Agent-Based Modeling of the Evolution of Vowel Harmony

K. David Harrison, Mark Dras, Berk Kapicioglu[1]

University of Pennsylvania, Philadelphia

1. Overview

A central tenet of the study of microparametric variation (MPV) is that closely related languages will reveal which observable parameters of languages are correlated; in particular, this correlation can lead to the discovery of more abstract parameters underlying the behavior of surface patterns. The study of MPV is a way of approximating diachronic experimentation on a language, where the linguist (if it were possible) would alter a parameter of a language and see what occurred. In this paper, we propose agent-based simulation as a complementary methodology for examining the feasibility of abstracting parameters through analysis of related languages. Agent-based simulations are a way to model the sort of complex system entailed by the evolution of grammars within a speech community.

The focus of our research is vowel harmony in Turkic (Altaic) languages. These systems exhibit a great deal of change—instability, even—over the more than one millennium during which these languages have been recorded in writing. Most Turkic languages have two autonomous harmony systems, one based on tongue backness and one on lip rounding. The systems range from robust, nearly exceptionless harmony to highly variable or restricted harmony to no harmony at all. They thus provide rich material for constructing typologies of harmony and models of language change.

With our model, we attempt to show how changes at the level of the individual aggregate to give language-wide evolution in backness harmony systems. Agent-based simulation, where a community of computational agents interacts over time, is appropriate here, as language changes takes place in a social context; moreover, the changes occur over multiple generations of individuals.

[1] The authors wish to thank Mark Liberman, Donca Steriade, Kie Zuraw, Andy Wedel and two anonymous reviewers for their helpful insights. We also gratefully acknowledge the expertise of native speaker consultants of Tuvan, Uzbek, and Turkish.

Our aim in building a simulation model is to suggest a set of inputs for the model—the factors we hypothesize to drive change in harmony systems—and then compare the outputs of the model to the corresponding observations of the phenomenon. If the outputs are close to the empirical observations, the inputs are plausible factors, although there could of course be many possible other inputs that give the same outputs. Of equal interest is the case where the outputs are not at all like the empirical observations; here we can dismiss that combination of factors as a possible cause of the evolution of vowel harmony.

It is for this reason that we are interested in modeling the historical trajectory of vowel harmony evolution. Other work on simulation modeling looks almost exclusively at the binary question, Does the phenomenon emerge (or decay) at all? In this situation there are effectively only two outputs to compare with observations: the start point (phenomenon not present) and the end point (phenomenon either present or not). Having such a small set for comparison does not allow one to determine with much confidence what the factors causing the phenomenon are: many possible factors could cause the same behavior. Modeling change as a trajectory constrains the simulation to a much greater degree, allowing us to rule out many possibilities.

In our research on vowel harmony systems, we analyzed historical data from a dozen Turkic language corpora. Based on our interpretation of the data, we propose that evolution of vowel backness harmony systems can be plausibly hypothesized to follow an S-shaped curve of the type attested in other language change phenomena.

We have constructed a model that attempts to incorporate all significant factors that drive change in harmony systems. We start with factors internal to the language. These include asymmetries in production or perception due to co-articulation, patterns in the lexicon which may affect how speakers treat new words, and vowel inventory structure. While co-articulation and lexical patterns tend to favor harmony, inventory structure may either favor or disfavor it. In the case of Uzbek, which we'll be looking at more closely, inventory changes due to vowel merger contributed to the loss of vowel harmony. We also include in our model of harmony a number of other factors, both internal and external, that affect harmony: these include loanwords, morphologization, contact phenomena and consonant co-articulation. We attempt to represent all these factors in our model.

We present preliminary findings that suggest that micro-level variation in simulated agents who have only weak preferences or patterns in their grammar and lexicon can lead at the macro level and over long periods of time to emergence or decay of entire harmony systems. Further, this change approximates the S-shaped curve attested in historical data.

2. Turkic Vowel harmony

In most Turkic (and Altaic) languages, backness (i.e., palatal) harmony is apparent both as an ambient pattern of vowel co-occurrence within word roots, and as a productive pattern of vowel alternations (e.g., in suffixes). Harmony also determines the quality of epenthetic vowels and the vowels of nonce words, reduplicants and loanwords, and often does so in ways that, though systematic, may differ from the patterns attested in root and suffix vowels in the language. A typical Turkic vowel inventory includes four front and four back vowels, neatly divisible into harmonic classes.

(1)

	front	*back*
high	i y	ɯ u
non-high	e ø	ɑ o

There are a number of hypotheses about how palatal harmony may have emerged. One school of thought holds that harmony arises from co-articulation (Boyce 1988, Ohala 1994, cf. Inkelas et al 2001). Another is the structuralist notion that symmetry in vowel inventories provides an impetus to harmony (cf. Trubetzkoy 1969). Thirdly, there is the view that harmony may cue word boundaries (Trubetzkoy 1969, Suomi 1983, Vroomen, Tuomainen, and de Gelder 1998) or aid the perception of difficult vowel contrasts (Kaun 1995). Our current model, which we describe in greater detail below, includes the factors of co-articulation, inventory structure, and also the possibility of misperception due to various factors.

Historically, Turkic vowel harmony systems are constantly in flux. Old Turkic as attested in 8^{th}-11^{th} century runic inscriptions from Siberia had an eight vowel system and fully regular backness harmony (Kondrat'ev 1981). Modern Turkic languages have from 5 to 10 vowels, and range from almost fully harmonic (Tuvan) to not harmonic at all (Uzbek). The Turkic family thus provides our model with over one millennium of documented stages and scenarios in the evolution of harmony. These serve as data points showing harmony evolution along a definable trajectory.

But not all points on this trajectory are discernable in the historical record. There are stages we know must have taken place that were not recorded. For example, 8^{th} century Old Turkic shows pervasive, almost exceptionless vowel harmony for backness. Prior stages in the emergence of this system were not documented, and the gap limits our empirical knowledge of how such systems originated. By contrast, the evolution of harmony systems in daughter languages of Old Turkic is quite well documented across a period of over 1,000 years, allowing us to precisely quantify stable or declining levels of harmony over time, up to the present day.

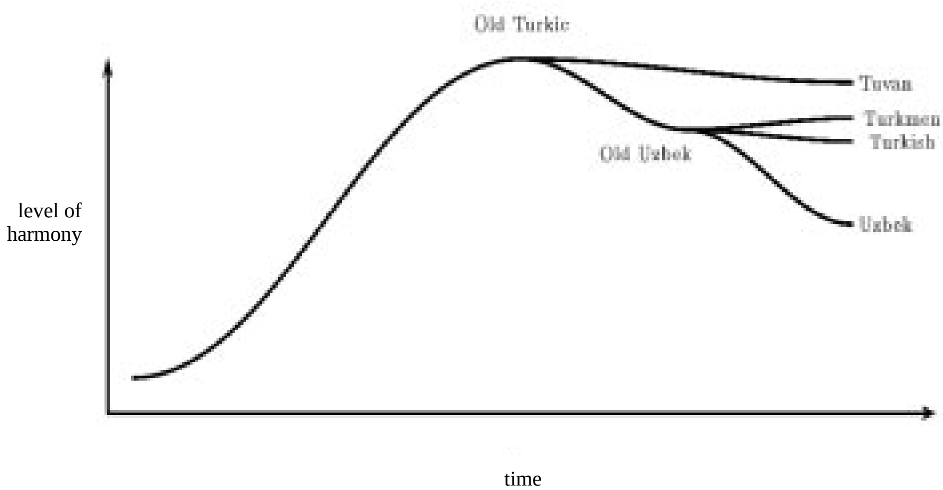
To quantify harmony evolution, we applied search algorithms to a dozen Turkic language corpora ranging in size from 1,000 to 40,000 words and in age from 1,000 years to contemporary. These algorithms allowed us to precisely tabulate the relative frequency of vowels (and, by extension their relative markedness), the pervasiveness of harmony as a pattern in roots and suffixes, and the presence of disharmony in the lexicon.

Our findings show the following levels of backness harmony for nine Turkic languages.[2]

(2)

Corpus	Level of harmony	Time period
Old Turkic	100.0 %	8th-9th century
Old Anatolian Turkish	90.4 %	13th century
Ottoman Turkish	81.0 %	17th century
Old Uzbek	77.8 %	17th century
Armeno-Kipchak	72.4 %	17th century
Tuvan (various dialects)	96.0 % to 99.0%	contemporary
Turkmen	83.0 %	contemporary
Turkish	75.0 %	contemporary
Uzbek	53.8 %	contemporary

(3) Schematic evolution of Turkic backness harmony systems[3]



Note from table (2) that the data overall follows a slow-fast-slow pattern of change over the given timespan, splitting it into three equal intervals (8th to 13th centuries, 13th to 17th, and 17th to present): harmony declined by 9.6% in the first interval, by 13.3% in the second interval (taking the average of the 17th century values), and by 6.5% in the third interval (taking the average of the contemporary values, excluding Tuvan as it is not directly descended from the various 17th century Turkish languages).

[2] Old Turkic (Fazylov 1972, Kondrat'ev 1981), Old Anatolian Turkish (Fomkin 1994), Ottoman Turkish (Gilson 1987), Old Uzbek (Batmanova 1971), Armeno-Kipchak (Schütz 1968), Tuvan (Anderson & Harrison 2000), Turkmen (Mämmedow 2001, Lastowka 1996), Turkish (Inkelas et al. 1997), Uzbek (Akobirov 1981, Ismatulla 1995, Dirks 2001).

[3] The schematic is not intended as a precise tree model of historical linguistic change. While all the languages shown are related, the schematic itself shows the comparative development of harmony systems across the Turkic family, not the finer details of interrelationships among Turkic languages.

3. Loanwords as diagnostics of change in harmony systems

Harmony evolution can be observed in a much compressed time-scale in some of these languages. The introduction of loanwords into a language can serve as a catalyst for harmony breakdown over a relatively short timeframe. Further, the degree to which loanwords are mutated to be harmonic provides a possible diagnostic for the state of the harmony system. The tendency to mutate loanwords depends partly on the social dynamic and evolves over time, sometimes more rapidly than the rest of the harmony system. In Tuvan, vowels (underlined below) in loanwords from an earlier period of language contact (18th-19th cent.) were uniformly mutated to obey backness harmony.

(4)	Tuvan word	source word	
	tʃurumal	← *ʒirumal*	‘pattern’ (Mongolian)
	hapajaq	← *kaˈpʲejek*	‘kopeck’ (Russian)
	maʃɯna	← *maˈʃina*	‘automobile’ (R)
	hine:k(e)	← *ˈkniga*	‘book’ (R)

By contrast, in a later period of language contact (20th century), vowels of loanwords in Tuvan are no longer mutated, but remain disharmonic:

(5)	*qambeːt*	← *kanˈfʲeta*	‘candy’ (R)
	maʃina	← *maˈʃina*	‘automobile’ (R)
	kniːga	← *ˈkniga*	‘book’ (R)
	kikpoqs	← *kikboks*	‘kick-boxing’ (R)

The data show evolution over a short period of time of one sub-system of harmony—its application to loanwords. What we see happening on a small scale in the sub-lexicon of Tuvan loanwords over the last three centuries mirrors the 1,000 year history of harmony evolution in Turkic. As there, the general trend is towards disharmony.

In Old Uzbek, of the 17th century, we see a similar tendency to mutate loanwords, but one that was already considerably weakened as the harmony system began to erode. At that time, the Uzbeks were absorbing a large population of Persian speakers, and Arabic and Persian loanwords were a prolific source of disharmony in the language. In Old Uzbek (Fazylov 1972) some Persian/Arabic disharmonic loanwords had been mutated to be fully harmonic, e.g., *rahmet* > *rahmat* ‘thank you’. However, the overall trend by this stage of the language was to allow disharmonic loanwords—of which there was a growing proportion—to remain disharmonic, e.g., *imam* ‘imam’, *fakiir* ‘fakir’.

But would the influx of disharmonic loanwords alone have been enough to destroy the harmony system? Probably not, since Old Uzbek still maintained a higher level of harmony (77.8%) than does modern Turkish, which displays a robust harmony pattern despite having a lexicon that is only 75% harmonic. Another factor, vowel merger, must also be considered.

4. Vowel markedness and vowel mergers

The Old Uzbek corpus shows significantly lower frequencies for the more highly marked vowels.

(6) Old Uzbek vowel frequency[4]

	Phoneme	*Frequency*	
front	i	16.2%	
	y	4.2 %	
	e	16.6 %	
	ø	9.6 %	
back	**ɯ**	**2.5 %**	← most marked
	u	12.7 %	
	ɑ	32.4 %	
	o	8.4 %	

The three vowels [y] [ø] [ɯ] above were soon to undergo merger and disappear, leaving modern Uzbek with just five vowels. These three vowels had a combined frequency of 16.3% in Old Uzbek. Was the disappearance of three of the original eight vowels enough to cause Uzbek harmony to disappear? Disharmonic words arising from the vowel merger plus disharmonic loanwords together yield a 38.5% disharmonic lexicon. This would have brought post-vowel merger Uzbek down to a 61.5% harmony level, a rather weak pattern that may fall below the threshold of a possible harmony system. Modern Uzbek, as noted above, is now only 53.8 % harmonic. It entirely lacks harmony as either an ambient pattern in roots or in suffix vowel alternations (e.g., plural suffix *-lar* ~ *-ler*) characteristic of Turkic harmony languages. We address in our simulation the question of what constitutes the minimal threshold for a harmony pattern.

To sum up, we can imagine a scenario in which the Persian population was responsible for the loss of Uzbek harmony, but was it their loanwords or their vowel merger that did it? Because these are intertwined, we cannot answer this question based on empirical facts alone. No Altaic language with declining harmony, as far as we know, exhibits only one of several possible harmony-eroding factors to the exclusion of all others. An agent-based simulation allows us to address this dynamic interaction by constructing many possible scenarios with differently weighted factors.

5. Trajectory of change

In building our simulation we adopted the S-curve as a logistic for evolutionary change. Several independent lines of research suggest that language change often proceeds along an S-shaped curve. The rising curve shows the advancement of a new form at the expense of an old one. On the curve, change begins slowly, accelerates, then slows again, over a period of many generations. It was originally proposed in Bailey (1973), as part of a "wave" model of linguistic change, with support coming from parallel behavior in population biology in the replacement of genetic alleles. As empirical support, Chen and Yang (1972) look at three case studies of historical data that demonstrate the S-shaped

[4] Frequency figures include monosyllables, which are excluded from calculations of harmony.

behavior of language change: the Chaozhou dialect of Chinese, where words have been shifting from one tone class to another, with the slow-fast-slow pattern in evidence; English diatones, for example, noun-verb pairs where nouns are moving to the stress pattern of accented first syllable (e.g. noun ˈ*addict* vs. verb *ad*ˈ*dict*); and the Swedish optional final *–d*, where the number of words that allow dropping of the final *–d* in ordinary Stockholm speech is decreasing in an S-shaped trajectory.

The earliest work in attempting to model language change, that of Kroch (1989) on the transition of Old English and Old French away from verb-second syntax, thus adopts the S-shaped curve, as has subsequent work.

In the case of Turkic harmony *emergence*, we are assuming an S-curve trajectory in the absence of historical data points. In the case of harmony *breakdown*, we also adopt an S-shaped curve, but we are guided here by a number of historical data points along the trajectory. We are not claiming that the S-curve is necessarily the right curve, merely that it is a plausible one for this type of change. (Though we plan to do further work in applied mathematics to apply curve-fitting to the results.) For now, the S-curve is more or less in accord with the evolution we have been able to map out for Turkic harmony.

6. Factors favoring harmony

The model of harmony evolution we adopt recognizes that change is driven by both internal and external factors. Internal factors include vowel shifts and mergers, markedness effects, consonant-induced disharmony, and assimilation of loanwords. External factors include language contact, bilingualism, language shift, and, once again, loanwords. Clearly there is some fuzziness in the boundary between internal and external factors, but we tried to focus first on the clearly internal factors.

We model the following internal factors favoring the emergence of vowel harmony harmony: (i) vowel co-articulation, (ii) inventory structure, (iii) patterns in the lexicon.

Inventory structure has at least two aspects relevant to harmony. The first aspect is inventory size: in a symmetrical 8 vowel system a uniform probabilistic distribution of 8 vowels yields a baseline 50% probability of harmony in disyllabic words. This favors harmony, albeit only in a minimal way. We note that an odd-numbered, seven-vowel inventory is slightly *more* favorable (at 51% probability) to harmony, given a uniform probabilistic distribution of vowels in disyllabic words. The second aspect of inventory structure relevant to harmony is symmetry. You still have an overall better chance of harmony if you have a symmetrical inventory, because you need paired vowels to get harmonic alternations. So, a system with unpaired vowels is not as conducive to harmony as a symmetrical inventory where all vowels are well-paired and amenable to alternation.

The third factor, distributional patterns in the lexicon, accounts for the treatment of new forms and may also influence the direction of errors in pronunciation or perception. We consider two types of patterns: (i) frequency of vowels (=markedness), and (ii) co-occurrence of vowels (=harmony). We assume a harmony pattern must robustly exceed the probabilistic minimum (the 50% threshold) to be noticed at all. There are also some minor factors favoring harmony not yet included in our model. These are: (i) harmony increases predictability in vowel sequences (this might facilitate cognitive processing), and (ii) harmony allows for economy (underspecification) in underlying representations. To sum up, three main factors: co-articulation, inventory structure, and patterns in the lexicon each contribute to the general probability that words will be harmonic, and in the case of new words—if they are disharmonic—the probability that they will be mutated to become harmonic.

We note that vowel mutation in loanwords as we showed in Tuvan and Uzbek can have multiple causes: for example, in the case of di-syllabic and longer loanwords, mutation of vowels seems to be driven mainly by co-articulation and harmony. But in the Old Uzbek corpus, we also found that 14% of all *monosyllabic* words had a mutated vowel form. This type of mutation cannot plausibly be driven by vowel co-articulation nor by harmony patterns. Clearly, markedness, misperception and other factors—for example—interference by consonants (e.g., palatal glides can cause adjacent vowels to be fronted)—must also contribute to the overall probability of mutation. In our simulation, *probability of mutation* is kept distinct from *probability of harmonizing*. Once the decision is made to mutate a word or not, then harmonic forces can take over and influence the direction of mutation.

7. Factors disfavoring harmony

Our model includes several internal factors disfavoring harmony. These are: (i) vowel merger (the Uzbek loss of [y] [ø] and [ɯ]), (ii) morphologization (the presence of fixed, non-harmonic suffixes such as the Uzbek possessive), (iii) pronunciation errors (low probability), and (iv) disharmonic lexemes (arising from unmutated loanwords). An additional factor dis-favoring harmony but not included in model is assimilation of vowels by consonants which can override vowel harmony.

At present only one external factor disfavoring harmony is included in our model. Our simulated Uzbek speech community includes a sub-population of Persians who lack the three marked vowels. They thus drive the vowel merger that gives rise to disharmony, as they can only mutate words from the more marked to the less marked vowels, but not in the reverse direction. Other external factors disfavoring harmony—including monolingual and bi-lingual sub-populations to better simulate the language contact environment—will be included in the model at a later stage. Language contact and bilingualism typically disfavor harmony, unless the contact language is also harmonic. Loanwords, although they originate in a donor language, are for the most part treated as an internal factor in our model. The initial conduit for loans in early contact situations is usually a minority of the population that is bilingual. The rest of the population must essentially treat these words as native, with the possibility of mutating them (or not) to conform to native phonology.

We now have in place a basic model of harmony. Where the uncertainty arises and the need for agent-based simulation begins is in determining what particular combination of factors in the model can lead to harmony breakdown or emergence. How can we assign relative weightedness to factors that favor or disfavor harmony?

8. Agent-based Simulation

Modeling diachrony in a mathematical or computational framework is useful for investigating the process of and the consequences of hypotheses about parametric variation, particularly for those changes that might occur over multiple generations in a community. However, there is only a small body of work on building mathematically-based models to evaluate language variation. One type (e.g., Clark and Roberts 1993, Briscoe 1999) models the co-evolution of a language acquisition device and the syntax of a language: the “language agent” has parameters that change according to agent interactions and the resulting “fitness” of the agents. The other type does not presuppose teleological agents that have a particular fitness goal; changes arise from community interaction. Early work in modeling language change in this way (e.g., Kroch 1989) imposed a particular S-shaped trajectory on the data. Kroch proposes:

> ... given the mathematical simplicity and widespread use of the logistic [a particular equation giving an S-curve], its use in the study of language change seems justified, even though, unlike in the population genetic case, no mechanism of change has yet been proposed from which the logistic form can be deduced. (Kroch 1989: 204)

Later work, including our own, is interested in how the S-shape observed in data can emerge from simple parameter interaction. The first of these are macro models that model the behavior of the whole speech community through mathematical recurrence relations (Niyogi and Berwick 1997). So, for example, the proportion of the population that is using the new variant at some time n, p_n, is a function of the population at time $n\text{-}1$ (i.e., p_{n-1}), of a form like the following:

(7) $p_n = A * p_{n-1}^{\;2} + B * p_{n-1} + C$

where A, B and C are coefficients determined by a model of language acquisition. However, these models have fundamental problems because they treat populations in the aggregate, and moreover non-stochastically (Briscoe 2000).

Incorporating stochastic behavior, and subdividing the aggregated population, leads logically to a computational agent-based simulation, as the mathematics otherwise becomes intractable. We therefore used the SWARM simulation platform (www.swarm.org), to build a simulated speech community with a mixed ‘Uzbek’ and ‘Persian’ population, composed of individual agents that differ individually in their lexicons, grammars (phonologies), and speech behavior. We used this community to enact scenarios for the emergence and/or decay of harmony systems over long periods of time (typically, a period of over one millennium).

A general premise of agent-based simulations is that you can use them to tease apart diverse factors that condition change in a synergistic fashion. As noted, we never get to observe in isolation the various factors affecting change in harmony. The collective interaction of these factors presents a computationally intractable problem. A second important premise of simulations is the possibility of emergent structures (Axtell and Epstein 1996, Satterfield 2000). Macro structures not programmed locally into simple, individual agents may emerge globally at the level of the speech community. We endow agents in the simulation with certain low-level preferences and behaviors, but no explicit knowledge of what kind of structure is supposed to emerge or along what path of development. Then we turn the agents loose and allow them to interact, and see what, if any, structures emerge over time.

A third premise is that agents' behaviors evolve in a context created by their collective interactions (Liberman 2001). Social interaction is a crucial dimension of the simulation. Agents provide feedback to each other, and the distribution of forms both in the individual agent and across the community ultimately depends on other agents' learning. In our simulation, agents' word-mutation algorithms are constantly changing based on what they hear in conversation with their neighbors and how they evaluate new words in light of their own internal lexicon. A fourth premise we adopt is to model the path of change, and not just the endpoints, to mirror the historical facts. Finally, building simulations allows us to address various questions about what kinds of language change can or cannot be usefully simulated.

An important principle in building our simulation—in choosing which factors to include in the model, and in choosing how to realize them—is, following Occam's Razor, to start with as simple a model as possible. If this fails to model the data accurately, the model is made incrementally more complex until (hopefully) the model properly fits the data. If we start with a complex model, it isn't possible to tell which factors are crucial for the outcome.

Our simulation contains agents with randomly varying lifespans and reproductivity. Agents in our simulation do the following:

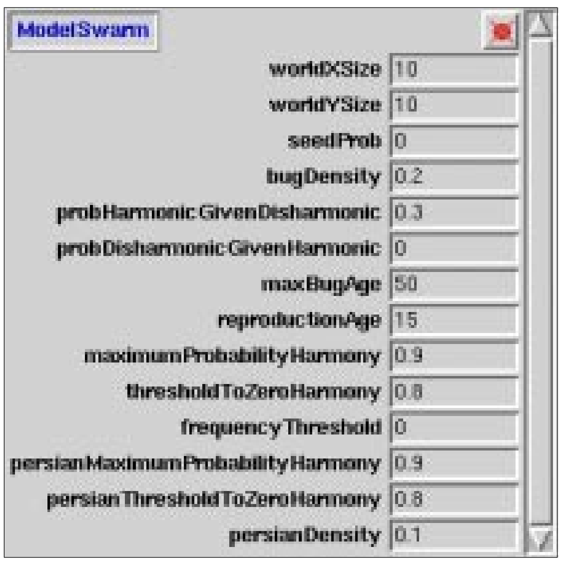
- maintain a lexicon
- strengthen lexical entries based on frequency of use
- learn new words from talking with neighbors
- assume others' grammars to resemble their own
- can mispronounce or mishear a word
- can mutate a word, changing its lexical entry
- co-articulate (more likely to err in favor of harmony)
- can ignore or attend to the fact of co-articulation by others
- evaluate the ambient harmony pattern in their own lexicon

Each of these can be varied with each new run of the simulation. In a typical run, which lasts about 1,500 "years", agents have one conversation a day in which they exchange a word with a randomly selected neighbor. Agents then evaluate the word, strengthening the lexical entry if it already exists. If it is new, they have a random probability of adopting it as is or mutating it. If they elect to mutate it, they assess their own internal lexicon and determine the probability that they will make it harmonic.

The tendency to co-articulate is represented in the MaxProbHarm parameter which also subsumes several other lesser factors favoring harmony (yielding a different granularity of parameters).

The parameters we adjust in the simulation are as follows. First, we vary the INPUT LEXICON of disyllabic words, which is based on a real Turkish lexicon. It's currently 50% harmonic—for disyllabic words, if front and back vowels are (in the aggregate) equally likely, this is the mean level of harmony that would occur just by chance—but we can adjust the exact level of harmony as a parameter. The relevant parameters are shown with arrows in the diagram below, a window from the actual SWARM simulation.

(8)



(not used in current version)
(not used in current version)

← factors favoring harmony
← pattern recognition
← relative strength of lexical entries

The next important parameters relate to the tendency of a word to mutate towards or away from harmony, because of coarticulation, misperception, and so on. Our probabilities of word change are conditioned on the harmony of the word; that is, we have different (conditional) probabilities depending on whether the word is harmonic. To start with, we choose the simplest case of uniform probabilities for all agents, values which are set as parameters. (We modify this later, in 9.2.)

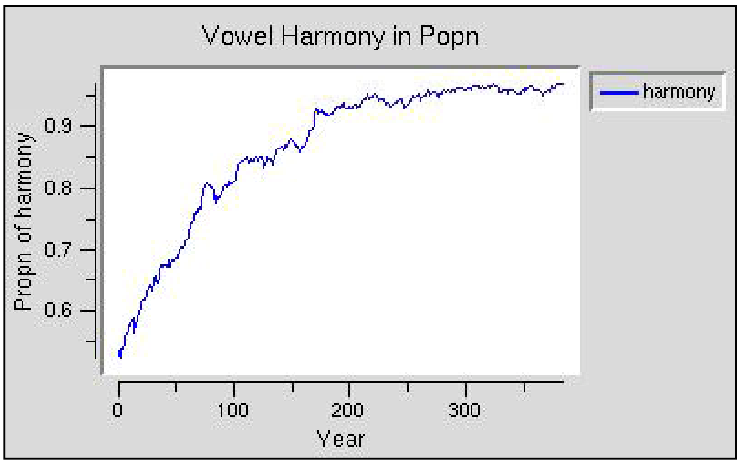
This is so far very simple. We could, in principle, impose a more complex formula, for example, one that incorporates Bayesian probabilistic reasoning, as modeled in Zuraw (2001). However, it is not clear that this level of complexity is necessary—in fact, alternative models for Zuraw's results could be proposed using simple uniform probabilities that explain the data equally well. We show later with our own results that we can model the trajectory successfully without Bayesian probabilities.

9. Results

In our preliminary results, based on approximately 200 simulation runs, we have been able thus far to generate a smooth trajectory for harmony emergence only. Harmony decay has proven to be more challenging. Our simulations generated the following trajectories.

9.1 A too-steep curve.

(9)



In this version of the simulation agents had a fixed probability (fixed at 0.3) of harmonizing words. That is, they all had exactly the same probability of harmonizing, as if they were endowed with a kind of power to peer directly into their neighbors' grammars and know how everyone else would behave. They behaved in a monolithic fashion, as if the goal of the simulation were to learn harmony, which they did in a relatively efficient manner. The curve rises, but too steeply: there is no period of slow change at the start. It is as if the agents were obeying a directive to "go forth and harmonize", rather than harmony evolving organically from social interactions.

Even with various settings of the co-articulation probability, we never generated a real S-curve. Given that this didn't produce the desired output, we moved to a slightly more sophisticated definition for the probability of word change. Here we recognize that not all people in the real world will be equally likely to modify a word. For example, in adopting the word *chauffeur* from French, a speaker from Istanbul is more likely to keep close to the original vowel sounds (as the word *ʃofør*), whereas a villager from eastern Turkey—whose lexicon contains many fewer disharmonic (foreign) words—may harmonize it, to *ʃøfør* (in fact, both variants are attested in colloquial Turkish). That is, the likelihood of harmonizing a word is correlated with the strength of the pattern of harmony. We model this by making the probability of word change linearly proportional to the proportion of harmony in the lexicon. In addition, recognizing that a "pattern" that covers, say, 2% of the lexicon is unlikely to be pervasive enough to be really considered a pattern for an individual, we can set a threshold below which a pattern is not recognized as significant. Mathematically, then, we calculate for an agent *x*:

(10) h(x) = MAXPROBHARM * harmony of x's lexicon

where MAXPROBHARM is a parameter we can set, representing the probability that an agent with a fully harmonic lexicon would mutate a word. Then:

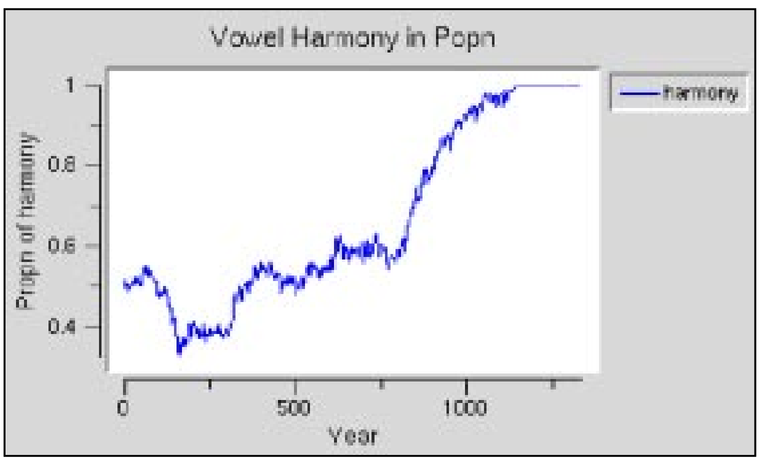
(11) Pr (word change | state of harmony) = h(x) if h(x) > threshold; 0 otherwise

In a sense, agents behave as if they assume that their neighbor's grammar resembles their own, even if they don't know exactly what it is.

9.2 A reasonably good S-curve

The simple heuristics outlined above get us a moderately satisfying result: our first S-shaped curve.

(12)

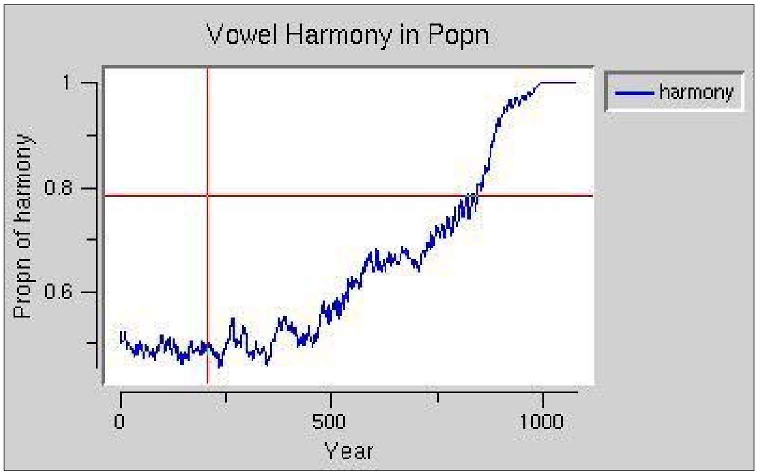


Note that the curve dips down below its starting level of harmony (at 50%) before it begins to rise in an S-curve. This demonstrates that we have note simply programmed into the model the goal of learning harmony.

9.3 A better S-curve

The graph in (13) represents that same simulation settings as the previous one (12), but the probabilistic structure of the model yields a different curve. Here, we get a good S-curve that reaches 100% harmony.

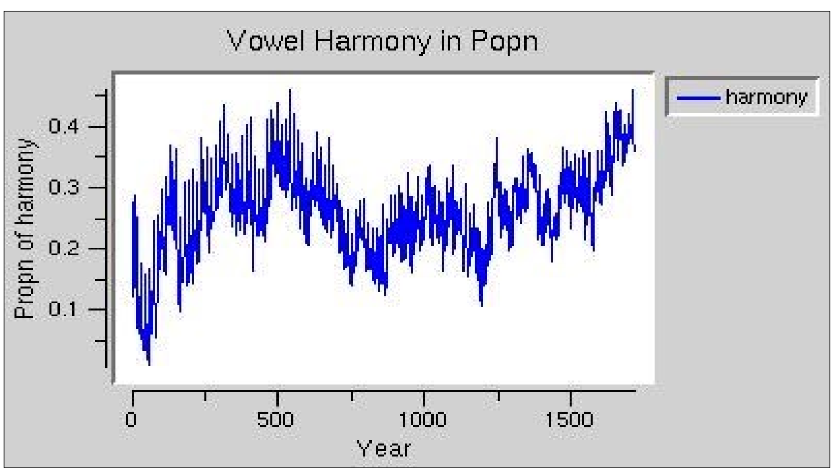
(13)



9.4 Vowel merging

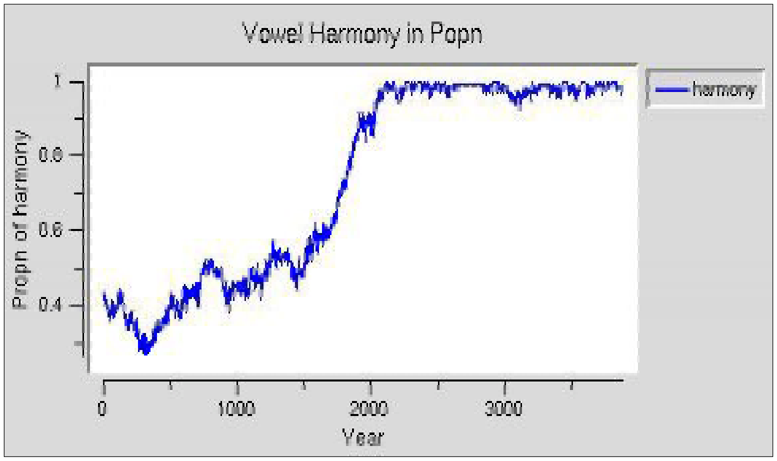
Until now we have only dealt with a single population. Here we introduce a second population (which we call ‘Persians’, as opposed to our original ‘Uzbeks’), who have their own set of parameter values. Their parameter types are the same as for the Uzbeks, except that for the Persians we have implemented vowel merging: the vowels we specify as marked are assimilated to a different vowel. Setting the population to be ≥30% Persians, and giving both the Uzbeks and Persians the same parameter values as in C, we find that the development of vowel harmony no longer follows the S-shaped trajectory. Vowel merging has interfered with the emergence of harmony.

(14)



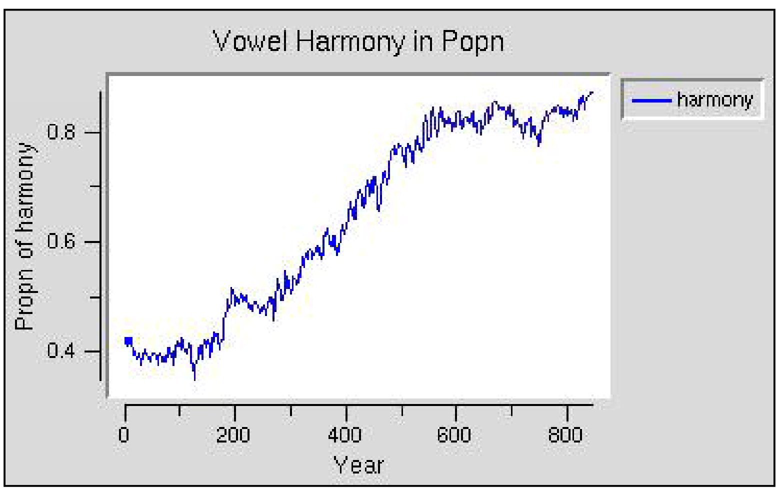
Further, with just a 10% population of Persians we get the curve shown in (15), an S-curve that never achieves perfection, due to long term-stable sub-population of bilinguals that maintains a distinct phonology (it may be an unlikely scenario for this state of affairs to continue for many generations).

(15)



Finally, graph number (16) shows a similar upward S-curve. it was generated by the same settings as simulation number (13), except that the harmony threshold is set lower, to 0.5. The difference between (13) the randomly oscillating one (14), and (16) the upward S-curve, may demonstrate that vowel merger of the Uzbek type is detrimental to harmony but not able to single-handedly destroy it, so long as speakers are proficient at recognizing a harmony pattern.

(16)



9.5 Additional factors needed to model harmony decay

The decay of harmony systems has proven more challenging to model: thus far we have been unsuccessful in generating a downward S-curve. We suspect that this is because the typical historical change scenario involves demographic and language contact factors, and these introduce elements of instability into the system which are not yet reflected in our simulation. We have also to consider the possibility that harmony decay is an entirely different mechanism, and we will not be able to get an S-curve of harmony decay using our current model. This seems reasonable since our model reflects a kind of payoff for having harmony, but there is no obvious payoff for getting rid of harmony (except perhaps that you can have shorter words or more words in your lexicon).

Our current model primarily reflects two factors in harmony decay: (i) loanwords, and (ii) change in vowel inventories (shift, merger). We are revising the model to better reflect additional factors: (iii) morphologization, and (iv) language contact. Morphologization is the widely attested case where harmony is reinterpreted as being the property of certain affixes or certain syllable types. Fixed, non-alternating suffixes can emerge. Formation of new morphemes via grammaticalization can lead to a decline in harmony. As a specific instance, both Turkish and Uzbek have a third person possessive suffix [-i], which was grammaticalized from an independent lexical third person pronoun (Poppe 1965). Early in its status as an affix it did not undergo harmony (Menges 1968); later in Turkish it developed harmonically alternating surface forms (underlyingly archiphonemic /-I/), while Uzbek kept only the original palatal (underlying /-i/). Thus far we have taught our Uzbek agents the possessive morpheme, which on its own exerts only a negligible effect on harmony. To fully model the suffixal aspects of (dis)harmony, we will have to teach our agents a great deal of morphology in addition to the phonology they already know.

A second factor that could prove important, and which will want our model to reflect, is language contact. No backness harmony system that we are aware of has been lost in the absence of intense and prolonged language contact with non-harmonic languages. Contact effects can be reflected in the specific demographics of a simulation, in which we introduce different sized sub-populations.

9.6 Preliminary Conclusions

Our first conclusion is that we have to some extent narrowed down how the S-shape in language change arises. Unlike earlier models, we are able to specify properties of individuals that lead, without explicit programming, to the emergent population behavior. In particular, we found that the most simplistic combinations of factors do not lead to the desired trajectory. However, if an agent does take into account its existing pattern of harmony in evaluating new words, effectively believing that its neighbors are similar to itself by projecting its own pattern onto other agents, it is possible to generate the S-curve of vowel harmony emergence. This incremental modification of the model, to one that produces the upward S-curve, gives us confidence that the model is a starting point for answering linguistic questions.

Our preliminary findings regarding the factors involved in harmony decay are that neither changes in vowel inventories of the type historically attested, nor the influx of foreign loanwords at the levels historically attested, nor the emergence of a disharmonic morphemes are individually strong enough to destroy a harmony system. Rather, these all have to be weighted and combined, in the context of varying population demographics. In Uzbek, the absorption of sizeable populations of Persians speakers who lacked harmony and had a different vowel inventory that was less amenable to harmony might have provided the tipping point which, in combination with the other harmony-destroying factors, led to the loss of harmony.

This suggests the utility of agent-based simulations: they allow us to combine purely internal factors with more external language contact factors and then to explore the effects of these different combinations. We propose that agent modeling opens a number of promising avenues of research, and could become a useful tool for understanding phonological phenomena that show evolution over long periods of time.

10. Addendum: some thoughts on how agent modeling compares to other models

10.1 How agent simulation differs from genetic algorithm models

Genetic algorithms mutate themselves, with lots of genetic agents trying to find the best solution. They evaluate themselves every so often, see which ones have done the best so far; these best ones mate, reproduce and mutate. To do this, you need a "fitness function" that evaluates the quality of the genetic agents. Different fitness functions will choose different agents, and possibly give different locally optimal solutions. Choosing fitness functions is an art, not a science.

In current models (e.g., Pulleyblank & Turkel 2000), the speaker is seen as a genetic agent. S/he has a set of parameters (in the sense of both program parameters and Principles & Parameters) for the language faculty; these might be the V2/non-V2 switch, the preposition/postposition switch, etc. If there are 12 of these, we have a 12-dimensional space, and it's too tricky for equations. What the model needs is to decide on a fitness function. But how do you decide on the fitness of a language (which here is some combination of parameters)?

We don't presuppose any measure of fitness in our agent-based model. It's mostly because we're 'working forwards'—it's like we're generating the data points, with no fixed idea about what parameters to tweak—and they're 'working backwards', trying to work out what the values are for parameters of an already chosen set. We're exploring, and they're trying to solve it as if it's a fixed, quantified problem.

10.2 How agent simulation differs from mathematical models of language change

What both we and they are measuring is the proportion of the phenomenon of interest at time n (call it p_n). For us it's harmony, for Niyogi et al. (see, in particular, Niyogi and Berwick 1997) it's speakers of V2 (i.e. those with verb-second syntax). There are two important differences: first, the *level of granularity*. We don't have to group speakers together in any way at all: each one's behavior and attributes are determined solely through the interaction with other agents. Each agent changes individually, and every agent can potentially have slightly different behavior and attributes. p_n is then determined by looking at all agents.

For mathematical models, the speakers have to be grouped. In current models there are two groups, speakers of language 1 and speakers of language 2 (L1 and L2 respectively), whose proportions are given by p_n and $(1-p_n)$. The groups are assumed to be homogeneous. p_n is then determined by an equation, say (from Niyogi and Berwick 1997):

(17) $p_n = p_{n-1} * (1-a) / ((1-b) + (b - a) * p_{n-1})$

for some coefficients *a* and *b*; it's a function of the proportion at time *n-1* (i.e., p_{n-1}).

The mathematical model could be finer-grained: e.g., the groups could be divided into say northern and southern version of speakers of languages 1 and 2, giving four groups. At the limit, we end up with our agent situation and an incredibly complicated formula.

A second difference is in the approach to *probabilistic behavior*: Our speakers are probabilistic, just like behavior in the real world. This isn't the case in current mathematical models of language. If $a=b$ in the equation above (which translates to there being an equal amount of evidence for language 1 and language 2 available to language learners *in the aggregate*), then it simplifies to $p_n = p_{n-1}$; that is, no change occurs (static equilibrium), and it has to be that way. It's like saying that if you've been flipping a coin, and the proportion of heads has been 0.5 (say, 50 out of 100 flips), that after the next coin flip (and the next and the next) that the proportion has to stay at exactly 0.5 (so there'll be 51 heads in 102 flips, 52 in 104 flips, etc). That's what stochastic models are for; they're not used currently in math models of language change, and it's likely the equations you'd end up with would be too complex. For example, the single variable p_n would have to be replaced by (probably) a binomial distribution.

References

Akobirov S.F., et al. 1981. *Uzbek tilining izohli lughati.* Moscow: Rus tili nashrieti.

Anderson, Gregory & K. David Harrison. 2000. Tuvan-Tofa-English Lexicon. At http://www.mpi.nl/world/celex/.http://www.mpi.nl/world/celex/.

Axtell, Robert and Joshua Epstein. 1996. *Growing Artificial Societies: Social Science from the Bottom Up.* Cambridge. MA: MIT Press.

Bailey, C.-J. 1973. Variation and linguistic theory. Washington, DC: CAL.

Batmanova, I. A. Ed. 1971. *Drevnie tyurkskie dialekty.* Frunze: Ilim.

Boyce, Suzanne. 1988. *The Influence of Phonological Structure on Articulatory organization in Turkish and in English: Vowel Harmony and Coarticulation.* Yale University Doctoral Dissertation.

Briscoe, E. 1999. Grammatical Acquisition and Linguistic Selection. In *Linguistic evolution through language acquisition: formal and computational models.* (ed.) Briscoe, E. Cambridge, UK: Cambridge University Press.

Briscoe, Ted. 2000. Macro and Micro Models of Linguistic Evolution. Paper presented at EvoLang 2000, Paris.

Clark, R. and I. Roberts. 1993. A Computational Model of Language Learnability and Language Change. *Linguistic Inquiry*. 24(2): 299-345.

Dirks, William D. 2001. *Zangâri kema: an Uzbek-English Dictionary*. At http://www.geocities.com/zangari_kema/.

Fazylov, E. I. 1972. *Starouzbekskiy yazyk.* Tashkent: Akademia Nauk.

Fomkin, M. S. 1994. *Sultan Veled i ego tiurkskaia poeziia.* Moskva: Nauka.

Gilson, Erika H. 1987. *The Turkish grammar of Thomas Vaughan: Ottoman-Turkish at the end of the XVIIth cent. from an English "Transkriptionstext."* Wiesbaden: Harrassowitz.

Inkelas, Sharon, et al. 2001. Turkish harmony and co-articulation Paper presented at LSA annual meeting.

______ 1997. Turkish Electronic Living Lexicon. at http://ist-socrates.berkeley.edu:7037/TELLhome.html.

Ismatulla, Khayrulla. 1995. *Modern literary Uzbek I.* Bloomington: Indiana University Press.

Kaun, Abigail, 1995. *The Typology of Rounding Harmony: An Optimality Theoretic Approach.* UCLA, Doctoral Dissertation.

Kondrat'ev, V. G. 1981. *Grammaticheskij stroj jazyka pamjatnikov drevnetjurkskoj pismennosti VIII-X vv.* Leningrad: Izdatel'stvo Leningradskogo Univeristeta.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. Language Variation and Change. 1:199–244.

Lastowka, Garret. 1996. *Turkmen-English Dictionary.* At www.Chaihana.Com/Dict.Doc.

Liberman, Mark. 2001. The lexical contract: Modeling the emergence of word pronunciation. Ms. University of Pennsylvania.

Mämmedow, Nazar. 2001. *Turkmen-English dictionary*. At http://kemine.tripod.com /edict/index.htm.

Menges, Karl. 1968. *The Turkic languages and peoples. An introduction to Turkic studies.* Wiesbaden: Harrassowitz.

Niyogi, P. & R. Berwick. 1997. Evolutionary Consequences of Language Learning. *Linguistics and Philosophy*. 20: 697-719.

Ohala, J. J. 1994. Towards a universal, phonetically-based, theory of vowel harmony. Proceedings, ICSLP 94, Yokohama, 18-22. Sept. 1994. 491-494.

Poppe, Nikolai N. 1965. *Introduction to Altaic linguistics.* Wiesbaden: Harrassowitz.

Pulleyblank, Douglas and & W.J. Turkel. 2000. Learning Phonology: Genetic Algorithms and Yoruba Tongue Root Harmony. In *Optimality Theory: Phonology, Syntax and Acquisition.* Eds. J. Dekkers, et al., 554-591. Oxford Univ. Press.

Satterfield, Theresa. 2000. The Socio-genetic solution: A new look at language genesis through swarm modeling. Ms. University of Michigan.

Schütze, Odön. 1968. *An Armeno-Kipchak chronicle on the Polish-Turkish wars in 1620-1621.* Budapest: Akadémiai Kiadó.

Suomi, Kari. 1983. Palatal vowel harmony: a perceptually motivated phenomenon? *Nordic Journal of Linguistics.*

Trubetzkoy, N. S. 1969. *Principles of Phonology.* trans. by Christiane A.M. Balthaxe Berkeley, University of California Press.

Vroomen, J.; Tuomainen, J.; de Gelder, B. 1998. The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory & Language.* Vol. 38(2): 133-149.

Zuraw, Kie. 2001. The extension of semiproductive morphology to novel forms: An optimality-theoretic model and simulation. Paper presented at WCCFL XX, USC.

K. David Harrison, Mark Dras, Berk Kapicioglu

K. David Harrison
Department of Linguistics
Swarthmore College
Swarthmore, Pennsylvania

dharris2@swarthmore.edu

Mark Dras
Department of Computing
Macquarie University
Sydney, Australia

madras@ics.mq.edu.au