing the proportion of their use of a linguistic variant across at least part of their adult lifespan.

Hence, it may not be the case that the trajectory of individual changes in linguistic behavior for a community change in progress is attributable to the role of the linguistic variable in the linguistic system.

Nevalainen and colleagues also compare individual change to community change in real time in the Corpus of Early English Correspondence (Nevalainen and Raumolin-Brunberg 2003; Nevalainen et al. 2011). They also observed that in some changes a greater proportion of speakers are relatively more categorical (their 'progressive' and 'conservative' categories) than the remaining speakers (their 'in-between' category). They compare the shift from negative concord to sentential negation with an indefinite and the shift from the gerund in *-ing* with an object introduced by *of* to an object without *of*; the former shift has many fewer in-betweens than the latter shift (Nevalainen et al. 2011:25). They argue that the rate of change contributes significantly to this difference: faster changes, such as the shift away from negative concord, leads to fewer in-between speakers than slower changes such as the shift to a gerund with a direct object (ibid., 25, 35).

5.2. Modeling the two patterns of language change across the lifespan

We agree with Nevalainen et al. 2011 that the rate of change plays a significant role in the different distribution of individual behavior that is observed. However, we propose that the differences are due not only to the relative weighting of the linguistic variants (that is, *b*), a primary determinant of the rate of change, but also to a second social variable the degree that a speaker accommodates to other speakers in the community (that is, *H*). A speaker who accommodates more easily to other speakers will undergo change

gradually, trending towards the community mean for the variant at a given point in time. A speaker that accommodates less easily to other speakers will retain the outgoing variant, and if and when she does change to the incoming variant, will do so rapidly. These two patterns of individual accommodation lead to two different distribution of individual behaviors at different stages of a community change.

By varying the value of *H* we can explore these different patterns of change in our model. In Figure 12 we show some examples of different cases. The solid black lines represent the average usage over the whole population, while the dots are the usage of a small number of randomly sampled speakers at multiple times. In all cases, the overall mean population usage of the incoming variant follows a smooth S-curve. However when *H* is small, as in the left-hand plot, speakers cluster near categorical usage of one or other variant (the distribution between them determining the population mean). On the other hand, when H is large, the right-hand plot, individual speakers change more smoothly through variable usage , more closely following the population wide S-curve. The middle plot illustrates a case between these two extremes, with still a large number of categorical speakers, but also a significant number with mixed usage.

[FIGURE 12 ABOUT HERE]

The difference between the patterns is most evident in the middle of the change. At the beginning and end of the change, the whole population uses a large fraction of either the old variant or the new one, and so necessarily all speakers will also have near-categorical usage. This can be seen clearly when we consider the development over time of the standard deviation of the individual speaker usage, as plotted in Figure 13. The standard deviation is very small at the beginning and end of the change, and is largest in

the middle. A peak standard deviation above about 0.29 indicates a polarized pattern, with speakers clustered near the boundaries (as seen in the left-hand plot, which is for a small value of *H*). A peak standard below this value indicates that speakers are centrally clustered, as seen in the right-hand plot (for large *H*). The boundary value 0.29 is simply the standard deviation we would find if speakers were uniformly distributed across the whole range from 0% to 100%. We also calculated the standard deviation separately for speakers 0-20 years of age (blue line) and 40-60 years of age (green line). For small *H*, the curves are very similar to the overall population curve, though peaking a little earlier for young speakers and a little later for older speakers. When *H* is large, however, we see that the youngest speakers are more diverse, and the older speakers more homogeneous than the population as a whole.

[FIGURE 13 ABOUT HERE]

Interestingly, the values of *H* for which we see the different patterns also depends on the strength of replicator selection (*b*). We can understand this in the following stylized way. Speakers encounter two conflicting forces. Replicator selection drives them toward categorical usage of the new variant, and accommodation drags them towards the average usage of their interlocutors. For the distribution of speaker grammars to be centralized, the accommodation parameter *H*, must be large enough to overcome not only a speaker's natural tendency towards categorical behavior but also the force of replicator selection *b*. This pattern is summarized in Figure 14a for the exponential λ decay function (1). The degree of polarization, as measured by the population standard deviation at the mid-point of the change, is indicated by the color in the plot, with blue representing small standard deviation, i.e. the centralized pattern, and orange-red representing large standard

deviation, the highly polarized pattern of change. The thick red line represents the crossover value, 0.29. We see that for larger *b*, the value of *H* needed to reach the centralized pattern also increases.

[FIGURE 14 ABOUT HERE]

Varying the parameters *b* and *H* also has an effect on the time taken for the change. As we might expect, the strength of the replicator selection, *b*, has the strongest effect. Increasing *b* increases the tendency of speakers to produce the incoming variant, leading to faster adoption. However, for a given value of *b*, changing *H* also affects the time taken for the change, with stronger accommodation (larger *H*) corresponding to longer times, as the tendency of younger speakers to switch over rapidly to use of the new variant is held back by the influence of older speakers still using the old variant. The times taken for the change are indicated in Figure 14a by white lines tracing contours of equal change time. The shortest times occur in the lower-right region of the plot, and we see that there is a minimum time for a change of approximately 60 years. This is due to the aging of the speakers. The change cannot be completed while there are older, inflexible speakers still using or partly using the original variant. In this region, then, the change time becomes independent of the model parameters, and depends only on the demographics of the population. The same effect can be observed in Figure 4a, in which the apparent time for the change (corresponding to the leading cohort of young speakers) can take very small values, but the real time for the change reaches an asymptotic value. This effect still occurs even when adult speakers are still able to change somewhat, as shown in Figure 14b, which was generated using the powerlaw λ decay function (2) with $\gamma = 2$.

The most detailed published data of the distribution of individual linguistic behaviors at different stages of a community change is for the Corpus of Early English Correspondence. Nevalainen and Raumolin-Brunberg (2003:101-9, Appendices 5.1-5.3; see also Raumolin-Brunberg 2005) gives percentages of use of the incoming variants by individuals in the Corpus of Early English Correspondence for three changes: 3rd person *-th* to *-s*, 2nd person subject pronoun *ye* to *you* and relativizer *the which* to *which* (only individuals for whom there were at least 10 tokens in the 20-year time intervals used were included; ibid., 92). The last change is more complex than the first two, in that *the which* and *which* were not only competing with each other, but both were competing with the outgoing WH-Prep (e.g. *whereby*), and themselves occurred in two variants for prepositional relatives (preposition before *which/the which* vs. stranded preposition). For this reason, we excluded *the which/which* from analysis. For the community change, published numerical data is available for the 3rd person singular verb form and 2nd person pronoun (Nevalainen 2000:360, Tables II-III). The *-s/-th* and *you/ye* changes are rapid, so we compare them to individual data on the slower change to gerund plus direct object generously provided to us by Terttu Nevalainen and Helena Raumolin-Brunberg (data on the community change for gerund plus direct object is found in Nevalainen and Raumolin-Brunberg 2003:66, 219).

[FIGURE 15 ABOUT HERE]

As we established in simulations, the peak value of the standard deviation of individual usage can be used as an indicator for the different patterns of change. In Figure 15a we plot the overall mean proportion of *-s* for each of the 20-year time intervals from 1560 to 1659 (squares). We also plot the standard deviation of the individual usage for

each window (triangles). Just as in the model, the standard deviation rises to a peak at about the midpoint of the overall S-curve, before reducing again as the change nears completion. In this example, the peak standard deviation value is about 0.33, indicating that this change is of the polarized type. The solid curves in the figure are model outputs. Model runs were carried out at a range of parameter settings (*b* and *H*), and the output most closely resembling the standard deviation and time of change of the data was chosen. The parameters corresponding to this particular choice are indicated in Figure 15a. The agreement with the standard deviation curve is excellent, and the S-curve is quite well fitted. Similar results were obtained using the powerlaw λ decay function. The agreement of the model with data is seen even more clearly in Figure 15b, where we plot the distribution of individual speakers (grey bars) at different times through the change, along with the corresponding population distribution found in the model. In particular, the weakly polarized distribution in the middle of the change is well predicted by the model (solid curves).

[FIGURE 16 ABOUT HERE]

The data for *you/ye* are not as clear, with the standard deviation fluctuating rather than following a smooth curve; see Figure 16a. Nevertheless, we estimated the peak standard deviation to be about 0.38, suggesting an even more polarized pattern of change than *-s/-th.* The simulation curve cannot be as closely matched; nevertheless the predicted highly polarized distribution, now found at several intervals throughout the change, once again matches the distribution of speaker usages quite well; see Figure 16b.

The data for the gerund plus direct object change are even more difficult to fit. Data were available for 20 year intervals from 1440 until about 1680, but there is a great deal

of fluctuation in the early part of the change, in large part due to the very small number of speakers recorded in those years. For this reason we chose to only fit the data from 1540-59 onwards, from which time the curve is more stable. At this point the average usage had already reached around 30%. The S-curve appears to saturate at around 80%, a possibility that is not catered for in the model (see Figure 17a). Furthermore, this causes the peak standard deviation in our fitted curve to occur at a different time to that in the data.

[FIGURE 17 ABOUT HERE]

A further caveat in this example is that the distribution for such centrally-peaked changes will be different for speakers of different ages. Assuming the authors in the CEEC corpus are all adults, the population standard deviation including children as well may well have a slightly higher peak. Nevertheless, we were able to find simulation curves that approximately matched the data. The largest standard deviation value observed, 0.26, is just below the threshold of 0.29, meaning that there are a significant fraction of mixed-usage speakers throughout the change. The histograms of individual usage are again well matched by the model simulation curves. We see indeed that the speakers are mostly in the middle of the range in the intervals 1580-99 and 1600-19, indicating that this change is a different pattern from the previous two.

A final example is available in the Sankoff and Blondeau's (2007) study of Montréal /r/ in 1971 and 1984. These panel studies tracked the changes in individual speakers. Nevertheless, we can estimate the population distribution of usages from the usages of the 36 speakers in each of the two surveys. Exact usage proportions were available for 9 highly variable speakers who had intermediate usage at one or both times. The ages and

usage percentages for the remaining speakers were estimated visually from Sankoff and Blondeau (2007:572, Figure 3). An apparent time S-curve for the change was derived using data from Sankoff, Blondeau and Charity (2001).

The standard deviation for the speakers in the 1971 study for which we have individual data is 0.39, while for 1984 data it was 0.40. If instead we divide the speakers into age cohorts, and find the standard deviation for each cohort, the largest standard deviation was for speakers born between 1937 and 1951, being 0.43 for the 1971 data or 0.44 for 1984 data. All of these values are far above the threshold value of 0.29, meaning that this example is the most polarized of the four we have described. This is of course clear from the individual data, in which 23 of the 36 individuals had near categorical usage of one variant or the other both in 1971 and 1984.

As we observed in §3, apparent time curves tend to underestimate the duration of a change. Looking back at Figure 4a, we see that an apparent time change over 30 years corresponds to a real time of 50-60 years. The best-match model parameters were found by calculating apparent-time curves from the simulations and matching them to the Montréal apparent time data. The result is shown in Figure 18. The red curve is apparent time, the black dotted curve is real time, both from the simulation. Also plotted is the evolution of the standard deviation, which also roughly matches the data. This example corresponds to the pink dot marked on the contour diagram in Figure 14a.

[FIGURE 18 ABOUT HERE]

The parameters corresponding to the best matches to the four data sets discussed are marked by pink circles in Figure 14a and 14b. We see that they all fall in different areas of the plane, with the gerund plus direct object change falling on the centrally peaked part

of the diagram, and the others falling in the region corresponding to the polarized pattern. There seems to be no particular pattern to their positioning. In particular, the *s/th, you/ye* and gerund plus direct object changes, the data for which all come from the same corpus, and hence occurred in the same population, neither agree in the value of H or in b.

We note also that the contours of time-of-change run almost parallel to the contours of the population standard deviation across a large part of the diagram. In particular, more centrally peaked areas tend to have longer times than more polarized areas, in agreement with the observation of Nevalainen et al. (2011) However we disagree as to the cause: it is not that faster changes cause more polarization, but that the parameters leading to faster changes also happen to lead to more polarization.

6. Conclusions

We have shown that a number of different mechanisms of language change in the model used here are responsible for different phenomena that have been observed in community change and in individual change across the lifespan. The most general property of community change across time, the S-curve, is determined by replicator selection, that is, differential weighting of variants of a linguistic variable, following the classic Labovian model (Blythe and Croft 2012). Under that umbrella of community change, a number of mechanisms interact to bring about variation and change across the individual lifespan. Our model offers another means to evaluate qualitative explanations that have been offered for individual change that cannot be practically tested (or cannot be tested at all) in a real-world, large-scale speech community.

The apparent time construct has been explained in terms of a decline in flexibility in adopting novel variants, or frequencies of novel variants in appropriate linguistic and sociolinguistic contexts, as a speaker ages. We modeled this decline in flexibility with a gradual function that declined rapidly around adolescence or early adulthood (§3.2). The results straightforwardly confirmed the basic account of the apparent time construct. Of course, adults are able to alter their linguistic behavior to some degree. We have accommodated for this possibility by using a decay function that allows some adult change. To model the specifics of such changes will require developing a model of social differentiation within a speech community; we have not done this here, although it is an avenue for future research.

The two patterns of individual change (more centralized and more polarized), observed by Sankoff and Blondeau and by Nevalainen et al. were explained by Nevalainen et al. as a consequence of the rate of change in the linguistic variable. While we agree that there is a correlation between the pattern of change and the rate of change for a linguistic variable, our model suggests that the correlation and the two patterns can be explained by the interaction of the differential weighting of the variants and the degree of accommodation of speakers. Speakers who accommodate more readily to their interlocutors will change more continuously (the centralized pattern), while speakers to accommodate less readily will change more suddenly, if they change at all (the polarized pattern).

Perhaps the most interesting result of the model has to do with the explanation for the adolescent peak. The adolescent peak has been explained as a consequence of a child being exposed to older caregivers before associating more widely with older peers and

incrementing their language use correspondingly. This can be modeled in terms of a specific dynamic network structure: a child interacts with her caregivers at first, then gradually comes to interact with more members of the speech community. In fact, however, the adolescent peak results from any network model in which a child is exposed to adult speakers of any age; only if the child interacts solely with caregivers in the adolescent cohort does the adolescent peak disappear. On the one hand, this might be a more realistic model of children's interactions: they interact with a wider range of speakers, including adults of differing ages, from early in childhood. On the other hand, it might suggest that at least for this phenomenon, social network structure does not play a significant role in bringing it about.

The model presented here is relatively simple in terms of the number of variables and mechanisms of interaction, variation and change posited. Yet it is able to model the interplay between community change and individual change across the lifespan. It is also able to model relatively fine-grained phenomena such as the adolescent peak and the two patterns of language change described above. As such, it suggests that there can be a fruitful marriage of empirical sociolinguistic research with mathematical modeling of proposed explanations for sociohistorical linguistic variation and change.

References

Arias, Elizabeth. 2014. United States Life Tables, 2009. *National Vital Statistics Reports* 62, 1-61.

Ashby, William J. 2001. Un nouveau regard sure la chute du *ne* en français parlé tourangeau: s'agit-il d'un changement en cours? *French Language Studies* 11.1-22.

Bailey, Guy. 2002. Real and apparent time. *Handbook of language variation and change*, ed. J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes, 312-32. Oxford: Blackwell.

Bailey, Guy, Tom Wikle, Jan Tillery and Lore Sand. 1991. The apparent time construct. *Language Variation and Change* 3.241-64.

Baxter, Gareth J. Richard A. Blythe, William Croft and Alan J. McKane. 2006. Utterance selection model of language change. *Physical Review E* 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* 21.157-96; "Mathematical Appendix" available online.

Beckner, Clayton, Richard A. Blythe, Joan L. 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. *Language Learning* 59:Supplement 1.1-26.

Blythe, Richard A. and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88.269-304.

Boberg, Charles. 2004. Real and apparent time in language change: late adoption of changes in Montreal English. *American Speech* 79.250-69.

Bowie, David and Malcah Yaeger-Dror. To appear. Language change in real time. To appear in *The Oxford handbook of historical phonology*, ed. Patrick Honeybone and Joseph Salmons. Oxford: Oxford University Press.

Cedergren, Henriette. 1988. The spread of language change: verifying inferences of linguistic diffusion. *Language spread and language policy: issues, implications, and case studies (Georgetown University Roundtable on Languages and Linguistics 1987)*, ed. Peter H. Lowenberg, 45-60. Washington, D.C.: Georgetown University Press.

Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow, Essex: Longman.

Cukor-Avila, Patricia. 2002. *She say, she go, she be like:* verbs of quotation over time in African American Vernacular English. *American Speech* 77.3-31.

De Decker, Paul. 2006. A real-time investigation of social and phonetic changes in post-adolescence. *University of Pennsylvania Working Papers in Linguistics* 12/2.65-76.

Eckert, Penelope. 1997. Age as a sociolinguistic variable. *The handbook of sociolinguistics*, ed. Florian Coulmas, 151-67. Oxford: Blackwell.

Eckert, Penelope. 2000. *Linguistic variation as social practice*. Oxford: Blackwell.

Giles, Howard. 1973. Accent mobility: a model and some data. *Anthropological Linguistics* 15.87-105.

Giles, Howard & Philip M. Smith. 1979. Accommodation theory: optimal levels of convergence. *Language and social psychology*, ed. Howard Giles and Robert N. St Clair. Oxford: Basil Blackwell.

Gompertz, Benhamon. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115.513-583.

Harrington, Jonathan, Sallyanne Palethorpe and Catherine I. Watson. 2000. Does the Queen speak the Queen's English? *Nature* 408.927.

Hull, David L. 1988. *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press.

Hull, David L. 2001. *Science and selection: essays on biological evolution and the philosophy of science*. Cambridge: Cambridge University Press.

Kagan, Jerome. 1981. *The second year*. Cambridge, Mass.: Harvard University Press.

Kerswill, Paul and Ann Williams. 2000. Creating a New Town koine: children and language change in Milton Keynes. *Language in Society* 29.65-115.

Labov, William. 1994. *Principles of linguistic change, vol. 1: internal factors*. Oxford: Basil Blackwell.

Labov, William. 2001. *Principles of linguistic change: vol. 2: social factors*. Oxford: Blackwell.

Labov, William, Sharon Ash and Charles Boberg. 2006. *Atlas of North American English*. Berlin: Mouton de Gruyter.

Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: Wiley.

Milroy, James and Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21.339-84.

Milroy, Lesley and James Milroy. 1992. Social network and social class: toward in integrated sociolinguistic model. *Language in Society* 21.1-26.

Milroy, Lesley. 1987. *Language and social networks* (2nd ed.). Oxford: Basil Blackwell.

Nahkola, Kari and Marja Saanilahti. 2004. Mapping language changes in real time: a panel study on Finnish. *Language Variation and Change* 16.75-92.

Naro, Anthony Julius and Maria Marta Pereira Scherre. 2013. Remodeling the age variable: number concord in Brazilian Portuguese. *Language Variation and Change* 25.1-15.

Nevalainen, Terttu. 2000. Processes of supralocalisation and the rise of Standard English in the early Modern period. *Generative theory and corpus studies: a dialogue from 10 ICEHL*, ed. Ricardo Bermúdez-Otero, David Denison, Richard M. Hogg and C. B. McCully, 327-71. Berlin: Mouton de Gruyter.

Nevalainen, Terttu and Helena Raumolin-Brunberg. 2003. *Historical sociolinguistics: language change in Tudor and Stuart England*. London: Pearson.

Nevalainen, Terttu, Helena Raumolin-Brunberg and Heikki Mannila. 2011. The diffusion of language change in real time: progressive and conservative individuals and the time depth of change. *Language Variation and Change* 23.1-43.

Pope, Jennifer, Miriam Meyerhoff and D. Robert Ladd. 2007. Forty years of language change on Martha's Vineyard. *Language* 83.615-27.

Prichard, Hilary and Meredith Tamminga. 2012. The impact of higher education on Philadelphia vowels. *University of Pennsylvania Working Papers in Linguistics* 18/2.87-95.

Prince, Ellen F. 1987. Sarah Gorby, Yiddish folksinger: a case study of dialect shift. *International Journal of the Sociology of Language* 67.83-116.

Raumolin-Brunberg, Helena. 2005. The diffusion of subject YOU: a case study in historical sociolinguistics. *Language Variation and Change* 17.55-73.

Roberts, Julie. 1997a. Hitting a moving target: acquisition of sound change in progress by Philadelphia children. *Language Variation and Change* 9.249-66.

Roberts, Julie. 1997b. Acquisition of variable rules: a study of (-t,d) deletion in preschool children. *Journal of Child Language* 24.351-72.

Sankoff, Gillian and Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83.566-88.

Sankoff, Gillian, Hélène Blondeua and Anne Charity. 2001. Individual roles in a real-time change: Montreal (r->R) 1947-1995. *'r-ratics: sociolinguistic, phonetic and phonological characteristics of* /r/ (Special issue of *Études & Travaux - Insitut des langues vivantes et de phonétique,* 2001 no. 4), ed. Hans Van de Velde and Roeland van Hout, 141-57.

Sankoff, Gillian, Suzanne Evans Wagner and Laura Jensen. 2012. The long tail of language change: Québécois French futures in real time. *University of Pennsylvania Working Papers in Linguistics* 18/2,107-16.

Shockey, Linda. 1984. All in a flap: long-term accommodation in phonology. *International Journal of the Sociology of Language* 46.87-95.

Sundgren, Eva. 2009. The varying influence of social and linguistic factors on language stability and change: the case of Eskiltsuna. *Language Variation and Change* 21.97-133.

Tagliamonte, Sali A. and Alexandra D'Arcy. 2007. Frequency and variation in the community grammar: tracking a new change through the generations. *Language Variation and Change* 19.199-217.

Tagliamonte, Sali A. and Alexandra D'Arcy. 2009. Peaks beyond phonology: adolescence, incrementation and language change. *Language* 85.58-108.

Tillery, Jan and Guy Bailey. 2003. Approaches to real time in dialectology and sociolinguistics. *World Englishes* 22.351-65.

Trudgill, Peter. 1988. Norwich revisited: recent linguistic changes in an English urban dialect. *English World-Wide* 9.33-49.

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

Vihman, Marilyn May. 1985. Language differentiation by the bilingual infant. *Journal of Child Language* 12.297-304.

Wagner, Suzanne Evans. 2012a. Age grading in sociolinguistic theory. *Language and Linguistics Compass* 6/6:371-82.

Wagner, Suzanne Evans. 2012b. Real-time evidence for age grad(ing) in late adolescence. *Language Variation and Change* 24.179-202.

Wagner, Suzanne Evas and Gillian Sankoff. 2011. Age grading in the Montréal French inflected future. *Language Variation and Change* 23.275-313.

Weinreich, Uriel, William Labov, & Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. *Directions for historical linguistics*, ed. Winfrid P. Lehmann, and Yakov Malkiel, 95-195. Austin: University of Texas Press.

	Individual	Community
1 Stability	*stable*	*stable*
2 Age-grading	*unstable*	*stable*
3 Generational change	*stable*	*unstable*
4 Community change	*unstable*	*unstable*

Table 1. Relationships between individual and communal language behavior (adapted from Labov 1994:83, Table 4.1).

Source	Study	Ages	Times	Variety/Corpus	Variable(s)	Comments
Change in adult behavior across real time in a community change < 10% difference						
Nahkola & Saanilahti 2004, Tables 1, 2, 6, 8	PA%	3	2	Virrat Finnish	(PRON), (COP), (SCHWA1), (GEM)	Described as 'steadily advancing changes'
Nevalainen and Raumolin-Brunberg 2003, Tables 5.1, 5.2, 5.3	TRn	3-4	3	Corpus of Early English Correspondence	(ITS), (S/TH), relative adverb	1550-69 cohort advances >10% for (S/TH); other cohorts varying <10% for all variables
Sundgren 2009, Tables 7, 16	TR%	2	2	Eskiltsuna Swedish	(DefPlurNeut), (PastPart2)	
Labov 1994:91, reporting Fowler 1986	TR%gr	2	2	New York City (Saks, Macys)	(r)	
Cedergren 1988:54, Fig, 6	TR%gr	6	2	Panamanian Spanish	(CH)	1949 age cohort advances >10%; other 5 cohorts advance <10%
Boberg 2004:159-60, Figs. 4, 5, 6, 7	TR%gr	2	2	Montreal English	/sk-/ in *schedule, zee* for *zed,* /ɒ/ in *progress,* /ɛ/ in *lever*	Adults slightly advancing (but <10%) for /ɒ/ in *progress*, slightly retreating (but <10%) for /ɛ/ in *lever*
Labov 1994:102, Table 4.2	PIf	1	2	Philadelphia English	vowel system	changes in F1, F2 means of vowel system of Jenny Rosetti; 'no overall changes in base of articulation' (Labov 1994:102)
Change in adult behavior across real time in a community change >10% difference (or significantly changing for continuous variables)						
—All adults advancing						
Pope et al. 2007:622, Fig. 2	TRgr	3	2	Martha's Vineyard	centralization of (ay), (aw)	Changes of >0.5 in centralization index
Boberg 2004:261-63, Figs. 8, 9, 11	TR%gr	2	2	Montreal English	*bathe* for *bath* (tv.), *dove* for *dived,* loss of *chesterfield* for *sofa*	
Nahkola and Saanilahti 2004, Table 7	TR%	3	2	Virrat Finnish	(SCHWA2)	Change reverses direction, but age cohorts continue to move in the same direction as they had in previous time point (Nahkola and Saanilahti 2004:81)
Harrington et al. 2000:927, Fig. 1	PIgr	1	2	Queen Elizabeth II's English	8 vowel variables	Queen's English at two time periods compared to current standard southern-British English
—Older adults advancing, but less so than younger adults						
Narro and Scherre 2013:10, Table 5	TR%	3	2	Brazilian Portuguese	(NPC), (SVC)	For SVC, two older cohorts advance <10%

Boberg 2004:262, Fig. 10	TR%gr	2	2	Montreal English	*couch* for *sofa*	Older cohort advances <10%
—Adults retreating from community change						
Wagner and Sankoff 2011:302-3, Figs. 3, 4 (also Sankoff, Wagner and Jensen 2012:112, Figs. 3, 4)	PI%gr	41	2	Montréal French	future (in affirmative contexts; periphrastic future has been replacing inflectional future)	Most high-SPS speakers retreat >10%, most low-SPS speakers advance <10% or retreat usually <10%
Boberg 2004: 264, Figs. 11,12, 13	TR%gr	2	2	Montreal English	/u:/ in *student*, merger of *whine* and *wine*	Older cohort retreats <10% for /u:/ in *student*
—No clear age-related pattern of adult changes						
Sankoff and Blondeau 2007, Tables 11, 12	PI%	52	2	Montréal French	[R]	Some older speakers as well as younger speakers advance >10%; see §5 for further discussion and modeling
Changes in adult behavior reported in above studies for variation that is not clearly community change, usually <10% difference						
Nahkola and Saanilahti 2004, Tables 3, 4, 5, 11, 12	PA%	3	2	Virrat Finnish	(VSEQ), (siCOND), (siPAST), (3INF), (tk)	Described as 'unsteadily advancing changes'; some appear to be reversals of changes; all involve <10% change in adult behavior
Sundgren 2009, Tables 4, 11, 19, 23	TR%	2	2	Eskiltsuna Swedish	(DefSingNeut), (PastPart1&4), (Pret1), (Become)	Generally <10% difference in adult behavior

Study: TR trend study; PA panel study, data aggregated by age cohort; PI panel study, individual data; n numerical data, % percentage data, gr graphed data only, f change in format frequencies (continuous variable)
Ages: number of age cohorts whose behavior is reported across at least two different times (TR, PA); or number of individuals whose data is reported across at least two different times (PI)
Times: number of different time samples

Table 2. Changes in adult individual linguistic behavior during a community change

Model	**Mean C2-C1**	**St. Dev. C2-C1**	**% Adolescent Peak**	**Mean Peak Size**
Caregiver 20-35	0.051	0.079	70	0.092
Peer influence 1-15	-0.015	0.097	37	-
Caregiver 30-45	0.082	0.090	81	0.112
Random caregiver	0.103	0.079	89	0.121
Static network	0.037	0.088	65	0.088
Fully connected	0.078	0.072	88	0.095

Table 3. Statistics of adolescent peak for several model variations. From left to right: mean and standard deviation of C2-C1 difference, percentage of instances for which C2 was leading C1, and mean size of this lead. Data are averaged over apparent time curves taken at ten year intervals through 100 realizations.

(a) (b)

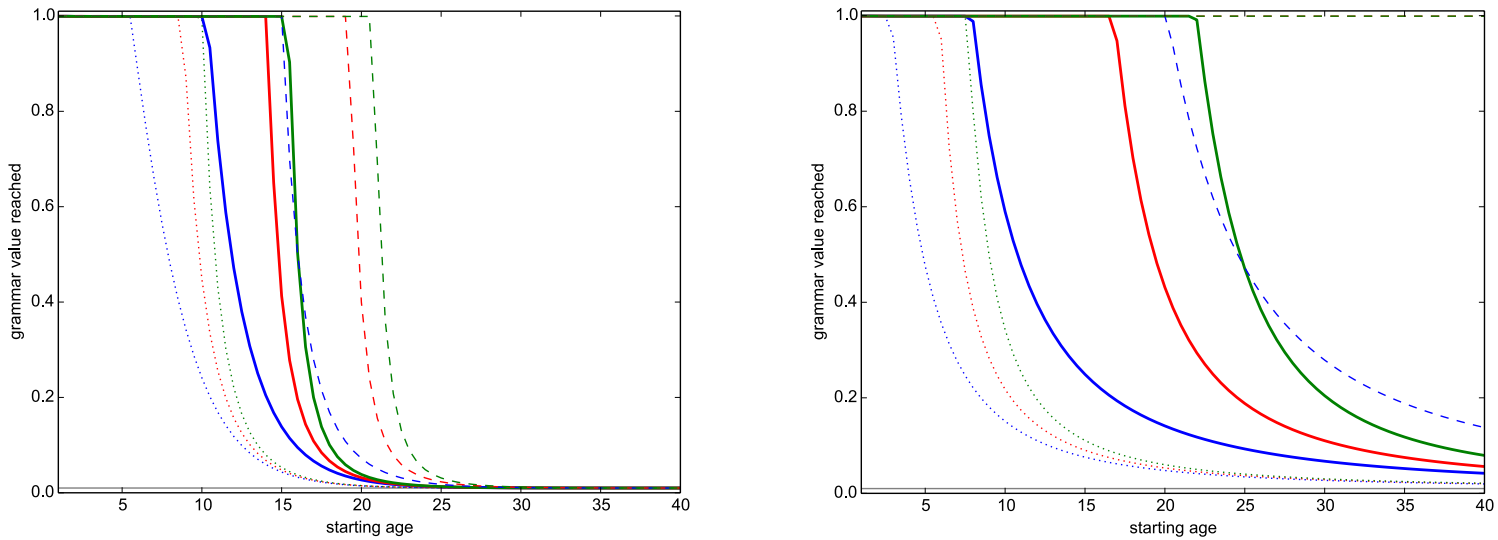


Figure 1

Final grammar value reached as a function of initial age, for a speaker of one variant immersed in a population of the new variant. (a) Speaker's λ value decays exponentially with parameter β = 0.4. Curves for three different replicator selection strengths, b = 0.001, 0.01, 0.02, (blue, red and green respectively) are shown. (b) Speaker's λ value decays as a powerlaw with parameter γ = 2.0. Curves

for the same three replicator selection strengths, b = 0.001, 0.01, 0.02, (blue, red and green respectively) are shown. In both plots, speaker use the given lingueme in 1000 interactions per year (solid lines), producing 10 tokens in each interaction. Curves for 100 interactions per year (dotted) and 10,000 per year (dashed) are also shown. Initial usage of 1% is marked by the gray line.

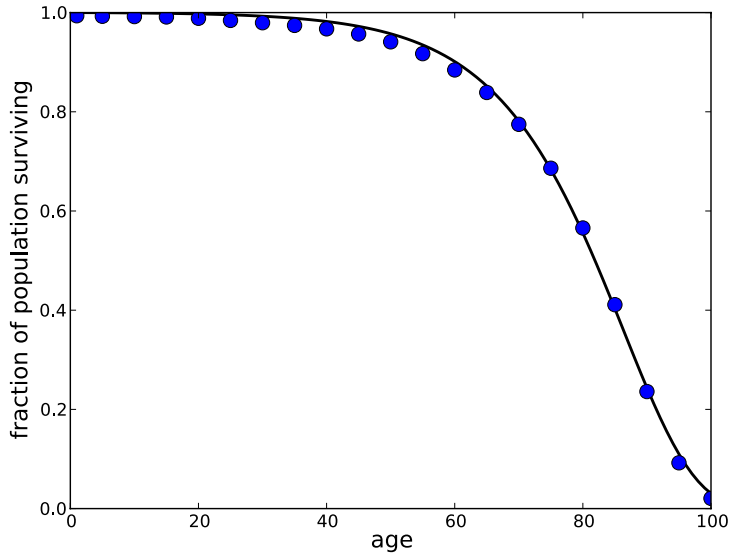


Figure 2

Probability of surviving to a given age. Circles represent 2009 United States mortality statistics (Arias 2014); the solid curve represents the age distribution used in simulations.

(a) (b)

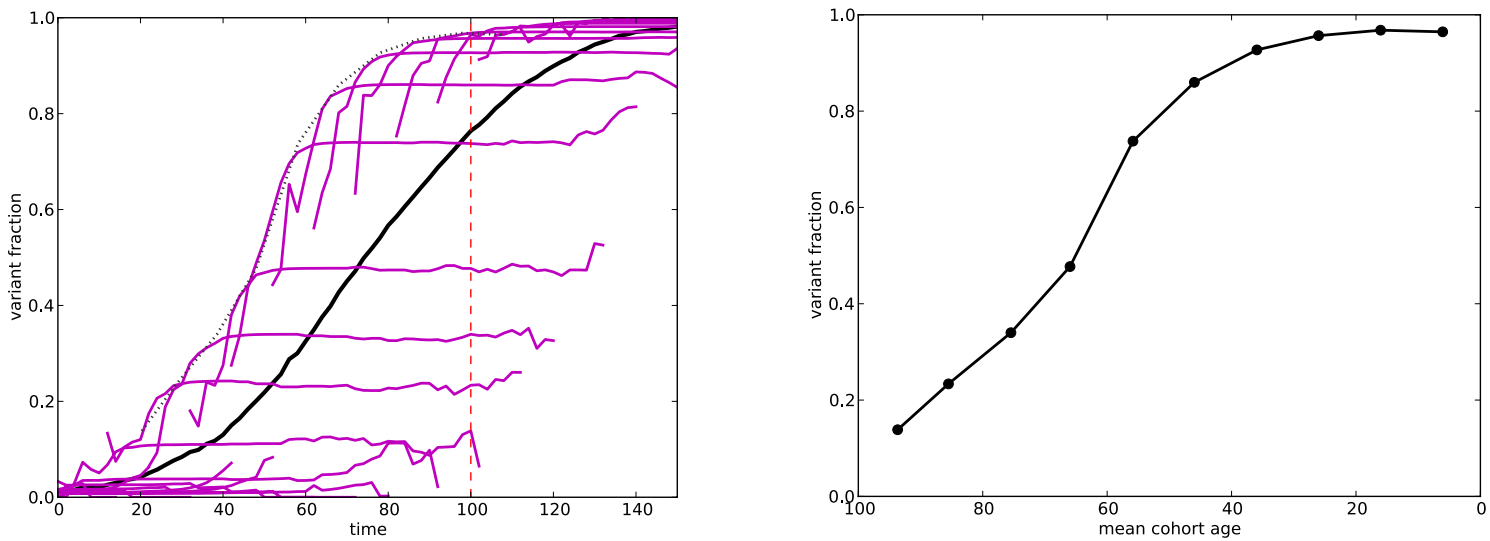


Figure 3

(a) Population mean (heavy black line) and cohort mean usage of the incoming variant over time for a typical simulation realization using the exponential λ decay function. (b) Cohort mean values for the same realization, plotted as a function of age at the time shown by the vertical red dashed line on the left.

(a) (b)

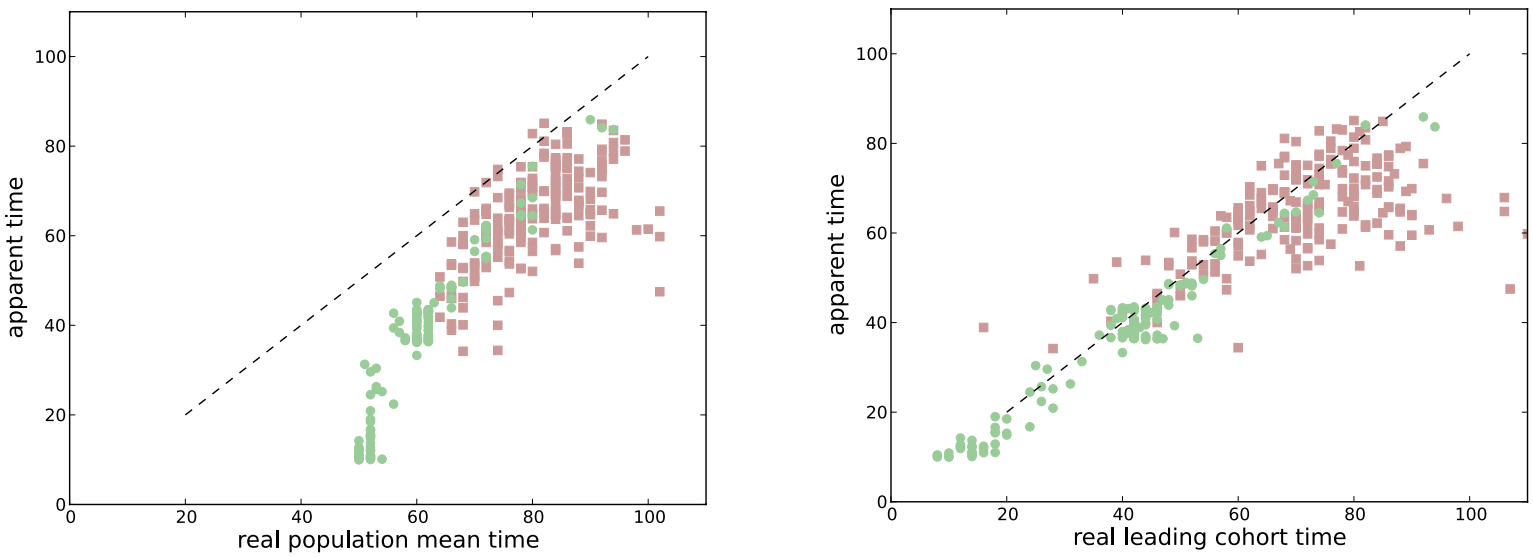


Figure 4

(a) *Apparent time vs real time intervals for the population mean grammar to change from 0.20 to 0.80, for multiple runs and various combinations of parameters b and H from 0.01 to 0.1 in a population of 1000 speakers. Circles represent a fully mixed population, while squares represent a sparse interaction network. The dashed line shows the line of equality, for reference. (b) Apparent time vs real time for the mean grammar of the leading cohort at each time.*

(a) (b)

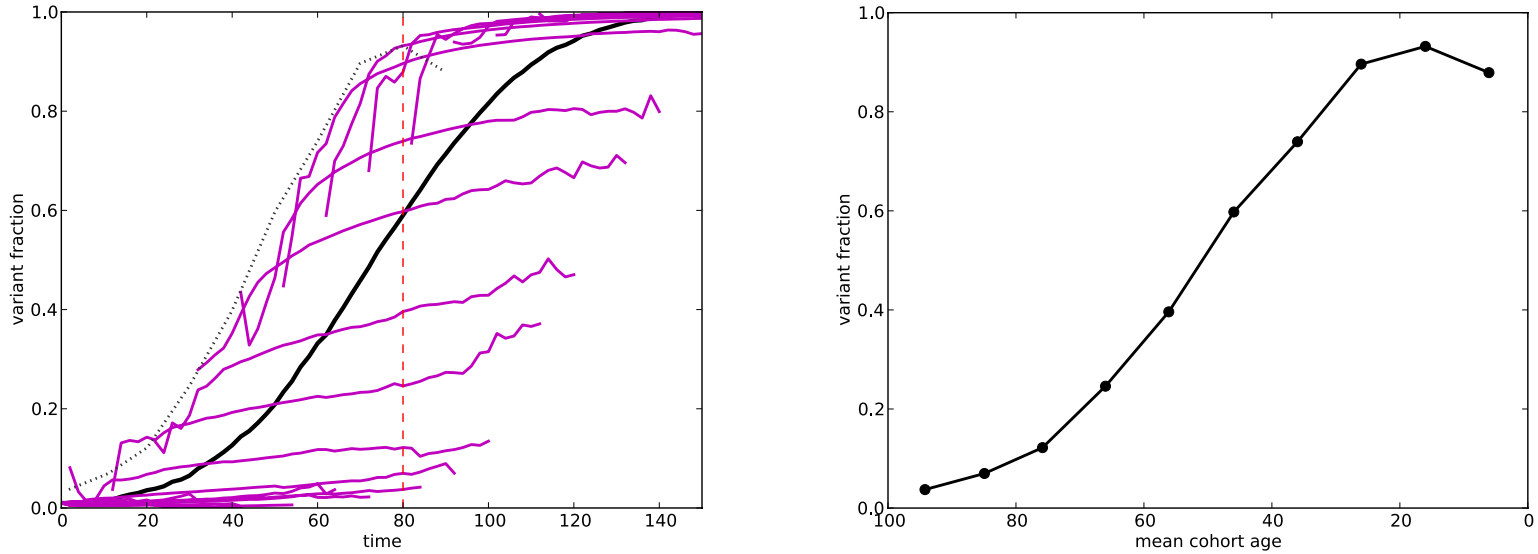


Figure 5

(a) Population mean (heavy black line) and cohort mean usage of the incoming variant over time for a typical smulation realization using the power law λ decay function. (b) Cohort mean values for the same realization, plotted as a function of age at the time shown by the vertical red dashed line on the left.

(a) (b)

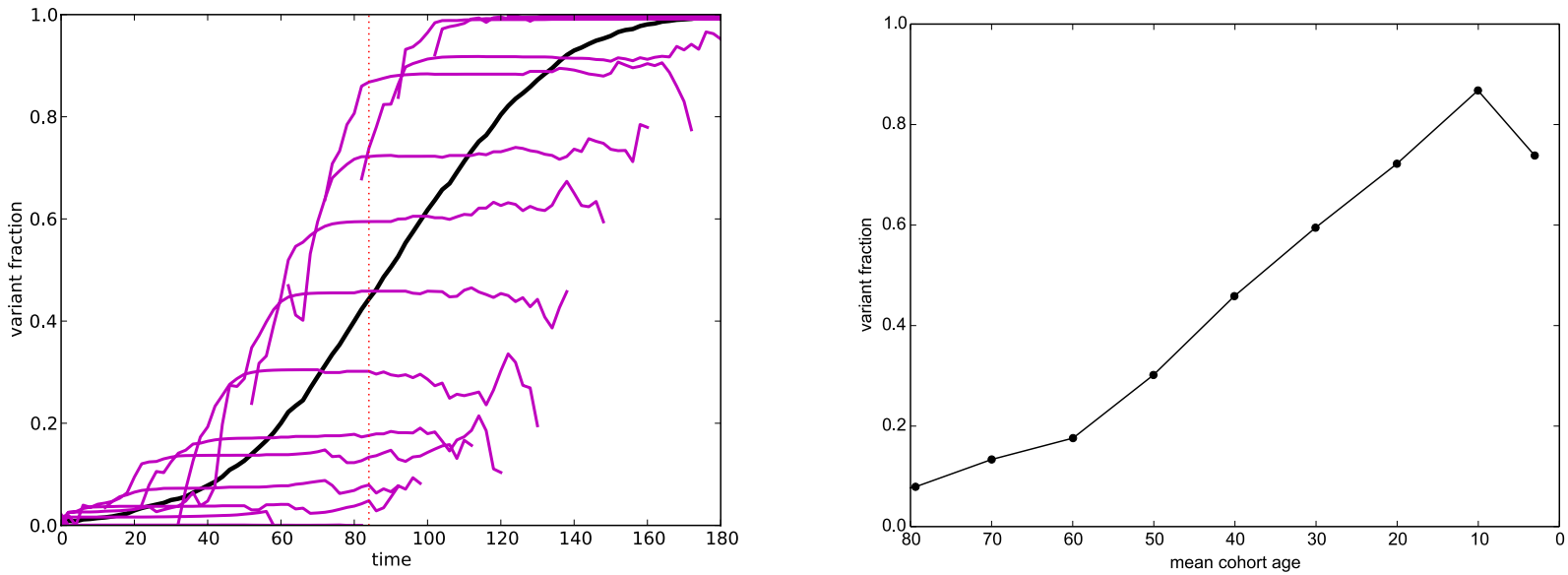


Figure 6

(a) Population mean (heavy black line) and cohort mean usage of the incoming variant over time for a typical simulation realization of the adolescent peak model. (b) The apparent time curve (cohort mean values for the same realization, plotted as a function of age at the time shown by the vertical red dashed line on the left) for the realization exhibits an adolescent peak.

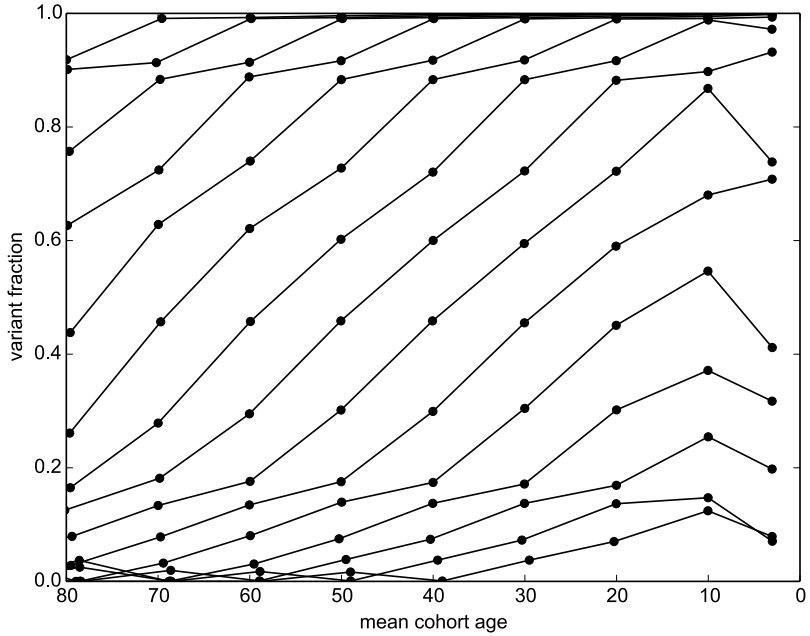


Figure 7

Apparent time curves taken at multiple time points through the same simulation run of the adolescent peak model. The adolescent peak is most prominent in the middle of the change, and disappears towards the end.

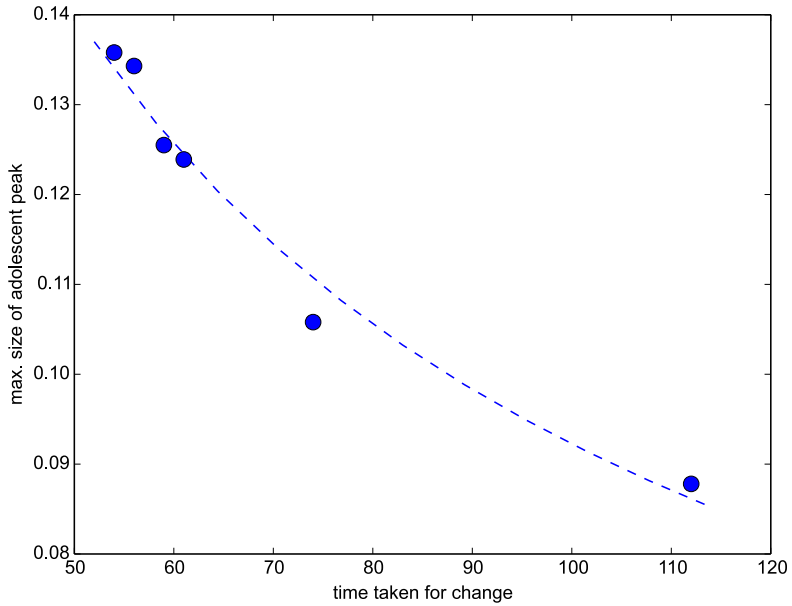


Figure 8

Size of the adolescent peak, calculated as the maximum difference between youngest two cohorts in each simulation, averaged over 10 realizations, as a function of the time taken for the community change.

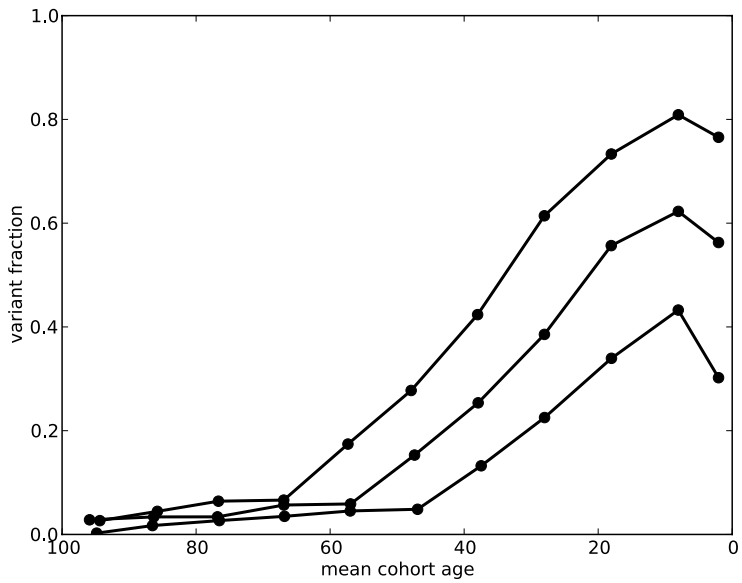


Figure 9

Distribution of differences in mean usage between the second youngest and youngest cohorts (C2-C1), for 100 realizations of the adolescent peak model using the same parameters (b=0.05, H=0.05). The mean value is marked with a vertical line. For comparison, the third to second (dashed) and fourth to third (dotted) difference distributions are also shown.

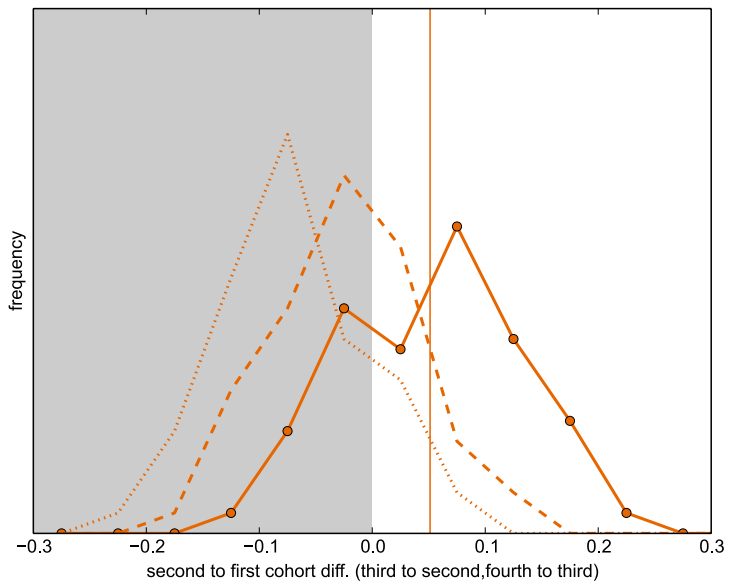


Figure 10

Apparent time curves for the adolescent peak model under the power law λ decay function, observed at three times during the same simulation realization. The adolescent peak also appears when using the power law λ decay function.

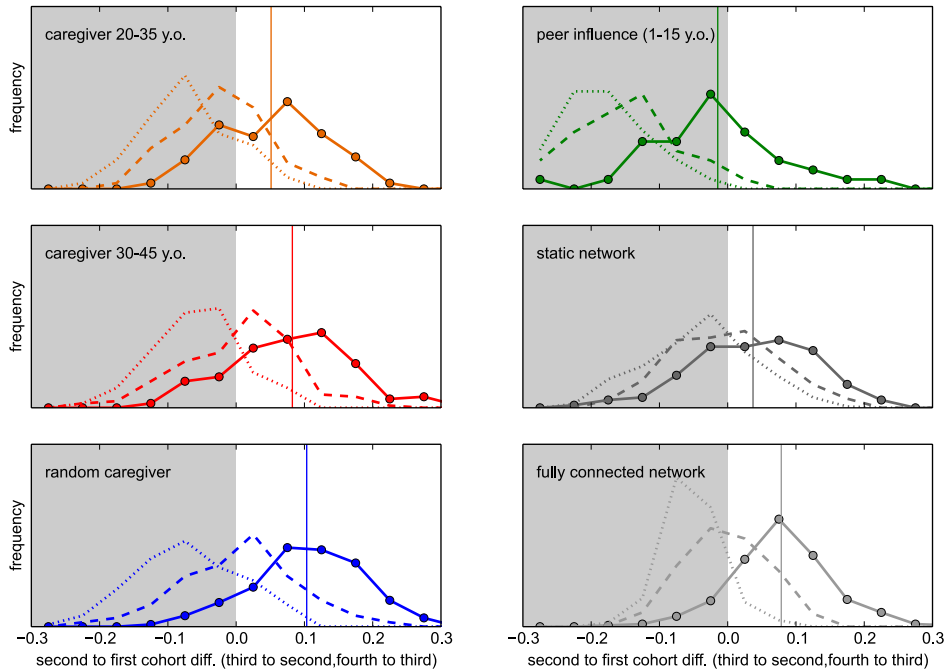


Figure 11

Distributions of differences C2-C1 for the main adolescent peak model (top left) compared with five other variations of the model. Middle left: main influence chosen to be 30-45 years old. Lower left: main influence chosen randomly from the entire population. Top right: main influence chosen to be 1-15 years old. Middle right: fixed network with the same structure as the adolescent peak model. Lower right: a fully connected network.

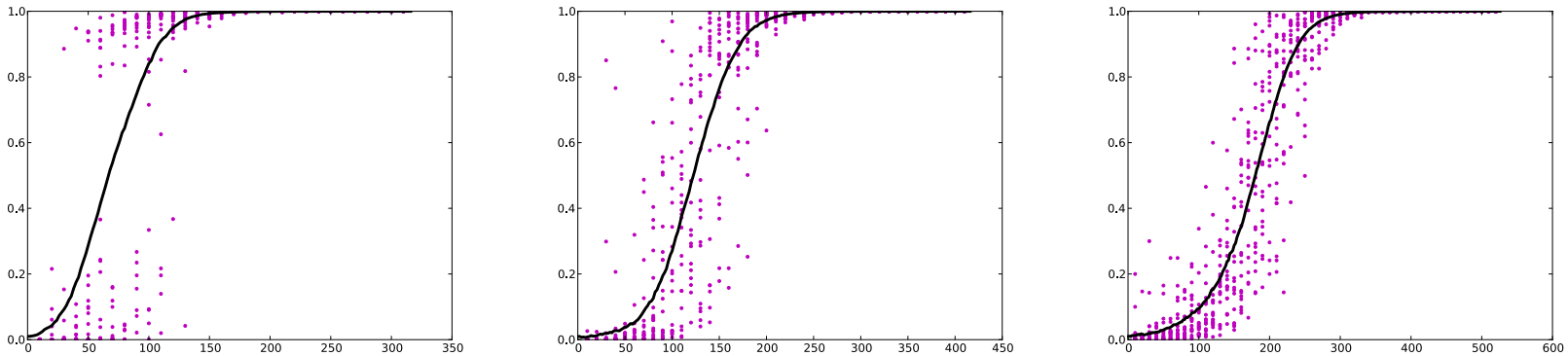


Figure 12

Change in individual "categoricalness" as accommodation increases. From left to right, H = 0.006, 0.014, 0.022, for fixed replicator selection strength b = 0.01. Solid lines give population mean grammar over time. Dots are grammars of 20 randomly sampled speakers from a population of 1000 (different sample at each time).

(a) (b)

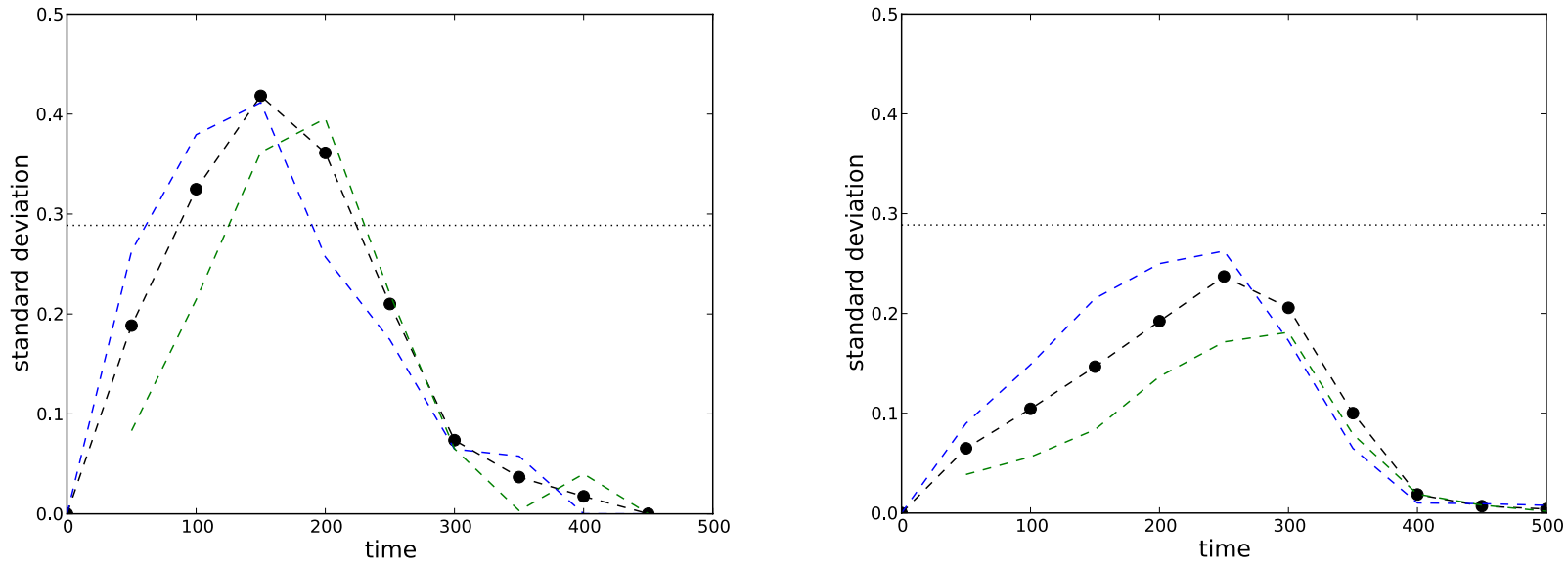


Figure 13

Standard deviation of population grammar values over time. (a) For a small value of the accommodation parameter H, resulting in a polarized pattern of change. (b) For a large value of H, resulting in a centralized pattern of change. Black line and circles are for the whole population. Blue line: speakers 20 years old or younger. Green line: speakers 40-60 years old.

(a) (b)

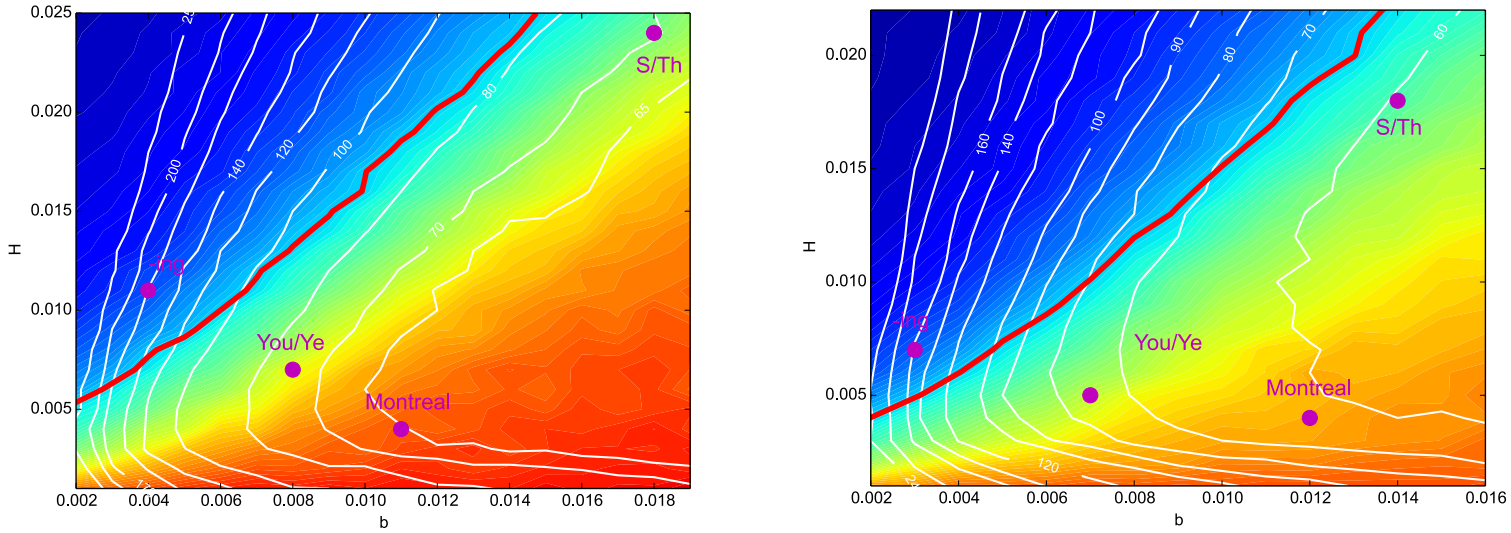


Figure 14

Summary of variation in patterns of change with the main parameters b and H, (a) using the exponential λ decay function, (b) using the powerlaw λ decay function. Color shading indicates population variance, blue smallest and orange largest values, corresponding to centralized and polarized patterns of change respectively. The red line marks the boundary value, variance = 1/12 (standard deviation = 0.29). White curves indicate lines of equal time of change, with value (in years) labeled. Pink circles mark the parameter values for simulations most closely resembling study data, as discussed in the text.

(a) (b)

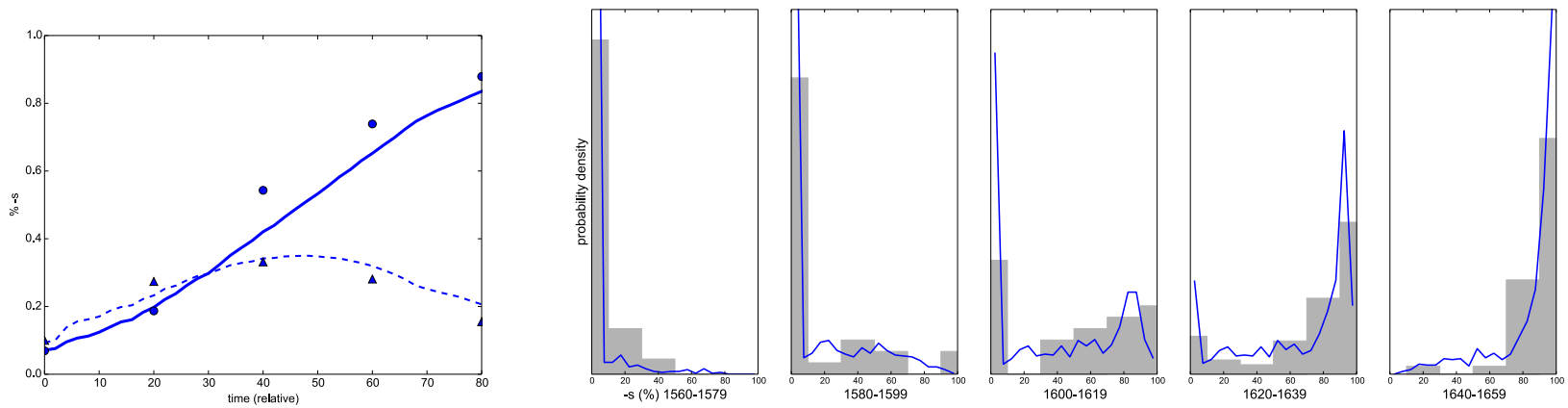


Figure 15

a) Time evolution of population usage of -s (circles) and standard deviation (triangles). Solid and dashed line are matching simulation values. b) Distribution of individual usage of -s in the population at several times through the change. Gray bars show density of speakers in the given interval, taken from data. Blue lines are the equivalent densities for matching simulation.

(a) (b)

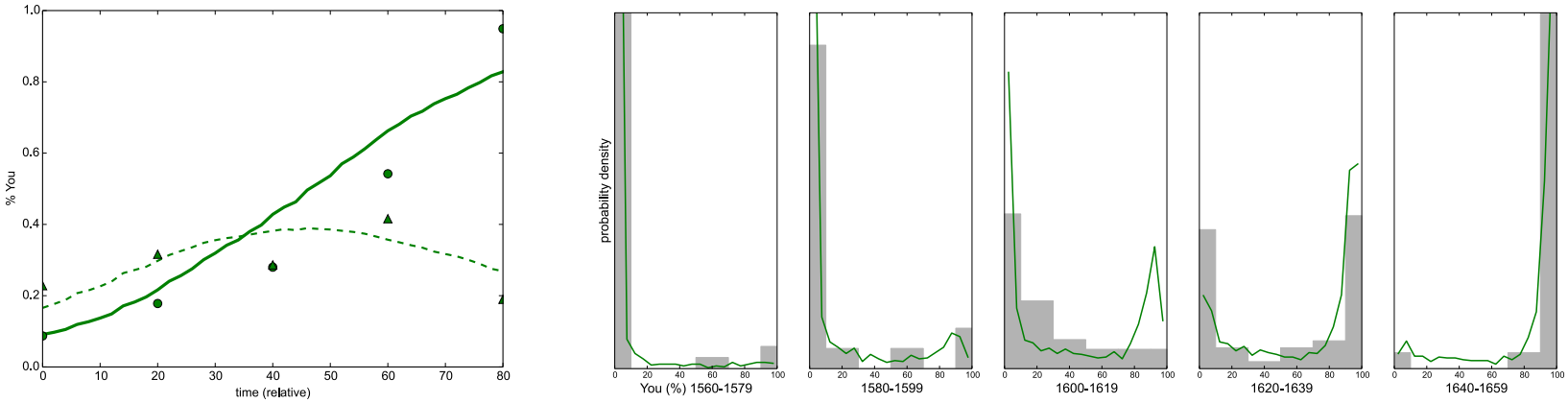
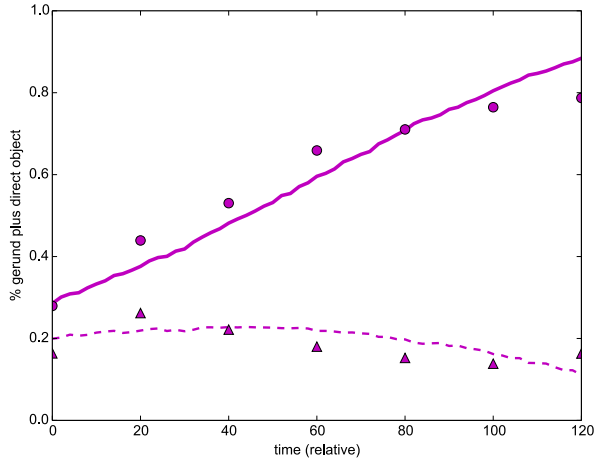


Figure 16

a) Time evolution of population usage of you/ye *(circles) and standard deviation (triangles). Solid and dashed line are matching simulation values. b) Distribution of individual usage of* you/ye *in the population at several times through the change. Gray bars show density of speakers in the given interval, taken from the data. Green lines are the equivalent densities for the matching simulation.*

(a)


% gerund plus direct object
time (relative)
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
120

(b)

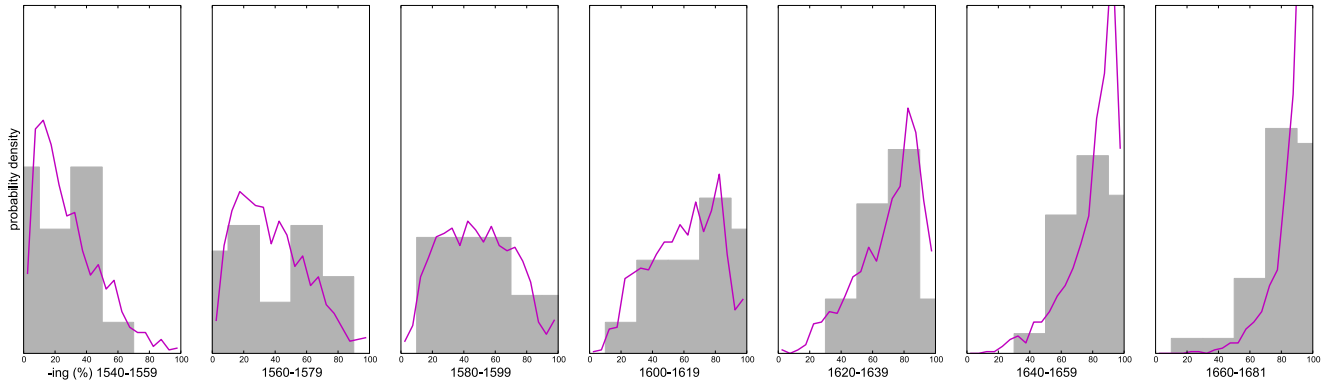

probability density
-ing (%) 1540-1559
1560-1579
1580-1599
1600-1619
1620-1639
1640-1659
1660-1681

Figure 17

a) Time evolution of population usage of gerund plus direct object (circles) and standard deviation (triangles). Solid and dashed lines are matching simulation values. b) Distribution of individual usage of gerund plus direct object in the population at several times through the change. Gray bars show density of speakers in the given interval, taken from data. Magenta lines are the equivalent densities for matching simulation.

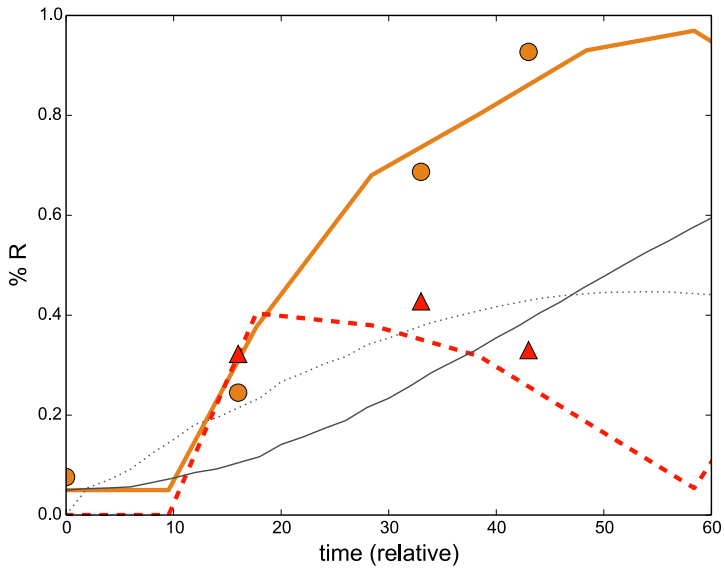


Figure 18

Simulation match to Montréal /r/ data. Circles are apparent time data from 1971, triangles are standard deviation values for the same. Solid orange line is the mean grammar value of the leading cohort for a simlation relization using the best matched parameters. Dashed line is standard deviation of leading cohort. For comparison, the real-time whole population curves (grammar and standard deviation) from simulation are also plotted (thin solid and dotted curves).