A Dynamical Model of Change in Phonological Representations: The Case of Lenition[1]

Adamantios Gafos and Christo Kirov

1. Introduction

This paper presents a model of diachronic changes in phonological representations. The broad context is that of sound change, where word representations evolve at relatively slow time scales as compared to the time scale of assembling a phonological representation in synchronic word production. The specific focus is on capturing certain key properties of an unfolding lenition process.

Diachronic changes in phonological representations accumulate gradually during repeated production-perception loops, that is, through the impact of a perceived word on the internal representation and subsequent production of that word or other related words. To formally capture the accumulation of such changes, we capitalize on the continuity of parameter values at the featural level. Our model is illustrated by contrasting it with another recent view exemplified with an exemplar model of lenition, which also embraces continuity in its representational parameters. A primary concern is providing a formal basis of change in phonological representations using basic concepts from the mathematics of dynamical systems.

2. The Case of Lenition

The term “lenition” is used to describe a variegated set of sound alternations such as voicing of obstruents between two vowels, spirantization of stops in a prosodically weak position, and devoicing of obstruents in syllable-final position which are either attested synchronically or have resulted diachronically in a restructuring of the phonemic inventory of a language. One diachronic example of lenition is Grimm’s Law, according to which Proto-Indo-European voiceless stops became Germanic voiceless fricatives (e.g. PIE *[t] > Gmc *[θ]). Other examples of sound changes described as cases of lenition are given below (for recent studies see Gurevich, 2004;

Cser, 2003 and Lavoie, 2001). In each case, a stop turns to a fricative similar in place of articulation to the original stop.

(1) a. Southern Italian dialects: [b d g] → [v ð ɣ] intervocalically.
b. Greek (Koine): [p^h t^h k^h] → [f θ x] except after obstruents.
c. Proto-Gaelic: [t k] → [θ x] intervocalically.
d. Hungarian: [p] → [f] word initially.

Consider any single transition between two states of a lenition process, say, starting with a stop [b] and resulting in a fricative [v], [b] > [v]. At a broad level, one can describe two kinds of approaches to this kind of transition. The symbolic approach, as exemplified by Kiparsky's classic paper on linguistic universals and sound change (Kiparsky, 1968), studies the internal composition of the individual stages (e.g. feature matrices at each stage) and makes inferences about the nature of the grammar and the representations. The continuity of sound change, that is, how the representation of the lexical item containing a [b] changes in time to one containing an [v], is not studied. This is in part due to the theoretical assumption that representations are discrete. That is, there is no symbol corresponding to an intermediate degree of stricture between that of a stop and a fricative. In the dynamical approach, the transition process between the stages is studied at the same time as the sequence of stages. In what follows, we instantiate a small, yet core part of a dynamical alternative to the symbolic model of sound change.

2.1. An Exemplar Model of Lenition

It is useful to describe the main aspects of our model by contrasting it with another model proposed recently by Pierrehumbert. This is a model of sound change aimed at accounting for certain generalizations about lenition, extrapolated from observations of synchronic variation or sound changes in progress. The model proposed in Pierrehumbert (2001) has two attractive properties. It offers a way to represent the fine phonetic substance of linguistic categories, and it provides a handle on the effect of lexical frequency in the course of an unfolding lenition process.

In Pierrehumbert's discussion of lenition, it is assumed that the production side of a lenition process is characterized by the following set of properties.

Table 1. Properties of Lenition

Properties of Lenition
1. Each word displays a certain amount of variability in production.
2. The effect of word frequency on lenition rates is gradient.
3. The effect of word frequency on lenition rates should be observable within the speech of individuals; it is not an artifact of averaging data across the different generations which make up a speech community.
4. The effect of word frequency on lenition rates should be observable both synchronically (by comparing the pronunciation of words of different frequency) and diachronically (by examining the evolution of word pronunciations over the years within the speech of individuals.)
5. The phonetic variability of a category should decrease over time, a phenomenon known as entrenchment. The actual impact of entrenchment on lenition is not clear, and Pierrehumbert does not cite any data specific to entrenchment for this particular diachronic effect. In fact, while a sound change is in progress, it seems equally intuitive (in the absense of any data to the contrary) that a wider, rather than narrower range of pronunciations is available to the speaker. Pierrehumbert uses the example of a child's productions of a category becoming less variable over time, but this may only apply to stable categories, rather than ones undergoing diachronic change. It may also be orthogonal to the child's phonetic representations, and rather be due to an initial lack of biomechanical control. For these reasons, therefore, our own model is not designed to guarantee entrenchment while sound change is taking place, but does show entrenchment effects for diachronically stable categories.

The first property, variability in production, does not apply exclusively to lenition process, but rather it is a general characteristic of speech production that any lenition model should be able to capture. The frequency related properties are based on previous work by Bybee who claims that at least some lenition processes apply variably based on word frequency (Bybee, 2003). Examples include schwa reduction (e.g. *memory* tends to be pronounced [mɛmri]) and t/d-deletion (e.g. *told* tends to be pronounced [tol]). Once a lenition process has begun, Bybee's claim amounts to saying that words with high frequency will weaken more quickly over time than rare words. Consequently, lenition effects can be seen both synchronically and diachronically. Synchronically, a more frequent word will be produced more lenited (with more undershoot) than a less frequent word in the current speech of a single person. Diachronically, all words in a language will weaken across all speakers, albeit at different rates.

What are the minimal prerequisites in accounting for the lenition properties above? First, it is clear that individuals must be capable of storing phonetic detail within each lexical item. We also need a mechanism for gradiently changing the lexical representations over time. To do this, the perceptual system must be capable of making fine phonetic distinctions, so that the information carried by these distinctions can reach the currently spoken item in the lexicon.

Pierrehumbert's exemplar-based model of lenition gives explicit formal content to each of these prerequisites (Pierrehumbert, 2001). The model is built on a few key ideas, which can be described in brief terms. Specifically, in the exemplar-based model, a given linguistic category is stored in a space whose axes define the parameters of the category. In Pierrehumbert (2001), it is suggested that vowels, for example, might be stored in an F1/F2 formant space. This space is quantized into discrete cells based on perceptual limits. Each cell is considered to be a bin for perceptual experiences, and Pierrehumbert views each bin to be a unique potential exemplar. When the system receives an input, it places it in the appropriate bin. All items in a bin are assumed to be identical as far as the perceptual system is concerned, and the more items in a particular bin, the greater the activation of the bin is. All bins start out empty and are not associated with any exemplars that have actually been produced and/or perceived (memory begins as a *tabula rasa*). When a bin is filled, this is equivalent to the storage of an exemplar. The new exemplar is given a categorical label based on the labels of other nearby exemplars. This scheme limits the actual memory used by exemplars. There is a limited number of discrete bins, and each bin only stores an activation value proportional to the number of exemplar instances that fall into it. Thus, not all the exemplar instances need to be stored. A decay process decreases the activation of an exemplar bin over time, corresponding to memory decay. Figure 1, taken from Pierrehumbert (2001), shows the F2 space discretized into categorically labeled bins.

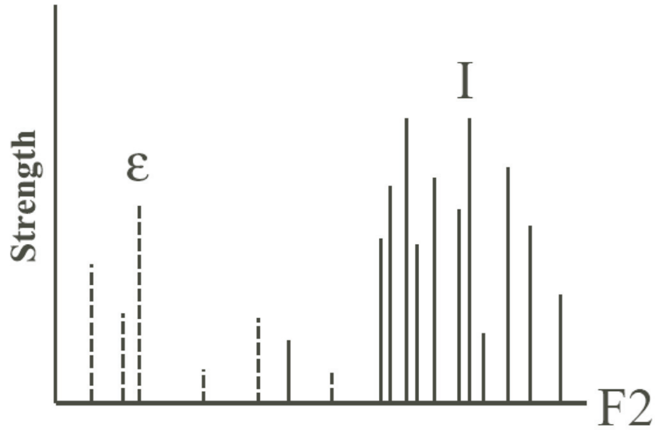

Figure 1. Exemplar bins with varying activations.

The set of exemplars with a particular category label constitutes an extensional approximation of a probability distribution for that category over the storage space. Given the coordinates in the storage space over which a category is defined, that distribution would provide the likelihood that a token with those coordinates would belong to that category (e.g. how likely is it that the token is an /a/). During production, a particular exemplar from memory is chosen to be produced, where the likelihood of being chosen depends on how activated the exemplar is. The chosen exemplar is shifted by a bias in the direction of lenition. This bias reflects the synchronic phonetic motivation for lenition. This includes at least the tendency to weaken the degree of oral constriction in contexts favoring segmental reductions, e.g. in non-stressed syllables, syllable codas, or intervocalically. For relevant discussion see Beckman, et al. (1992) and Wright (1994). To account for entrenchment (see Table 1(5)), Pierrehumbert extends this production model by averaging over a randomly selected area of exemplars to generate a produced candidate. Since the set of exemplars defines a probability distribution (in an extensional sense), weighing the average by each exemplar's probability results in a production candidate pushed toward the center of the distribution.

The exemplar scheme described in this section derives the five properties of lenition discussed earlier as follows. Variability in production is directly accounted for since production is modeled as an average of the exemplar neighborhood centered around a randomly selected exemplar from the entire set stored in the system. Each lexical item has its own exemplars, and each production/perception loop causes the addition of a new exemplar to the set. This new exemplar is more lenited than the speaker originally intended due to biases in production, so the distribution of exemplars skews over time. In a given period of time, the number of production/perception loops an item goes through is proportional to its frequency. Thus, the amount of lenition associated with a given item shows gradient variation according to the item's frequency (Dell, 2000). As all processes directly described by the exemplar model occur within a single individual, lenition is clearly observable within the speech of individuals. Diachronically, lenition will proceed at a faster rate for more frequent items because they go through more production/perception loops in a given time frame. The synchronic consequence of this is that at a point in time, more frequent items will be more lenited in the speech of an individual than less frequent items. Finally, entrenchment is a consequence of averaging over several neighboring exemplars during production, shifting the resulting production towards the mean of the distribution described by all the exemplars.

In sum, the exemplar-based model offers a direct way to encode phonetic details, and captures the assumed effects of frequency on lenition. Pierrehumbert further claims that the exemplar model is the only type of model that can properly handle the above conception of lenition (Pierrehumbert, 2001:147). In what follows, we will propose an alternative dynamical model of lenition. The dynamical model can also account for the lenition properties reviewed above. But it is crucially different from the exemplar-based model in two respects. The dynamical model encodes phonetic details while maintaining unitary category representations as opposed to representations defined extensionally by collections of exemplars. In addition, the dynamical model also admits a temporal dimension, which is currently not part of the exemplar-based model.

2.2. A Dynamical Model of Lenition

2.2.1. Description of the model

Studying language change as a process occurring in time broadly motivates a dynamical approach to modeling. A *dynamical model* is a formal system whose internal state changes in a controlled and mathematically explicit way over time. The workings of the proposed model are based on a dynamical formalism called *Dynamic Field Theory* (Erlhagen & Schöner, 2002).

A central component of our model is the spatio-temporal nature of its representations. Take a lexical item containing a tongue tip gesture as that for /d/. We can think of the specification of the speech movements associated with this gesture as a process of assigning values to a number of behavioral parameters. In well-developed models that include a speech production component, these parameters include constriction location and constriction degree (Guenther, 1995; Saltzmann & Munhall, 1989; Browman & Goldstein, 1990). A key idea in our model is that each such parameter is not specified exactly but rather by a distribution depicting the continuity of its phonetic detail.

Although our model does not commit to any specific phonological feature set or any particular model for the control and execution of movement, to illustrate our proposal more explicitly let us assume the representational parameters of *Articulatory Phonology* (Saltzmann & Munhall, 1989; Browman & Goldstein, 1990). Thus, let us assume that lexical items must at some level take the form of gestural scores. A gestural score, for current purposes, is simply a sequence of gestures (we put aside the intergestural temporal relations that also must be specified as part of a full gestural score). For example, the sequence /das/ consists of three oral gestures - a tongue tip gesture for /d/, a tongue dorsum gesture for /a/, and a tongue tip gesture for /s/. Gestures are specified by target descriptors for the *vocal tract variables* of *Constriction Location* (CL) and *Constriction Degree* (CD), parameters defining the target vocal tract state. For example, /d/ and /s/ have the CL target descriptor {alveolar}. The CD descriptor of /d/ is {closure} and for /s/ it is {critical}. These descriptors correspond to actual numerical values. For instance, in the tongue tip gesture of a /d/, {alveolar} corresponds to 56 degrees (where 90 degrees is vertical and would correspond to a midpalatal constriction) and {closure} corresponds to a value of 0 mm.

In our model, each parameter is not specified by a unique numerical value as above, but rather by a continuous activation field over a range of values for the parameter. The field captures among other things a distribution of activation over the space of possible parameter values so that a range of more activated parameter values is more likely to be used in the actual execution of the movement than a range of less activated parameter values. The parameter fields then resemble distributions over the continuous details of vocal tract variables. A lexical item therefore is a gestural score where the parameters of each gesture are represented by their own fields. Schematic fields corresponding to the (oral) gestures of the consonants in /das/ are given in Figure 2.

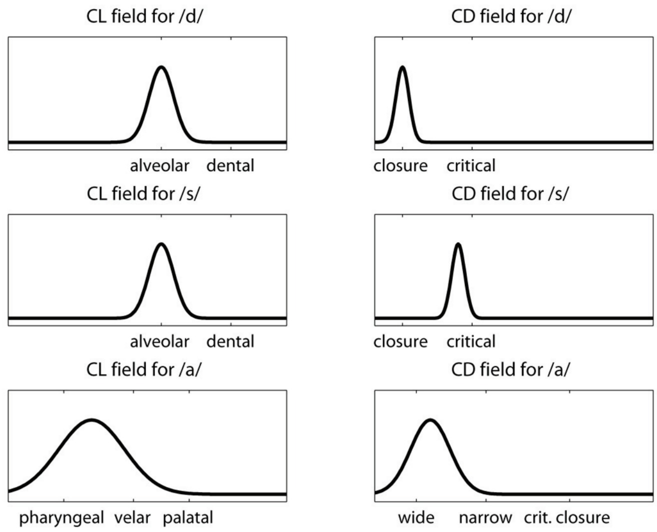

Figure 2. Component fields of /d/, /s/, and /a/. *y*-axis represents activation. /d/ and /s/ have nearly identical CL fields, as they are both alveolars, but they differ in CD.

Formally, parameters are manipulated using the dynamical law from Dynamic Field Theory (Erlhagen & Schöner, 2002). The basic dynamics governing each field are described by:

$$\tau dp(x,t) = -p(x,t) + h + input(x,t) + noise \qquad (1)$$

where *p* is the field in memory (a function of a continuous variables *x*, *t*), *h* is the field's resting activation, *dp*(*x*,*t*) is the change in activation at *x* at time *t*, τ is a constant corresponding to the rate of decay of the field (i.e. the rate of memory decay), and *input*(*x*,*t*) is a field representing time dependent external input to the system (i.e. perceived token) in the form of a localized activation spike.

The equation can be broken down into simpler components to better understand how it functions. The core component $\tau dp(x,t) = -p(x,t) + h$ is an instance of exponential decay. If we arbitrarily select a value for *x*, and plot *p*(*x*,*t*) over time, we will see behavior described by the exponential decay equation. In the absence of any input or interaction, the activation at *p*(*x*,*t*) will simply decay down to its resting level, *h*, as shown in Figure 3. If *p*(*x*,*t*) starts at resting activation, it will remain there forever. In the terminology of dynamical systems, the starting activation of a point is known as an *initial condition*, and the activation it converges to, in this case the resting activation, is known as an *attractor*. If the input term, *input*(x,t), is non-zero, then the system will move towards a point equivalent to its resting activation plus the input term. The speed of the process is modulated using the τ term.

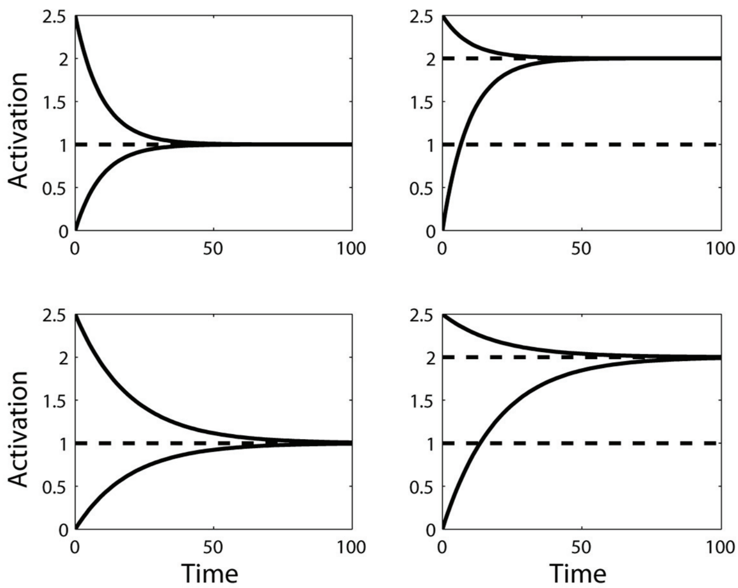

Figure 3. Top left: In the absence of input, field activation at a particular point converges to the resting level h = 1 (dashed line) (τ = 10). Top right: With added input *input*(x,t) = 1, activation converges to resting level h = 1 plus input (top dashed line)(τ = 10). Bottom left: In the absence of input, node activation converges to resting level r = 1 (τ = 20). Bottom left: With added input *input*(x,t) = 1, activation converges to resting level r = 1 plus input (τ = 20).

Fields are spatio-temporal in nature. Thus specifying the value of a gestural parameter is a spatio-temporal process in our model. We describe each of these aspects, spatial and temporal, in turn. The spatial aspect of the gestural specification process corresponds to picking a value to produce from any of the fields in Figure 2, e.g. choosing a value for Constriction Location for /d/ and /s/ from within the range of values corresponding to the [alveolar] category. This is done by sampling the Constriction Location field, much as we might sample a probability distribution. Since each field encodes variability within the user's experience, we are likely to select reasonable parameter values for production. A demonstration of this is shown in Figure 4. The noisy character of the specification process allows for variation in the value ultimately specified, but as the series of simulations in Figure 4 verifies the selected values cluster reliably around the maximally activated point of the field.

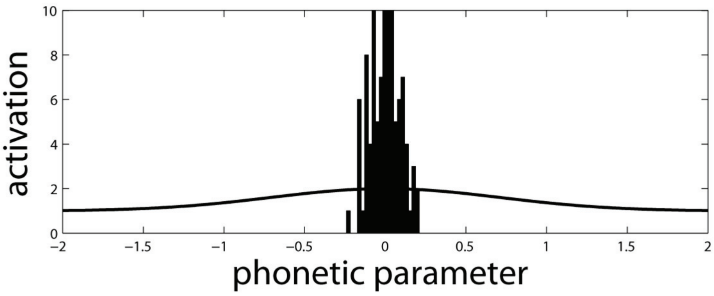

Figure 4. Variability in production. Histogram of selected values over 100 simulations of gestural specification. Histogram overlaid on top of field to show clustering of selected values near the field maximum.

The specification process presented here is similar but not identical to the sampling of a probability distribution. Fields have unique properties that make them useful for modeling memory. Unlike distributions, fields need not be normalized to an area under the curve of one. The key addition here is the concept of *activation*. Fields can vary from one another in total activation while keeping within the same limits of parameter values. Because of this added notion of activation, the specification process is more biased towards the maximally activated point in the field (i.e. the mean of the distribution) than a true random sampling would be. This leads to an entrenchment effect for categories not undergoing change. This behavior is shown in Figure 5. In addition, fields have a resting activation level (a lower-limit on activation). This level slowly tends to zero over time, but increases every time the field is accessed during production or perception. Thus, lexical items whose fields are accessed more frequently have higher resting activation levels than lexical items whose component fields are accessed less frequently. Finally, much as memory wanes over time, activation along a field decays if not reinforced by input.

The other crucial aspect of the specification process is its time-course. Formalizing gestural parameters with fields adds a time-course dimension to the gestural specification process. Thus, if a lexical representation contains a /d/, the CD and CL parameters for this /d/ are not statically assigned to their (language- or speaker- specific) canonical values, e.g., CL = [alveolar]. Rather, assigning values to these parameters is a time-dependent process, captured as the evolution of a dynamical system over time. In short, lexical representations are not static units. This allows us to derive predictions about the time-course of choosing or specifying different gestural parameters.

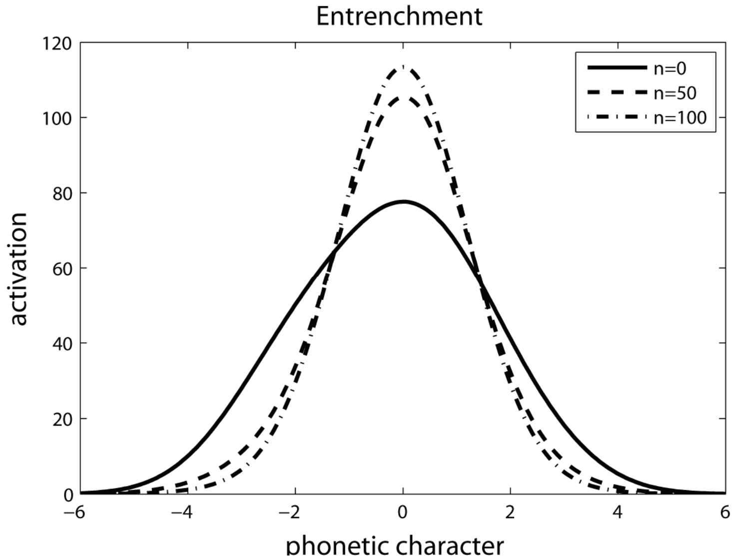

Figure 5. Output of entrenchment simulation. The *x*-axis represents a phonetic dimension (e.g. constriction degree). The field defining the distribution of this parameter is shown at various points in time. As time progresses, the field becomes narrower.

The specification process begins by a temporary increase in the resting activation of the field, i.e. pushing the field up, caused by an intent to produce a particular lexical item (which includes a gesture ultimately specified for a parameter represented by this field). Activation increases steadily but noisily until some part of the field crosses a decision threshold and becomes the parameter value used in production. This scheme ensures that the areas of maximum activation are likely to cross the decision threshold first. After a decision has been made, resting activation returns to its pre-production level. The following equation represents this process mathematically:

$$\tau dh / dt = -h + h_0 + \delta\big(d, \max(p)\big) * h + noise \qquad (2)$$

where h is the temporarily augmented resting activation, τ is a time scaling parameter, h_0 is the pre-production resting activation level, $\delta(d, \max(p))$ is a nonlinear sigmoid or step function over the distance between the decision threshold d and the maximum activation of field p, and *noise* is scaled gaussian noise. While the distance is positive (the decision threshold has not yet been breached), the δ function is also positive and greater than 1, overpowering the $-h$ term and causing a gradual increase in the resting activation h. When the decision threshold is breached, the δ function becomes 0, and remains clamped at 0 regardless of the subsequent field state, allowing the $-h$ term to bring activation back to h_0.

The gestural specification process is affected by the pre-production resting activation of the field, in that a field with high resting activation is already "presampled", and thus automatically closer to the decision threshold. This leads to faster decisions for more activated fields, and by extension more frequent parameter values. The relevant simulations are described below. Figure 6 shows representative initial fields, and Figure 7 shows the progression of the featural specification process over time. We see that given two fields identical in all respects except for resting activation, the field with the higher resting activation reaches the decision threshold first.

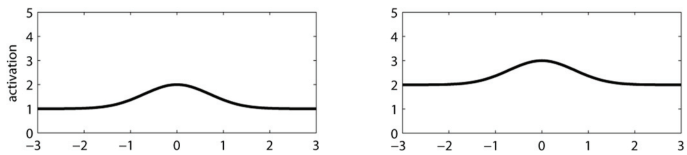

Figure 6. The two fields are identical except for resting activation: $h_0 = 1$ (left), $h_0 = 2$ (right). The x-axis is arbitrary.

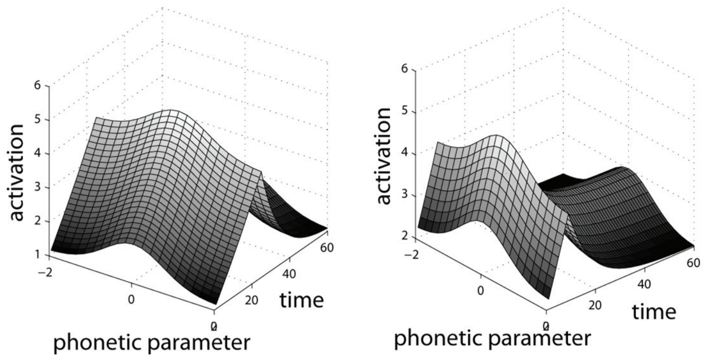

Figure 7. Sampling was simulated with a decision threshold $d = 5$, $\tau = 10$, and *noise* = 0. The first field (left) reached the decision threshold at $t = 25$,

and the second field (right) reached the decision threshold at $t = 9$ (where t is an arbitrary unit of simulation time). The field with higher initial resting activation reached the decision threshold faster. Both fields return to their pre-production resting activation after decision threshold is reached.

We now discuss the ways in which representing gestural parameters by fields relates to other proposals.

The field equation used in our model parallels the exemplar model in many ways, but encapsulates much of the functionality of that model in a single dynamical law which does not require the storage of exemplars. Memory wanes over time as the field decays, much as older exemplars are less activated in the exemplar model. Input causes increased activation at a particular area of the field, much as an exemplar's activation is increased with repeated perception. This activation decays with time, as memory does.

Perhaps the most crucial difference between our model and the exemplar model described earlier is the time-course dimension. In the exemplar model discussed, the assignment of a value to a parameter does not have any time-course. The process is instantaneous. The same is true for the relation between our model and those of Saltzman & Munhall (1989), Browman & Goldstein (1990).

Using fields is a generalization of a similar idea put forth in Byrd & Saltzman (2003), where gestural parameters are stored as ranges of possible values. In our model, each range is approximated by an activation field in memory. Finally, representing targets by activation fields is also a generalization of two well-known proposals about the nature of speech targets, Keating's "windows" (Keating, 1990) and Guenther's "convex regions" (Guenther, 1995). In Guenther's model of speech production, speech targets take the form of convex regions over orosensory dimensions. Unlike other properties of targets in Guenther's model, the convexity property does not fall out from the learning dynamics of the model. Rather, it is an enforced assumption. No such assumption about the nature of the distributions underlying target specification need be made in our model.

2.2.2. Lenition in the Dynamical Model

When a lexical item is a token of exchange in a communicative context, phonetic details of the item's produced instance may be picked up by perception. This will have some impact on the stored instance of the lexical item. Over longer time spans, as such effects accumulate, they trace out a path of a diachronic change. Our model provides a formal basis for capturing change at both the synchronic and the diachronic dimensions.

We focus here on how a single field in a lexical entry is affected in a production-perception loop. The crucial term in the field equation is the input term *input*(x,t). This input term *input*(x,t) represents sensory input. More specifically, input is a peak of activation registered by the speech perception module. This peak is located at some detected x-axis value along the field. This value is assumed to be sub-phonemic in character. For example, we assume that speakers can perceive gradient differences in Voice Onset Time values, constriction location, and constriction degree within the same phonemic categories. In the current model, the input term is formulated as $e^{-(x-off)^2}$, where *off* is the detected value or offset along the x-axis of the field.

The spike corresponding to the input term *input*(x,t) is directly added to the appropriate field, resulting in increased activation at some point along the field's x-axis. A concrete example is presented in Figure 8. Once input is presented, a system can evolve to a stable attractor state, that is, a localized peak at a value corresponding to the input. The state is stable in the sense that it can persist even after the input has been removed. In effect, the field for the lexical item has retained a memory of the sub-phonemic detail in the recently perceived input. The process of adding gaussian input spikes to an existing field is analogous to the storage of new exemplars in the exemplar model. The field, however, remains a unitary function. It is an intensional representation of a phonetic distribution. A growing set of exemplars is an extensional representation.

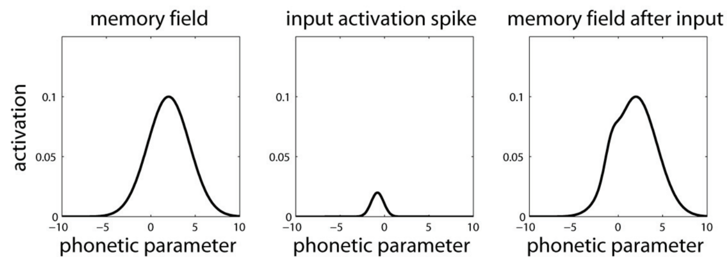

Figure 8. (Left) Field representing a phonetic parameter of a lexical item in memory. (Middle) Input function (output of perception corresponding to *input*(x,t) in Equation 1). Represents a localized spike in activation along the field, corresponding in location to, for example, the constric-

tion degree of the input. (Right) Field of lexical item in memory after input is added to it. Field shows increased activation around area of input.

Since activation fades slowly over time, only areas of the field that receive reinforcement are likely to remain activated. Thus, a peak in activation may shift over time depending on which region of the field is reinforced by input. In terms of the lenition model this means that regions of the field representing a less lenited parameter fade while regions representing a more lenited parameter are kept activated by reinforcement from input.

The interaction between localized increase in activation based on input and the slow fading of the field due to memory decay is the basic mechanism for gradual phonetic change. Since activation fades slowly over time, only areas of the field that receive reinforcement are likely to remain activated. So, a peak in activation may shift over time depending on which region of the field is reinforced by input. Regions of the field representing a less lenited parameter fade while regions representing a more lenited parameter are kept activated by reinforcement from input.

Given an initial field (a preshape) representing the current memory state of a lexical item, we can simulate lenition using the model described above. Figure 9 shows the results of one set of simulations. Shown are the state of the simulation at the starting state, after 50 samples of a token, and after 100 samples (in the simulations, the number of samples is small but each sample produces a large effect on the field). Each time step of the simulation corresponds to a production/perception loop. Production was performed as described above by picking a value from the field and adding noise and a bias to it. This produced value, encoded by an activation spike of the form $e^{-(x-off)^2}$, where $off = sample(p) + noise + bias$, was fed back into the system as input.

As can be seen in Figure 9, at the point when lenition begins, the field represents a narrow distribution of activation and there is little variability when sampling the field during production. As lenition progresses, the distribution of activation shifts to the left. During this time the distribution becomes asymmetrical, with a tail on the right corresponding to residual traces of old values for the parameter. It also grows wider, corresponding to an increase in parameter variation while the change occurs.

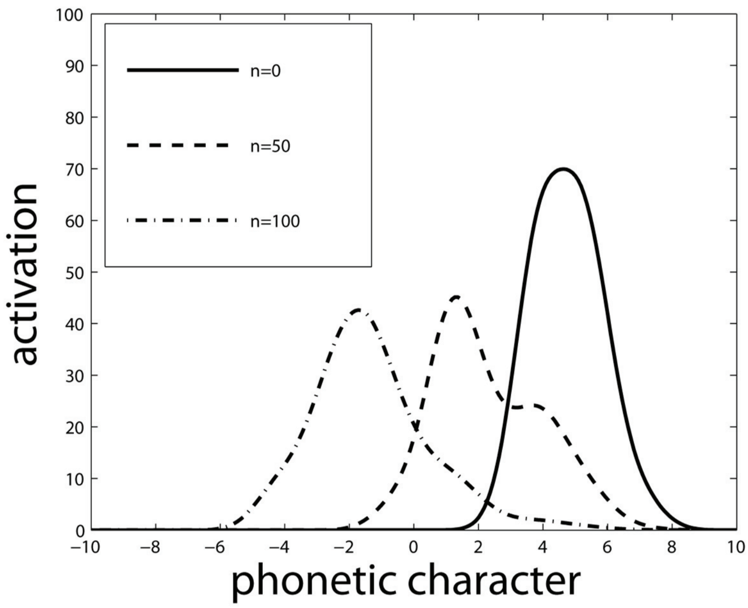

Figure 9. Output of lenition simulation. The *x*-axis represents a phonetic dimension (e.g. constriction degree during t/d production). Each curve represents a distribution of a particular category over the *x*-axis at a point in time. As time progresses, the distribution shifts to the left (i.e. there is more undershoot/lenition) and becomes broader.

With small changes in parameterization, our model can more closely represent the entrenchment behavior seen in Pierrehumbert (2001). In Figure 10, lowering the strength of memory decay by adding a constant $\varepsilon < 1$ factor in the $-p(x,t)$ term in Equation 1, results in less flattening of the parameter field as lenition proceeds. However, the distribution retains a wide tail of residual activation around its base.

To keep the field narrow as time proceeds, we can alternate between production/perception cycles with a production bias and without. This resembles production of the category in contexts where the phonetic motivation for the bias is present versus contexts where it is absent (e.g. prosodically weak versus strong positions). In effect, Figure 11 was created by biasing only every other simulated production. This was done in addition to lowering the strength of memory decay as discussed above.

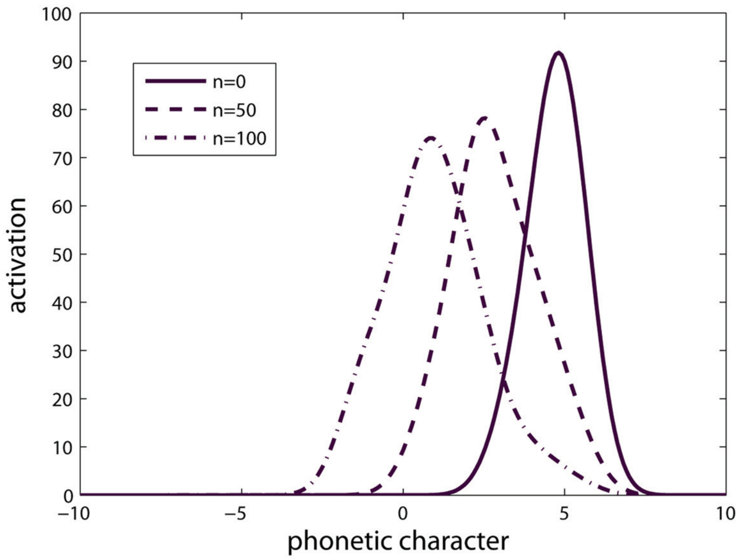

Figure 10. Lowering memory decay results in less flattening of the field as the lenition simulation proceeds.

Like the exemplar model above, the model described in this section can derive the properties of lenition assumed by the exemplar model. Here we enumerate the functional equivalence of the two models with respect to the properties of lention assumed by the exemplar model. In the dynamical model, variability in production is accounted for by noise during the gestural specification process. Each lexical item has its own fields and each production/perception loop causes a shift in the appropriate field towards lenition due to biases in production (see Figure 8 for an example of a field starting to skew to the left). In a given period of time, the number of production/perception loops an item goes through is proportional to its frequency. Thus, the amount of lenition associated with a given item shows gradient variation according to the item's frequency. All the processes described here occur within a single individual, so lenition is clearly observable within the speech of individuals. Diachronically, lenition will proceed at a faster rate for more frequent items, again because they go through more production/perception loops in a given time frame. This same mechanism is evident synchronically as well, since at any single point in time, more frequent items will be more lenited than less frequent items.

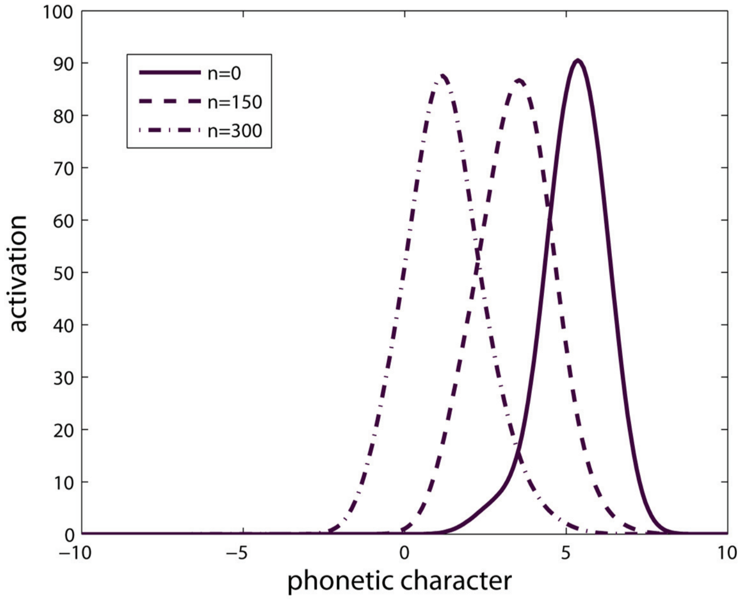

Figure 11. Interleaving biased and non-biased productions leads to consistently narrow field.

In sum, the broad proposal of this section is that diachronic change can be seen as the evolution of lexical representations at slow time scales. The specific focus has been to demonstrate that certain lenition effects, described in a previous exemplar model, can also be captured in our model of evolving activation fields.

3. Conclusion

We have presented a dynamical model of speech planning at the featural or vocal tract variable level. This model allows us to provide an alternative account for lenition in lieu of an exemplar-based model. The dynamical and exemplar models cover the same ground as far as their broad agreement with the assumed properties of an evolving lenition process are concerned. However, there are fundamental high level differences between the two. Tables 2 and 3 contrast properties of the exemplar and dynamical models.

Table 2. Properties of the Exemplar Model

Exemplar Model
1. Every token of a category (where category could mean any item capable of being recognized - word, phoneme, animal cry, etc.) is explicitly stored as an exemplar in memory. A new experience never alters an old exemplar (Hintzman, 1986).
2. the complete set of exemplars forms an extensional definition of a probability distribution capturing variability of a category.
3. Distributions are altered by storing more exemplars.

Table 3. Properties of the Dynamical Model

Dynamical Model
1. Every token of a category is used to dynamically alter a single representation in memory associated with that category, and is then discarded. No exemplars are stored.
2. Variability is directly encoded by the singular representation of a category. The parameters of a category exist as field approximations to probability distributions which are defined intensionally. That is, they are represented by functions, rather-than a set of exemplars.
3. Distributions are altered by dynamical rules defining the impact of a token on a distribution, and changes to the distribution related to the passage of time.

Two key differences are highlighted. First, the dynamical model remains consistent with one key aspect of generative theories of representation. Instead of representing categories extensionally as arbitrarily large exemplar sets, linguistic units and their parameters can have singular representations[2]. These are the fields in our specific proposal. It is these unitary representations, rather than a token by token expansion of the exemplar sets, that drifts in sound change. In this sense, our model is similar to other non-exemplar based models of the lexicon such as Lahiri & Reetz's (2002) model while still admitting phonetic detail in lexical entries (see previous chapter of this volume by Nguyen, Wauquier & Tuller, for relevant discussion).

Second, the dynamical model is inherently temporal. Since both the exemplar and the dynamical model are at least programmatically designed to include production and perception, which unfold in time, this seems to be a key property. In an extension of the present model, we aim to link perceptual to motor representations and to provide an account of the effects of certain lexical factors (such as neighborhood density and frequency) on the time-course of speech production. Such an account would contribute to the larger goal of establishing an explicit link between the substantial literature on the time-course of word planning and linguistic theories of representation.

Acknowledgments

We wish to thank Matt Goldrick for his comments on an earlier draft. Many thanks also to the two anonymous reviewers and the editors for detailed and cogent commentary on the manuscript. Research supported by NIH Grant HD-01994 to Haskins Labs. AG also acknowledges support from an Alexander von Humboldt Research Fellowship.

Notes

[1] Authors names are listed in alphabetic order. Correspondence should be addressed to both authors, adamantios.gafos@nyu.edu, kirov@cogsci.jhu.edu

[2] It is useful to distinguish the exemplar approach from a version of the dynamical one where multiple different instances of a category are stored, corresponding to different registers, different speakers, etc. For our purposes, each of these subcategories is considered unique and has a singular representation.

References

Beckman, Mary E., de Jong, Ken, Jun, Sun-Ah, & Lee, Sook-hyang.
1992 "The Interaction of Coarticulation and Prosody in Sound Change". *Language and Speech*, 35, 45–58.

Browman, Catherine P., & Goldstein, Louis.
1990 "Gestural Specification Using Dynamically Defined Articulatory Structures". *Journal of Phonetics*,18, 299–320.

Bybee, Joan
2003 "Lexical Diffusion in Regular Sound Change" in Restle, D., & Zaefferer, D. (eds), *Sounds and Systems: Studies in Structure and Change*. Mouton de Gruyter, Berlin, 58–74.

Byrd, Dani & Saltzman, Elliot
2003 "The Elastic Phrase: Modeling the Dynamics of Boundary-Adjacent Lengthening". *Journal of Phonetics*, 31(2), 149–180.

Cser, András
2003 "The Typology and Modelling of Obstruent Lenition and Fortition Processes". Akadémiai Kiadó.

Dell, Gary S.
2000 "Lexical Representation, Counting, and Connectionism" in Broe, Michael B., & Pierrehumbert, Janet (eds), *Papers in Laboratory Phonology V*. Cambridge University Press, Cambridge UK, 335–348.

Erlhagen, Wolfram, & Schöner, Gregor
2002 "Dynamic Field Theory of Movement Preparation". *Psychological Review*, 109, 545–572.

Guenther, Frank H.
1995 "Speech Sound Acquisition, Coarticulation, and Rate Effects in a Neural Network Model of Speech Production". *Psychological Review*, 102, 594–621.

Gurevich, Naomi
2004 "Lenition and Contrast: Functional Consequences of Certain Phonetically Conditioned Sound Changes", in *Outstanding Dissertations in Linguistics Series*. Routledge, NY.

Hintzman, Douglas H.
1986 ""Schema" Abstraction in a Multiple-Trace Memory Memory Model". *Psychological Review*, 93(4), 411–428.

Keating, Patricia A.
1990 "The window model of coarticulation: articulatory evidence" in Beckman, M. E., & Kingston, J. (eds), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, 451–470.

Kiparsky, Paul
1968 "Linguistic Universals and Linguistic Change" in Emmon, Bach, & Robert, Harms (eds), *Universals in Linguistic Theory*. New York: Rinehart and Winston, 170–202.

Lahiri, Aditi & Reetz, Henning
1990 "Underspecified recognition" in Gussenhoven, C., & Warner, N. (eds), *Papers in Laboratory Phonology VII*. Berlin, Germany: Mouton de Gruyter, 637–675.
Lavoie, Lisa
2001 "Consonant Strength: Phonological Patterns and Phonetic Manifestations" in *Outstanding Dissertations in Linguistics Series*. Garland, NY.
Pierrehumbert, Janet
2001 "Exemplar Dynamics: Word Frequency, Lenition, and Contrast" in in Bybee, J., & Hopper, P. (eds), *Frequency Effects and the Emergence of Linguistic Structure*. John Benjamins, Amsterdam, 137–157.
Saltzman, Elliot L., & Munhall, Kevin G.
1989 "A Dynamical Approach to Gestural Patterning in Speech Production". *Ecological Psychology*, 1, 333–382.
Wright, Richard
1994 "Coda Lenition in American English Consonants: An EPG Study". *The Journal of the Acoustical Society of America*, 95, 2819–2836.