Positional Neutralization in an Exemplar Model: The Role of Unique Inflectional Bases*

Abby Kaplan
University of Utah

March 7, 2016

Exemplar models of phonology have shown promise in accounting for phenomena less easily handled by traditional generative approaches (e.g., frequency effects). However, the extent to which exemplar models can (or cannot) account for certain basic patterns has still not been fully explored. I present an exemplar model that simulates final devoicing of obstruents, leading to morphological alternations, as in German. I show that a model using only basic elements proposed in previous exemplar work incorrectly predicts that such a pattern will inevitably lead to leveling of neutralized segments throughout the paradigm. However, when relationships among paradigmatically related words are asymmetrical – such that a contrastive form may influence neutralized forms, but not vice versa – the model successfully reproduces the German pattern. I conclude that if the exemplar approach is to provide a viable account of sound patterns, it must constrain the space of possible relationships among morphologically related words.

1 Introduction

Research under the umbrella of exemplar theory (Pierrehumbert 2001, 2002; Bybee and McClelland 2005; Johnson 2006) explores the feasibility of a model in which words are stored with large amounts of phonetic detail. This approach contrasts with classic generative theories of phonology, which posit that only unpredictable information is stored in the lexicon, and that generalizations about sound patterns are located elsewhere (e.g., in rules or in constraint rankings).

Exemplar theory has shown encouraging success in its ability to provide natural accounts of phenomena that are more difficult to handle under other approaches. Exemplar models are particularly well suited to explaining frequency effects, if we assume that the frequency with which speakers encounter tokens of a given category affects the way that category is stored (Pierrehumbert 2001; Bybee and McClelland 2005; Bybee and Torres Cacoullos 2008). Furthermore, if stored exemplars are tagged with social information in addition to other types of data, then exemplar models provide a promising method of integrating sociolinguistic factors with other sources of variation (Foulkes and Docherty 2006; Johnson 2006; Pierrehumbert 2006).

However, while it is useful to know that exemplar models advance our ability to account for detailed sound patterns, it is also important to ensure that these approaches do not lose the insights

*Acknowledgements here.

of decades of prior phonological research. It is crucial, in other words, to determine whether exemplar theory is capable of modeling the kinds of phenomena that have been the bread and butter of (symbolic) generative phonology. And, indeed, there exists a body of research showing that exemplar approaches retain several basic, long-standing observations about sound patterns:

- Given an appropriate feedback loop between the storage of exemplars and the production of new ones, it is possible to model category entrenchment: phonological categories do not spread out to cover the entire range of possible values, but instead sharpen into identifiable clusters that look very much like discrete segments (Steels 1998; Pierrehumbert 2001; Oudeyer 2006; Oudeyer and Kaplan 2007; de Boer and Zuidema 2010).

- Because ambiguous tokens are more difficult for listeners to categorize, distinct categories tend to be well separated in the phonetic space, even without an explicit bias in favor of category dispersion (Steels 1998; de Boer 2000; Wedel 2004; Oudeyer and Kaplan 2007).

- Pressure for exemplars of different categories to resemble each other leads to the spread of (near-)categorical patterns across the entire lexicon, of the type traditionally modeled with phonological rules (Wedel 2004; Oudeyer 2006; Oudeyer and Kaplan 2007; Wedel 2007).

In this paper, I attempt to model positional neutralization in an exemplar-based simulation. Positional neutralization patterns, such as final devoicing (Lombardi 2001), are quite common and are a staple of introductory phonology courses; thus, any adequate model of phonology must be able to account for them. Pierrehumbert (2001) and Yu (2007) argue that exemplar frameworks can simulate loss of contrast, resulting in (near-)neutralization; however, these studies do not explicitly model positional effects. Thus, it remains to be demonstrated that an exemplar approach is capable of handling this type of pattern.

I present a working agent-based exemplar simulation of a German-style final devoicing pattern that results in paradigmatic alternations: stem-final obstruents may be either voiced or voiceless when they precede a suffix, but when they are word-final, they are uniformly voiceless. I show that a basic model combining standard design features from previous exemplar work is *not* capable of producing this type of pattern; instead of ordinary positional neutralization, such a model invariably predicts leveling, spreading neutralized segments throughout the morphological paradigm. An outcome with stable, alternating paradigms is possible only when paradigmatic relationships are asymmetrical: neutralized forms (those that lack a suffix and are thus subject to final devoicing) are under pressure to resemble non-neutralized forms (those with a suffix), but not vice versa. This model of paradigm uniformity is similar to Albright's (2002) proposal that the most informative member of an inflectional paradigm serves as its unique base.

Thus, the simulations described in this paper advance our understanding of the potential of exemplar models in two ways. First, they show that it is indeed possible to model positional neutralization (with morphological alternations) under an exemplar approach, thus putting exemplar models on firmer typological ground. Second, they show that this is possible *only under certain conditions*; relationships within morphological paradigms must be asymmetrical. As a consequence, these results constrain the space of possible exemplar models; if the exemplar approach is to be a viable model of phonological patterns more generally, we now know more about the specific form that approach must take.

```
1. Production (Agent 1)
   (a) Choose a wordform at random
   (b) Choose an exemplar at random from the wordform's cloud; copy into
       a new exemplar for production
   (c) Modify production
        i. Entrenchment:  make the production more similar to other
           exemplars
       ii. Phonetic bias:  apply pressure towards final devoicing
      iii. Noise
2. Perception (Agent 2)
   (a) Determine the lemma to which the production belongs
   (b) Store the production in the appropriate cloud, replacing a
       randomly selected old exemplar
3. Return to step 1 and switch roles
```

Figure 1: Overview of one simulation round.

2 Initial Simulations: Assembling the Necessary Ingredients

This section describes the basic design of a working exemplar simulation. Although there is great diversity in how models published in the literature are constructed, the choices made here result in a fairly standard model of interacting agents who converge on a shared lexicon that contains morphologically related words.

2.1 Basic Model Design

2.1.1 Segment Inventory

The consonant inventory of the simulations presented here consists only of stops; each stop receives a value between 0 and 100 that represents its voicing. The voicing scale is intended as a rough analogue of VOT; values above 50 are interpreted as voiceless and values below 50 as voiced. Place of articulation is not manipulated, and all stops are interpreted as labial. Thus, there are two possible consonants: {p, b}. A single vowel category is used in the simulations; I represent it here as [i].

2.1.2 Lexicon and Exemplar Clouds

A single agent's lexicon contains two lexical entries, 'Lemma 1' and 'Lemma 2'. Each lemma is realized with two different wordforms: an 'absolutive' form with no suffix, and an 'ergative' form with the suffix [-i]. (The choice of 'absolutive' and 'ergative' as names for these morphological categories is, of course, arbitrary.)

Each wordform is associated with a cloud of exemplars. At the beginning of a simulation, the cloud of each wordform is initialized with 13 randomly generated exemplars. The stems of these exemplars have the shape CVC; values for place and voicing of the consonants are chosen randomly for each exemplar, with the result that any degree of similarity among exemplars, whether within or between clouds, must emerge as a result of the agents' behavior during the simulation.

2.1.3 Linguistic Agents, Production, and Perception

This study follows previous work (Batali 1998; Steels 1998; de Boer 2000; Wedel 2004; Oudeyer 2006; Oudeyer and Kaplan 2007; de Boer and Zuidema 2010; Kirby 2010; Chirkova and Gong 2014) in simulating agents who take turns 'talking' to each other: one agent produces a token based on her stored exemplars, and the other agent perceives it and stores it in one of her own clouds. Over the course of many iterations, the result is that the agents converge on a set of similar, shared linguistic forms.

Each simulation includes two linguistic agents, each with her own lexicon. During one round of a simulation (summarized in Figure 1), one agent chooses a random wordform (e.g., the ergative of Lemma 2) and selects a random exemplar from that wordform's cloud to serve as the base of her production. That exemplar is copied into a new exemplar (which is *not* stored in the agent's own cloud) and subjected to three types of influence:

- Entrenchment (step 1(c)i, section 2.1.4): the wordform may be altered so that it more closely resembles other exemplars in the agent's cloud.
- Phonetic bias (step 1(c)ii, section 2.3): the wordform may be subject to pressure towards final devoicing.
- Noise (step 1(c)iii): each consonant in the wordform has a probability of .2 of being subjected to noise. If it is, a real number drawn from a normal distribution with mean 0 and standard deviation 10 is added to its VOT.

In step 2, the final production is supplied to the other agent, who determines the category to which the production belongs (Lemma 1 or Lemma 2); this is done in one of two ways. The first scenario, which occurs randomly with a probability of .2, models the role of context in cueing listeners to the categories they are hearing: the agent is simply 'told' the category from which the exemplar was drawn. In the second scenario, the listener compares the production to all of her stored exemplars and chooses the category whose exemplars are most similar to the production (Johnson 1997; Pierrehumbert 2001; Wedel 2004; Ettlinger 2007; Kirby 2010). The overall similarity of a given category to the production is defined as the mean of the similarity of that cloud's exemplars to the production; the similarity S_{ep} of exemplar e to production p is defined as

$$S_{ep} = \frac{1}{[\max(.1, \sum_{f,i} |d_{fi}|)]^2},$$

where d_{fi} is the distance between e and p for feature f of segment i:

$$d_{fi} = \frac{f_{ei} - f_{pi}}{\max(f) - \min(f)}.$$

The agent chooses the category with the highest mean S_{ep}. If there is a tie, the agent chooses one of the tied categories at random. Candidate categories are restricted to wordforms of the same case as the production; thus, only words of the same case are in competition with each other.

The result of using direct categorization a small percentage of the time, in addition to inference based on category similarity, is that agents converge on a set of shared labels (Wedel 2004, 165–166). In the absence of such 'deixis', agents acquire a set of shared categories, but their labels for those categories may be different (i.e., what one agent calls 'Lemma 1' may correspond to what the other calls 'Lemma 2', and vice versa).

2.1.4 Entrenchment

Previous work on exemplar-based models has shown that distinct, coherent categories can emerge in a system where exemplars are under pressure to resemble other exemplars (Pierrehumbert 2001; Wedel 2004; Oudeyer 2006; Wedel 2007, 2009; Tupper 2014). This *entrenchment* can apply during production, perception, or both. On the production side, the idea is that motor routines that have already been used are more likely to be used again (see Oudeyer 2006 for an implementation); on the perception side, the attraction of percepts to existing categories is well known as the Perceptual Magnet Effect (Kuhl and Iverson 1992). For simplicity, the simulations presented here apply entrenchment during production only.

The process of entrenchment in step 1(c)i of Figure 1 begins with collecting the set V of all values of feature f from some cloud c. We want to make the value of f in production p more similar to the values in the set V. However, there is no guarantee that V will have a simple structure; for example, the points in V might be clustered into several distinct groups. Therefore, a simple procedure, such as setting f_p to the mean of V, is not appropriate. Instead, the simulations presented here use kernel density estimation (KDE) to impose a category structure on V. The basic idea is to put a Gaussian density function at each point in V; when the Gaussians are summed, points that are close together will produce a single local maximum, while clusters of points that are distant from each other will produce a multimodal distribution. (In addition, the more points there are in a cluster, the higher that cluster's mode will be.) Modes are interpreted as basins of attraction for e; f_p can be set either to the location of the nearest mode (that is, a local maximum) or to the location of the largest mode (the global maximum). KDE was implemented using the `kde` module of the `pyqt_fit` library for Python with a bandwidth of 5; the locations of the modes were identified with a resolution of .1.

In the basic simulation, entrenchment applies on two levels: the wordform level and the segment level. At the wordform level, a segment is compared only to segments at the same position in other exemplars of the same lemma and case; for example, C_1 of an exemplar of the ergative of Lemma 2 would be compared to C_1 of other exemplars of the ergative of Lemma 2. What emerges from this wordform-level entrenchment is a coherent cloud of exemplars for each wordform in the lexicon. During wordform entrenchment, each feature f of each segment has a .7 probability of being affected; the value of f for an affected segment is set to the global maximum of f's KDE.

At the segment level, each segment is compared to all other segments of the same type; thus, a consonant is compared to all other consonants in the agent's cloud. The result of segment-level entrenchment is that the agent's exemplars tend to reuse a handful of categories, rather than (for example) allowing the cloud for each consonant in each word to cluster around an idiosyncratic VOT. Segment entrenchment applies after wordform entrenchment; each feature f of each segment has a .5 probability of being affected, and the value of f for an affected segment is set to the nearest local maximum of f's KDE.

2.1.5 Output of the Basic Model

Figure 2 illustrates a typical run of the basic simulation. Each pair of graphs shows the distribution of VOT over time in the cloud of one wordform for one agent. The first graph in each pair plots C_1, and the second graph C_3, in the $C_1V_2C_3$ stem. Note that Agent 1 (top) and Agent 2 (bottom) quickly converge to a set of shared wordforms; this effect was the same in all versions of the simulation, and so only Agent 1 is shown in subsequent figures. The dark line shows the median of each cloud; the light gray areas show the range of each cloud, and the dark gray areas the interquartile range. (Because the size of the clouds was set to 13, the median always represents an actual exemplar – not the midpoint between two exemplars – as do the endpoints of the interquartile range.)

The entrenchment described in section 2.1.4 has resulted in three clearly identifiable VOT clusters: around 5, around 55, and around 85. The 'voiceless' category at 55 is the most frequent, claiming five of the eight segments in the lexicon. These simulations commonly produce one category that is stronger (more frequent) than the other(s); this behavior is not unexpected, since a slightly more frequent category forms a deeper basin of attraction for new exemplars, which in turn strengthen the dominant category further; without some countervailing pressure, the larger category will eventually swallow up the smaller one (Pierrehumbert 2001; Tupper 2014). Some of that behavior can be seen in Figure 3, where a fourth category around 35 can be seen early in the simulation for C_1 of the absolutive of Lemma 1; after about 500 iterations, that category merges with the stronger category at 55.

Ultimately, though, in these simulations a single-category state is rendered unstable by the categorization procedure, which amplifies any small differences that arise by chance between categories until they are sufficiently distinct (Steels 1998; de Boer 2000; Wedel 2004; Oudeyer and Kaplan 2007). Figure 3 illustrates this behavior: the simulation was initialized with a single cluster of VOT values at 50, but around 1000 iterations the absolutive of Lemma 1 finds a distinct realization of C_1, eliminating the homophony between the absolutive forms of Lemma 1 and Lemma 2. (Recall that only wordforms of the same case are in competition.) Once this second category is established, C_1 of the ergative of Lemma 1 eventually follows, eliminating homophony in the ergative as well.

2.2 Paradigm Uniformity

In the simulation runs shown in Figures 2 and 3, the two agents converge on a set of distinct, stable wordforms, but wordforms of the same lemma are not necessarily related. In Figure 2, for example, the stem of Lemma 1 is [pip-] in the absolutive but [pib-] in the ergative.

Wedel (2004, 2007) proposes that the existence of (near-)categorical patterns across the lexicon can be attributed to cross-category blending: essentially, exemplars are under pressure to resemble all other exemplars, not just the ones in their own cloud. Steriade (2000) argues that morphologically related words share fine details of their phonetic implementation, and Pierrehumbert (2002) suggests that this tendency could be accounted for in an exemplar approach. The simulations here use a similar idea to encourage paradigm uniformity: after the wordform entrenchment within the production's own cloud (step 1(c)i, section 2.1.4), the same entrenchment procedure applies again, using data from the other cloud of the same lemma. In other words, a production is entrenched with respect to its own cloud, and then with respect to other clouds in the paradigm. This opposite-case

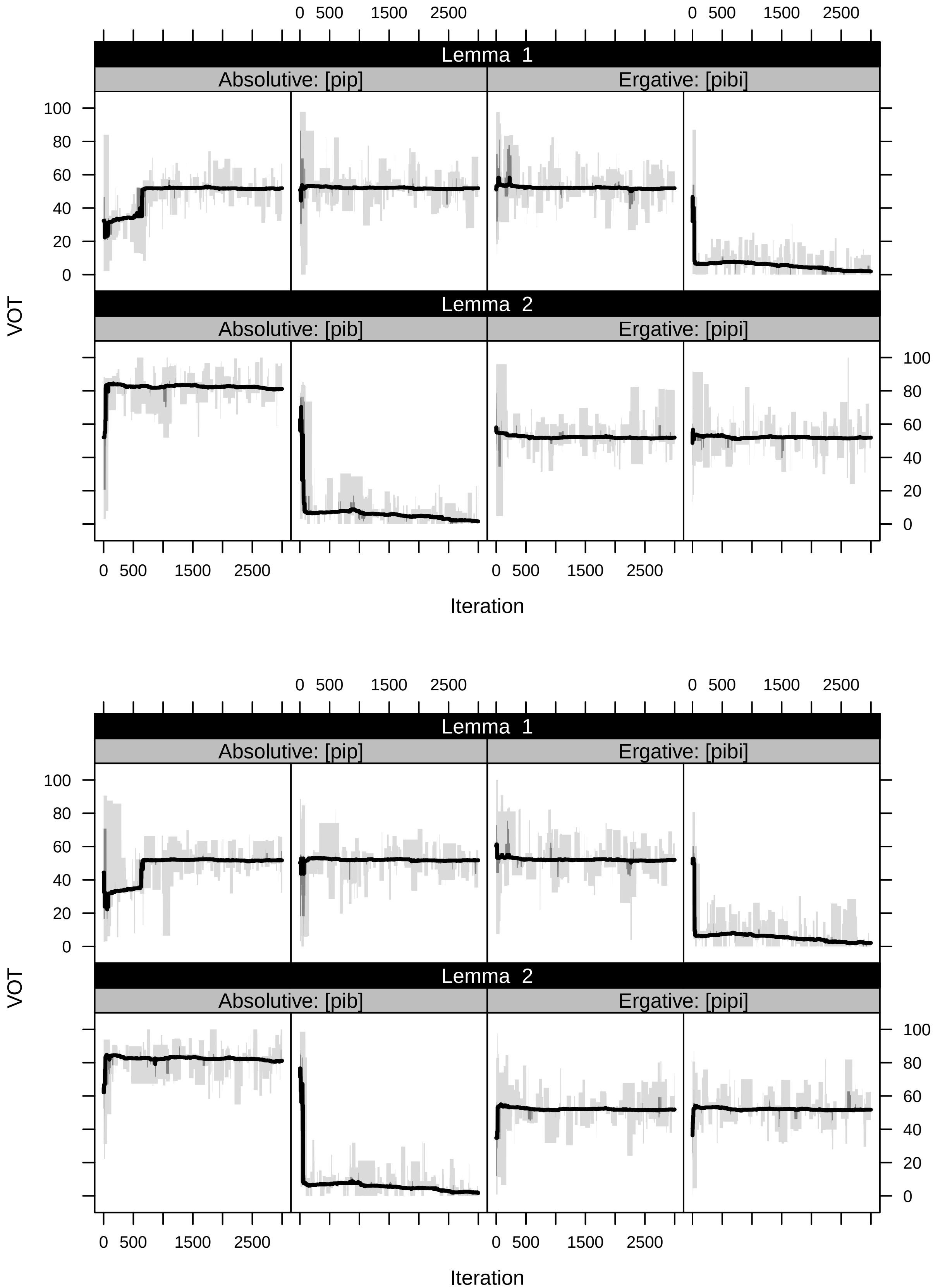

Figure 2: Results of a representative simulation run showing category entrenchment. Agent 1 is shown in the top set of graphs and Agent 2 in the bottom set.

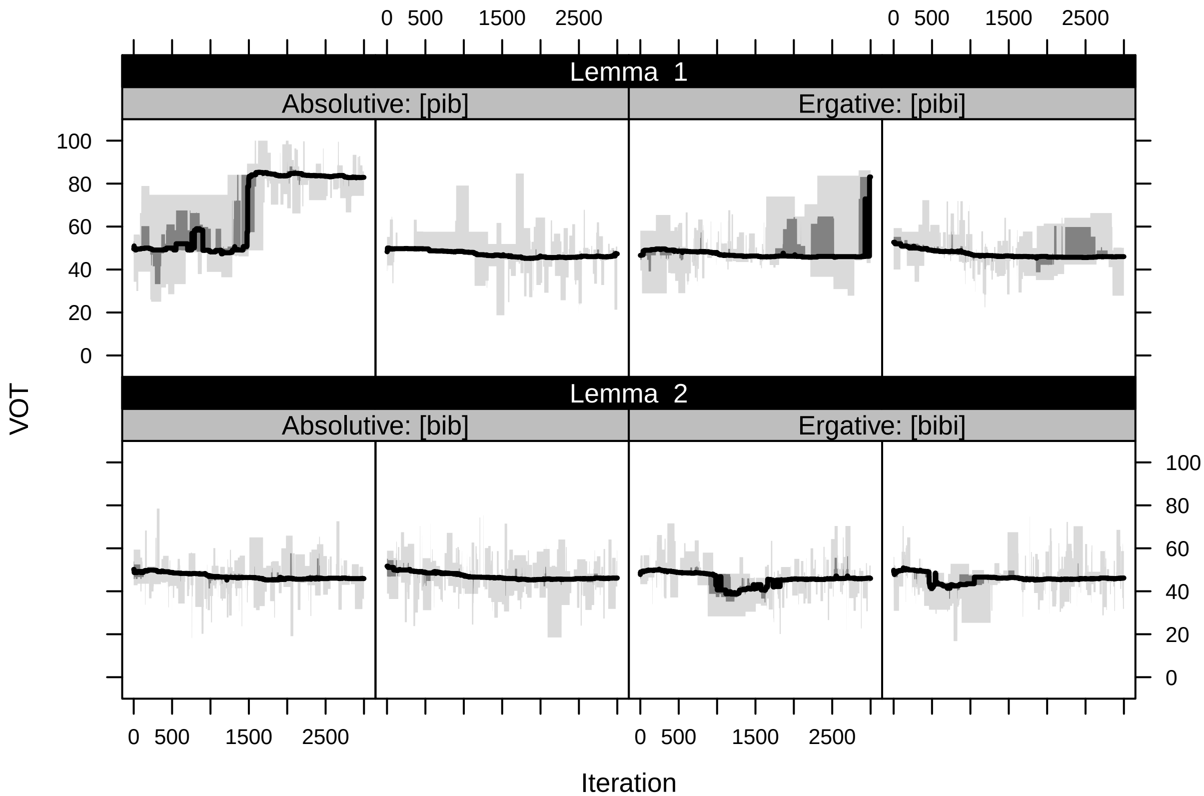

Figure 3: Results of a simulation run showing dispersion of competing wordforms.

entrenchment applies with a probability of .28 (20% of the probability of entrenchment within the production's own cloud).

Figure 4 shows the results of this added analogical pressure. The simulation initially coverges on non-uniform paradigms: in Lemma 1, the absolutive stem is [pip-] and the ergative is [pib-]; in Lemma 2, the absolutive stem is [pib-] and the ergative is [pip-]. At around 1000 iterations, C_3 of the absolutive switches categories in both lemmas, resulting in uniform stems: [pib-] for Lemma 1 and [pip-] for Lemma 2. Note that the rapid shift between [p] and [b] illustrates the categorical behavior of the system: although values of VOT are on a continuous scale, clusters of exemplars act as basins of attraction that create what is effectively a discrete set of options.

2.3 Phonetic Bias

Many implementations of exemplar models introduce a small bias in production to simulate the phonetic pressures that favor some forms over others (Pierrehumbert 2001; Wedel 2004; Oudeyer 2006; Wedel 2007; Kirby 2010). To model final devoicing, the simulations shown here introduce a bias towards voiceless final stops during production (step 1(c)ii of Figure 1). This bias has a .6 chance of applying to any given production; if it does, the final consonant of the word (if there is one; recall that ergative forms have a vowel suffix) is moved halfway to an 'ideal' VOT target of

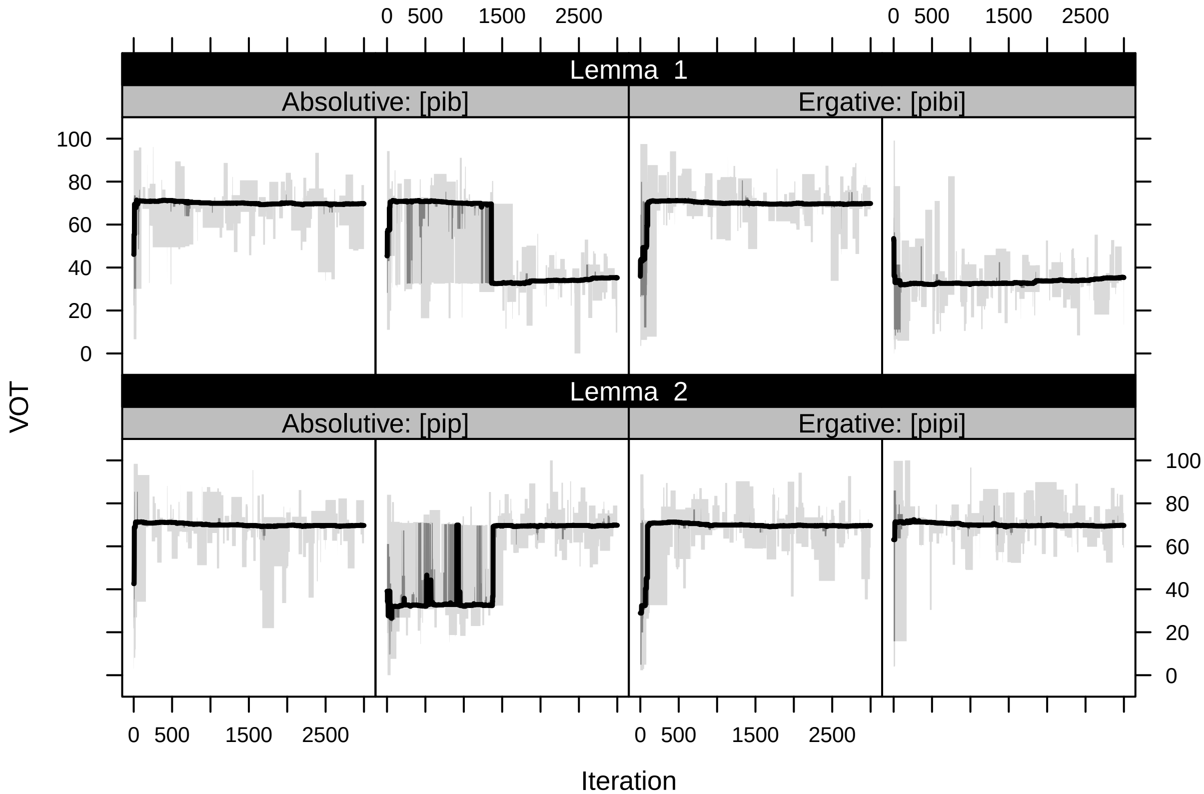

Figure 4: Results of a simulation run showing paradigm uniformity.

75.

Figure 5 shows the output of a simulation with this production bias but *without* the paradigm uniformity described in section 2.2. In this simulation, in order to show the effects of devoicing more clearly, exemplars of Lemma 1 were initialized with a consistent stem of [bib-] (both consonants with a VOT of 25), and exemplars of Lemma 2 with a stem of [pip-] (both consonants with a VOT of 75). The two lemmas are thus clearly distinguished by both C_1 and C_3; although final devoicing quickly pulls C_3 of the absolutive of Lemma 1 up to the voiceless category, the distinct C_1s mean that category contrast is not threatened.

In this simulation, the paradigm uniformity that we see everywhere except C_3 of Lemma 1 is an accident; it is the result of the fact that the stems were initialized with consistent stems. Because analogical bias was omitted, paradigm uniformity is not being actively enforced.

3 Simulations of Final Devoicing: What Doesn't Work, and What Does

With the basic ingredients of a model of final devoicing in place, we can now explore what happens when all of the necessary components are put together. In section 3.1, I show that a simple

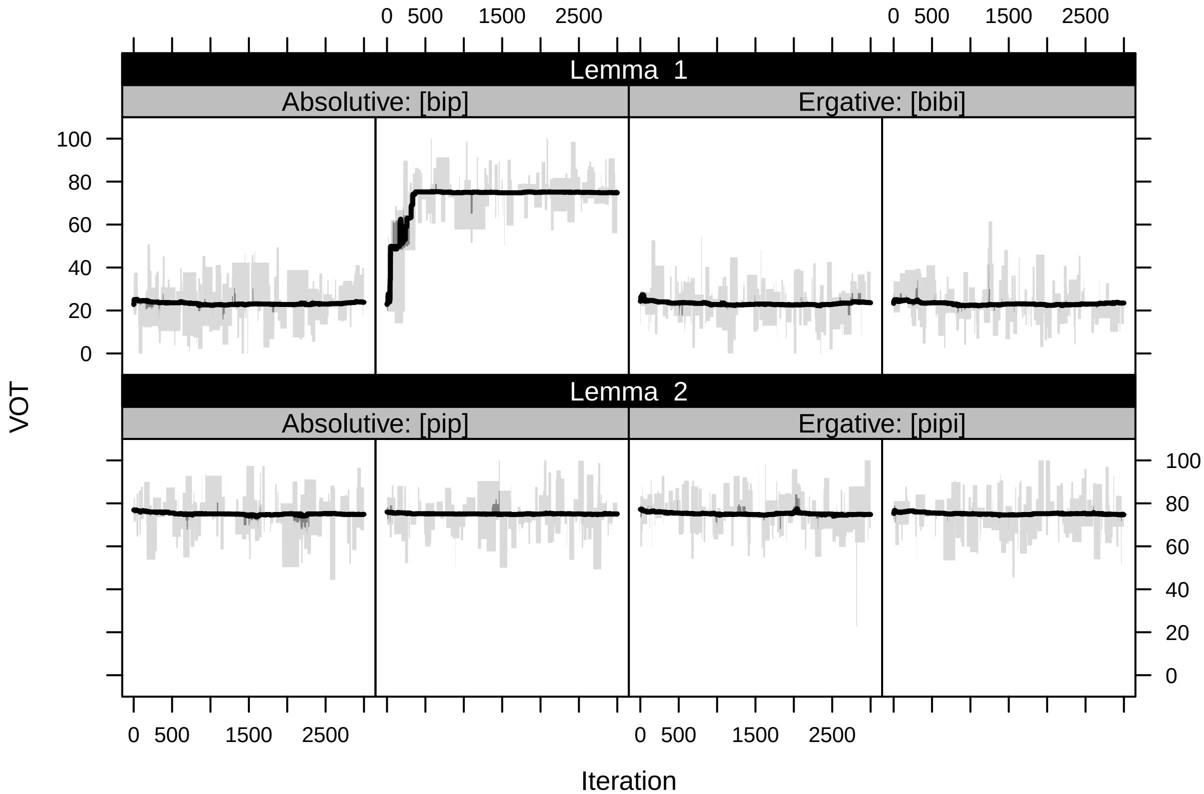

Figure 5: Results of a simulation run showing a production bias towards final devoicing.

model that combines the paradigm uniformity of section 2.2 and the phonetic bias of section 2.3 does *not* produce the classic final devoicing pattern; instead, the neutralized value in the absolutive is propagated throughout the paradigm.

3.1 Combining Paradigm Uniformity with Phonetic Bias

To model final devoicing, the most obvious approach is to combine the two ingredients introduced above: the paradigm uniformity of section 2.2 and the phonetic bias of section 2.3. Figure 6 illustrates the results of this procedure.

Unfortunately, the output of this kind of simulation is *not* the classic final devoicing pattern. Although the production bias successfully encourages devoicing of C_3 in the absolutive, the pressure for paradigm uniformity causes C_3 of the *ergative* to devoice as well. This devoicing of the ergative is due solely to analogical bias; recall that in Figure 5, where paradigm uniformity was not operative, C_3 of the ergative was not affected.

The pattern produced in Figure 6 is not entirely bad; paradigm leveling is, after all, an attested phenomenon. For example, as noted in Martin (1992) and discussed in Albright (2002, 72-73), many contrasts among final stops in Korean are neutralized in non-suffixed forms; some of these contrasts are currently undergoing leveling even in suffixed forms. Similarly, some Māori verb stems

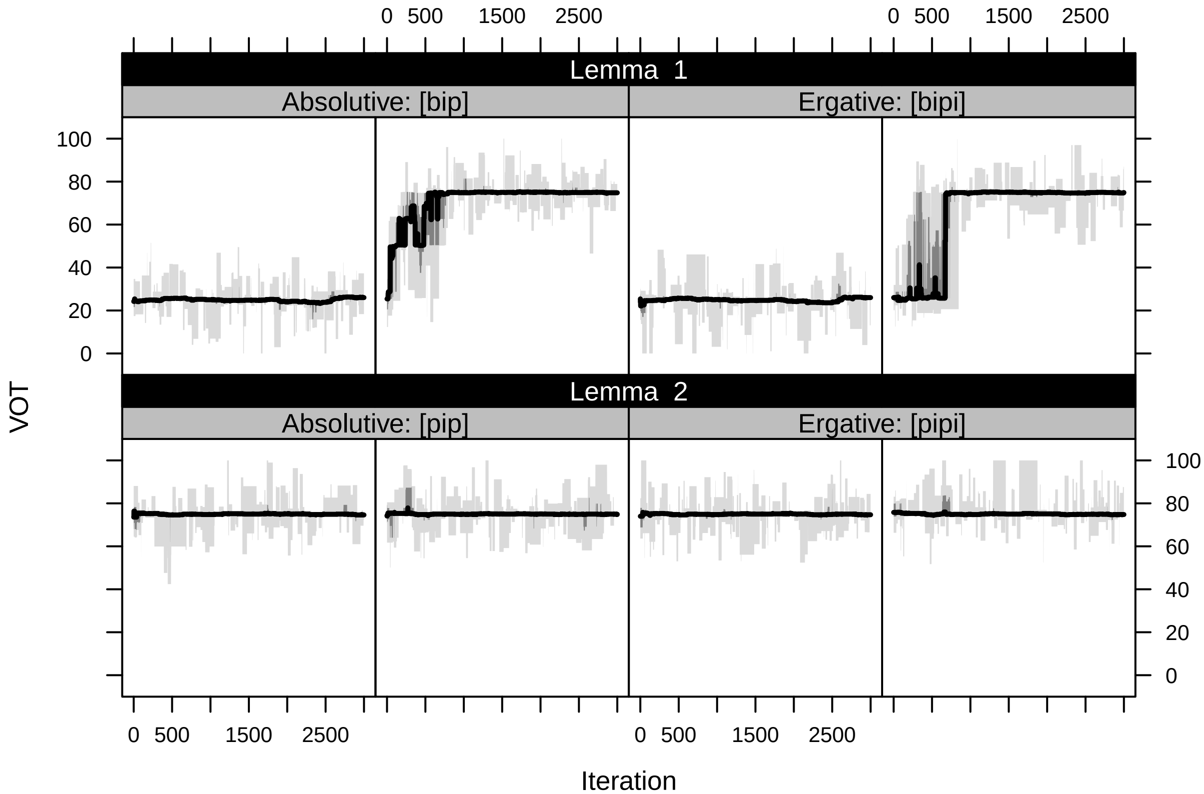

Figure 6: Results of a simulation run with both paradigm uniformity and a phonetic bias towards final devoicing.

have famously been analyzed as underlyingly consonant-final, in order to account for differences among passive forms; but Hale (1973, 1991) notes that many of these contrasts have been leveled in favor of a default *-tia* or *-a* passive suffix.

Thus, the problem with this type of simulation is not that it produces paradigm leveling, but that it cannot produce anything else: out of 10 simulation runs, all converged on a uniformly voiceless stem for Lemma 1 within 1500 iterations. To put it another way, this approach can model Korean but not German; but a viable model of phonological patterns must be able to do both.

3.2 A Non-Solution: Using the Categorization Regime to Push Wordforms Apart

The fundamental problem of the simulation illustrated in Figure 6 is that paradigm uniformity is symmetrical: the absolutive is trying to look like the ergative, and the ergative is trying to look like the absolutive. The phonetic bias is strong enough to overcome this pressure in the absolutive and cause C_3 to devoice; but once devoicing is accomplished, nothing prevents C_3 of the ergative from following suit. Leveling is, in a sense, the best solution to the simultaneous pressures of phonetic bias and paradigm uniformity.

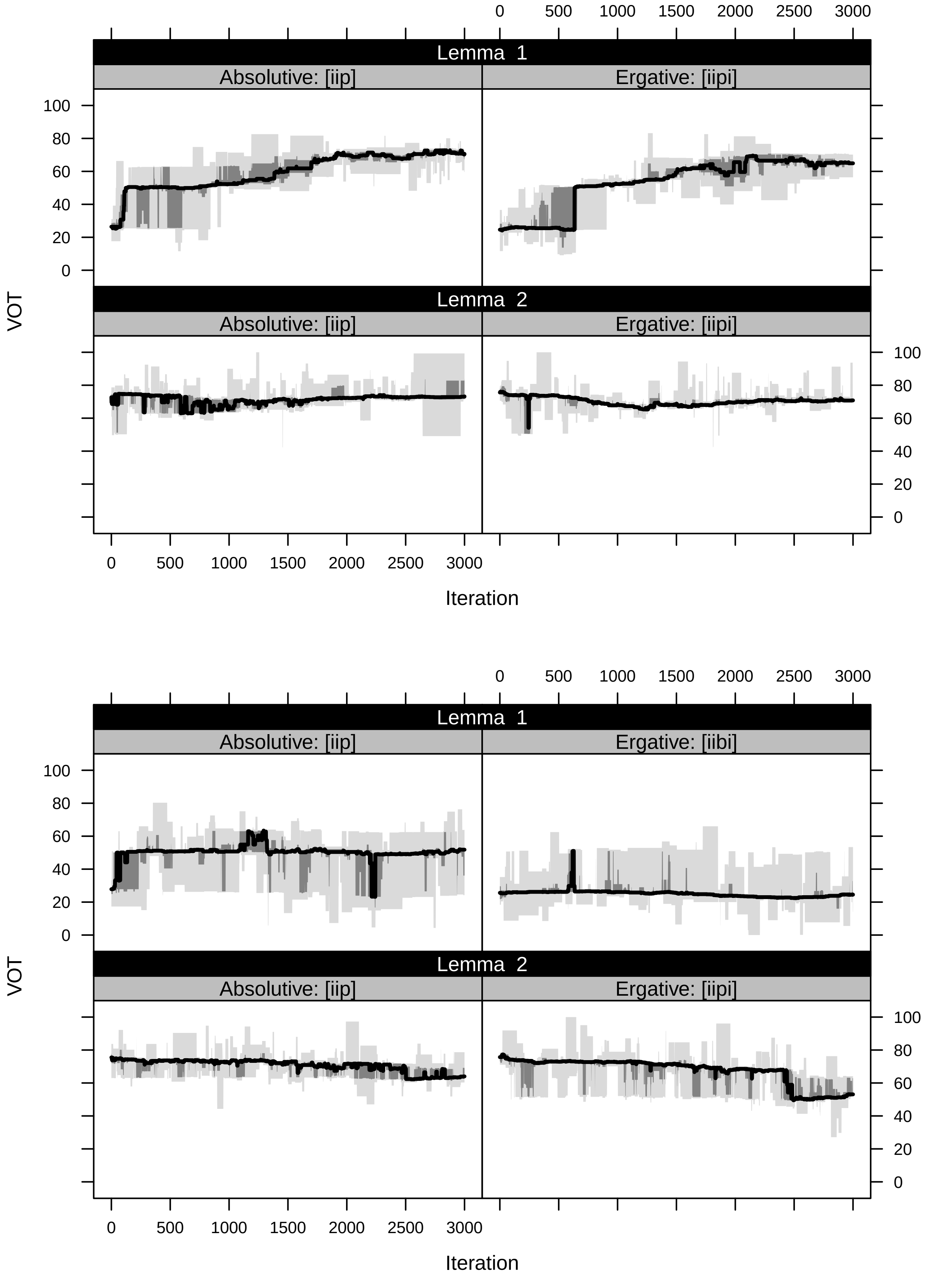

Figure 7: Results of two simulation runs with both paradigm uniformity and a phonetic bias towards final devoicing, with stems of the shape VVC.

One way to approach this problem would be to try to find a additional pressure for the ergative forms that could counteract the effects of paradigm uniformity; the categorization procedure of step 2 of Figure 1 could, in principle, supply this pressure. As noted above, the categorization regime of these simulations forces wordforms to compete with each other and thus encourages different wordforms to be phonologically distinct. In Figure 6, the ergative of Lemma 1 is free to become [bipi] (instead of remaining [bibi]) because C_1 distinguishes the word from [pipi] of Lemma 2. It is worth exploring whether the leveling observed in Figure 6 could be prevented by the threat of homophony if final devoicing spreads from the absolutive to the ergative.

In the simulation runs depicted in Figure 7, stems were given the shape VVC; the lone C_3 of the stem is the only segment capable of distinguishing between Lemma 1 and Lemma 2. Thus, the fact that the categorization procedure encourages distinct categories exerts a pressure on wordforms counter to the phonetic bias, which nudges all final Cs towards a value of 75. Under these conditions, the simulation produced two basic types of outcomes. The first, shown in the top half of Figure 7, was more common, occurring in 8 out of 10 simulation runs. Here, the final consonant of the absolutive form of Lemma 1 (initialized as [iib]) devoices, but settles on a value for VOT short of the 75 encouraged by the phonetic bias – apparently because of competition from the [iip] of Lemma 2. The ergative of Lemma 1 is initially unchanged, but after about 500 iterations it transitions quickly to the VOT of the absolutive, thus exhibiting the same paradigm leveling seen in simulation runs with CVC stems. After settling on a uniform [iip-] stem, both members of Lemma 1's paradigm drift towards a VOT of 75, but never quite get there (again, presumably because of competition from Lemma 2).

The second type of outcome, shown in the bottom half of Figure 7, seems to have settled on a stable alternation in Lemma 1: the absolutive is [iip], while the ergative is [iibi]. (Note that the range of exemplars within each category suggests a continuing pressure for the two wordforms to resemble each other.) The final consonant of [iip] for Lemma 1 has stabilized at a VOT much lower than 75; it would not be unreasonable to interpret this as a new, third voicing category, especially since C_3 of the ergative of Lemma 2 switches rapidly to this category around 2500 iterations.

Although this second outcome produces a stable alternation, it nevertheless fails to provide us with a satisfying model of positional neutralization. Crucially, the stable, non-leveled paradigm seen in the bottom half of Figure 7 is possible *only* for wordforms that are in danger of homophony – in other words, the ergative of Lemma 1 can remain [iibi] only because Lemma 2 has [iipi].

Figure 8 illustrates the problem more concretely. This simulation has an expanded lexicon: two VVC stems, and one CVC stem initialized as [bib-]. As before, the two VVC lemmas can be distinguished only by C_3, and the simulation sometimes settles on a slightly devoiced C_3 for the absolutive of Lemma 1 and a fully voiced C_3 for the ergative. (This result occurred in 4 out of 10 simulation runs.) By contrast, Lemma 3 is clearly distinguished from the other two lemmas by C_1, and thus C_3 of the ergative (as well as the absolutive) is free to devoice with no danger of homophony. The paradigm of Lemma 3 was leveled in all 10 simulation runs.

Thus, this type of simulation predicts that in a language with positional neutralization, neutralization in one form will invariably spread to other members of the same paradigm *unless* those other words are in danger of becoming homophones. But the facts of attested positional neutralization patterns are clearly otherwise: in German, for example, of the 2,053 words with underlying final voiced stops listed in CELEX (Baayen et al. 1995, using orthographic representations to infer underlying forms), a mere 41 (1.99%) are part of such a minimal pair. Natural languages do not depend on potential homophones to maintain alternating paradigms.

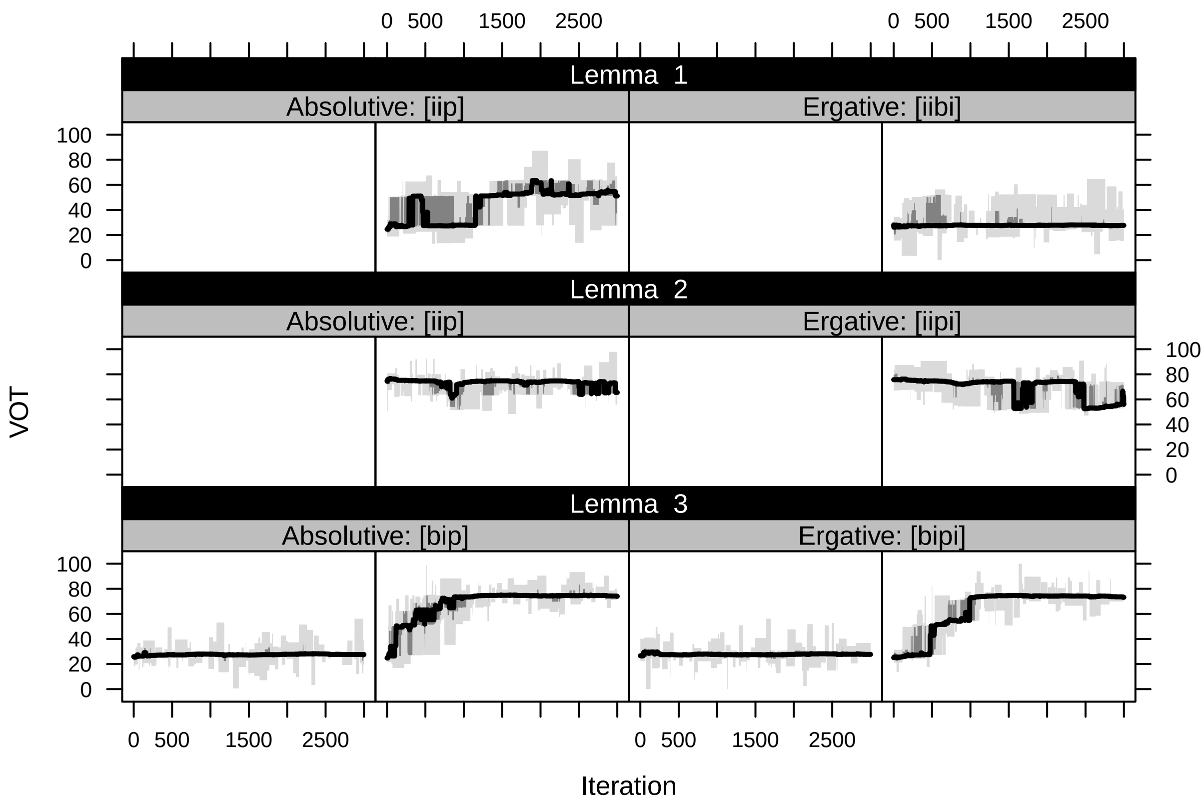

Figure 8: Results of two simulation runs with both paradigm uniformity and a phonetic bias towards final devoicing, with two stems of the shape VVC and one of the shape CVC.

3.3 A Good Solution: Unique Inflectional Bases

We have seen that competition among wordforms can solve the problem of leveling on a word-by-word basis, but not for the system as a whole. Nor can we introduce a new phonetic bias that encourages stops to be voiced between vowels – although this is a typologically well-motivated tendency, it does not accomplish what we need because it does not encourage the maintenance of a *contrast* in C_3 of the ergative; it encourages voicing in all forms. The result is that C_3 remains voiced in the ergative of Lemma 1 ([bibi]), but C_3 becomes voiced in the ergative of Lemma 2 as well (resulting in [pibi]).

There is, then, no plausible candidate for a counter-pressure that would encourage C_3 of the ergative to remain voiced in Lemma 1. If we cannot introduce a new bias to counteract the effects of paradigm uniformity in Figure 6, the only alternative is to remove the existing bias – that is, the ergative form of Lemma 1 must not be attracted to the absolutive form. But the absolutive should still be attracted to the ergative, or else there will be no paradigm uniformity at all. What we need, then, is *asymmetrical* paradigm uniformity: only some forms in the paradigm can influence others.

Albright (2002, 2008, 2010) and Albright and Kang (2009) have argued for precisely this kind of asymmetry. The proposal is that inflectional paradigms depend on a unique base from which all

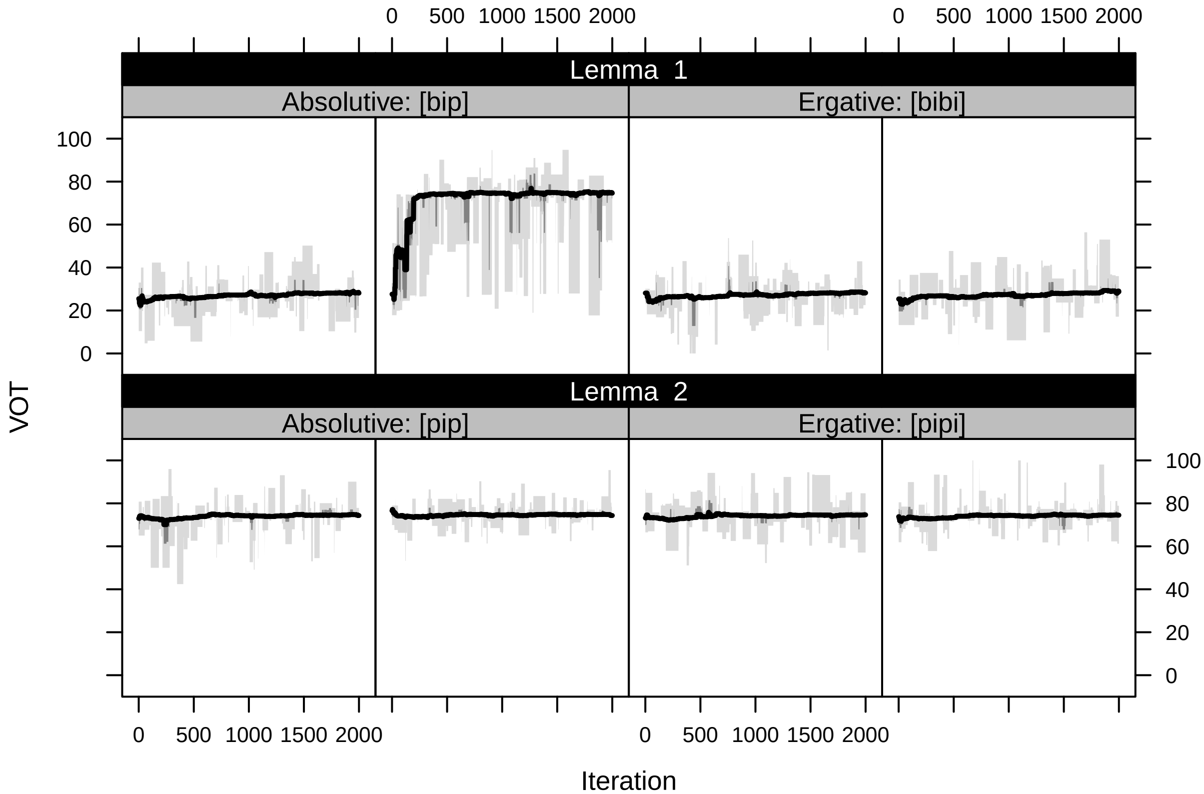

Figure 9: Results of a simulation run with asymmetrical paradigm uniformity.

other forms are derived, and that the most informative member of the paradigm is systematically chosen to be this base. As applied to the simulations here, the spirit of this proposal is exactly what is needed to avoid leveling: After phonetic bias has produced final devoicing, C_3s of absolutive forms are no longer informative (because they are all voiceless). But C_3 still contrasts in ergative forms; thus, the ergative is clearly the most informative member of the paradigm.

Although Albright's proposal is that a single form serves as the base of the paradigm, 'basehood' in the simulations presented here was determined on a segment-by-segment basis (indeed, on a feature-by-feature basis). For a given feature f_s of segment s, the informativity of f for each case c was operationalized as the performance of a classifier that attempted to distinguish between all the exemplars of case c on the basis of f_s alone. (The classifier used the same similarity-based criterion described in section 2.1.3 above.) Under this procedure, after C_3 has been devoiced in the absolutive of Lemma 1, the VOT feature of C_3 is not useful for distinguishing between the absolutive forms of Lemma 1 and Lemma 2: they both end in voiceless stops, and the classifier will perform essentially at chance. But for the ergative, as long as C_3 of Lemma 1 remains voiced, the classifier will perform quite well.

Informativity was applied to paradigm uniformity using a 'winner-take-all' rule: the most informative case c was identified as the unique base of the paradigm; during paradigm-driven entrenchment (step 1(c)i, section 2.2), a production was entrenched relative to the exemplars of c, but not of any other cases. Informativity (and therefore basehood) was calculated anew with each production. Thus, as soon as final devoicing has rendered voicing of C_3 less reliable for categorization in the absolutive than in the ergative, the ergative is established as the most informative case, and ergative forms are no longer influenced by absolutive forms.

Figure 9 shows that this modification succeeds in producing a system with final devoicing and robust paradigm alternations. I conclude that in a model where similarities among paradigm members are enforced via surface-to-surface relations (but *not* via reference to an abstract underlying stem), the only viable approach is one in which informative members of the paradigm can influence non-informative members, but not vice versa.

Note also that although the phonetic bias towards devoicing of C_3 in the absolutive is strong enough to overcome the pressure for the absolutive to resemble the ergative, the effects of paradigm uniformity can still be observed. Specifically, the lower bound of the range of exemplars for C_3 of [bip] in Figure 9 is lower than in Figure 5, where paradigm uniformity is not in effect. The cloud for this segment, then, contains more exemplars with relatively small VOTs than we would otherwise expect. This effect looks very much like incomplete neutralization (Port and O'Dell 1985; Slowiaczek and Dinnsen 1985; Port and Crawford 1989; Charles-Luce 1993; Gerfen and Hall 2001; Warner et al. 2004; Simonet et al. 2008; Kharlamov 2014) and is consistent with the hypothesis that incomplete neutralization is itself a kind of sub-phonemic paradigm uniformity (Ernestus and Baayen 2007; Braver 2013, although note that Ernestus and Baayen explicitly propose a symmetrical paradigm uniformity of the type rejected here).

3.4 Asymmetrical Paradigm Uniformity without Unique Inflectional Bases

We will consider one final modification to the implementation of paradigm uniformity. Following Albright (2002), the algorithm described in section 3.3 uses a 'winner-take-all' rule: the most informative member of the paradigm can influence others; less informative members have no effect whatsoever. But other effects of differences in informativity are logically possible; for example, suppose that the influence a cell has on other cells in the paradigm is proportional to its informativity (so that less informative members of the paradigm are penalized but are not rendered completely inert). Does such an approach still succeed in modeling positional neutralization?

The answer is no. As shown in Figure 10, this 'proportional' implementation of informativity is no more effective at preventing leveling than a simulation with symmetrical paradigm uniformity. The problem is that the slight difference between the devoiced C_3 of [bip] and the non-alternating C_3 of [pip] (what I identified in section 3.3 as incomplete neutralization) is enough to allow even the mostly neutralized absolutive form to pull the ergative form toward itself. Out of 10 simulation runs of this type, not one produced the paradigm alternation seen in Figure 9. I conclude that an adequate model of positional neutralization that relies on surface-to-surface relationships to enforce paradigm uniformity must employ *unique* inflectional bases.

4 Discussion and Conclusion

Exemplar approaches have shown great promise in accounting for properties of sound patterns that are more difficult to capture in traditional formal generative theories. However, it is important that these models not abandon the insights of previous phonological research. Specifically, it is a

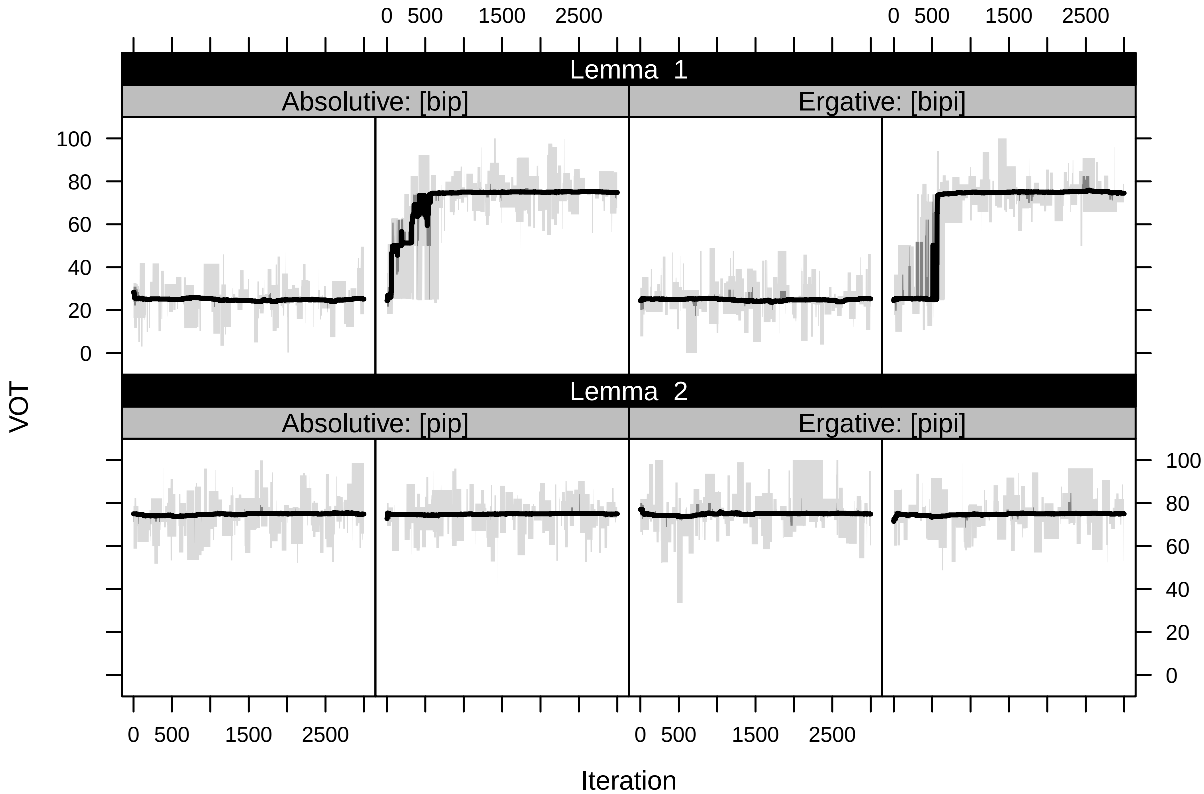

Figure 10: Results of a simulation run with asymmetrical paradigm uniformity and no 'winner-take-all' rule.

basic observation that languages can exhibit positional neutralization (where a contrast is realized in some positions but not others) and that this neutralization can lead to alternations within morphological paradigms. I have shown that in order for an exemplar model to account for this kind of pattern, in must be implemented in a specific way: in particular, it must make relationships among members of a morphological paradigm asymmetrical, such that the most informative member of the paradigm can influence others, but not vice versa. Converging evidence for this approach comes from the fact that such unique basehood has been previously proposed to account for typological phenomena (Albright 2002), and from the fact that the model predicts the well-known phenomenon of incomplete neutralization (Port and O'Dell 1985).

Work on exemplar approaches to phonology is still in its early stages, and there is little consensus on many of the specifics of how these models should work: although there is broad agreement on some of the basic ingredients, there is also a tendency for each researcher or research group to build a new model from scratch. The present study takes a step towards constraining the space of possible implementations by showing what is *not* successful in modeling natural language, as well as what is.

References

Adam Albright. *The Identification of Bases in Morphological Paradigms.* PhD thesis, University of California, Los Angeles, Los Angeles, CA, 2002.

Adam Albright. Explaining universal tendencies and language particulars in analogical change. In Jeff Good, editor, *Linguistic Universals and Language Change*, chapter 7, pages 144–181. Oxford University Press, Oxford, 2008.

Adam Albright. Base-driven leveling in Yiddish verb paradigms. *Natural Language and Linguistic Theory*, 28:475–537, 2010.

Adam Albright and Yoonjung Kang. Predicting innovative alternations in Korean verb paradigms. In *Current Issues in Unity and Diversity of Languages: Collection of the Papers Selected from the CIL 18, Held at Korea University in Seoul*, pages 1–20, Seoul, Korea, 2009. Linguistic Society of Korea.

R. Harald Baayen, Richard Piepenbrock, and H. van Rijn. The CELEX lexical database. Philadelphia, PA: Linguistic Data Consortium, 1995. Release 2 (CD-ROM).

John Batali. Computational simulations of the emergence of grammar. In Hurford et al. (1998), chapter 24, pages 405–426.

Aaron Braver. *Degrees of Incompleteness in Neutralization: Paradigm Uniformity in a Phonetics with Weighted Constraints.* PhD thesis, Rutgers University, 2013.

Joan Bybee and James L. McClelland. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, 22:381–410, 2005.

Joan Bybee and Rena Torres Cacoullos. Phonological and grammatical variation in exemplar models. *Studies in Hispanic and Lusophone Linguistics*, 1(2):399–413, 2008.

J. Charles-Luce. The effects of semantic context on voicing neutralization. *Phonetica*, 50(1):28–43, 1993.

Katia Chirkova and Tao Gong. Simulating vowel chain shift in Xumi. *Lingua*, 152:65–80, 2014.

Bart de Boer. Self-organization in vowel systems. *Journal of Phonetics*, 28(4):441–465, 2000.

Bart de Boer and Willem Zuidema. Multi-agent simulations of the evolution of combinatorial phonology. *Adaptive Behavior*, 18(2):141–154, 2010.

Mirjam Ernestus and R. Harald Baayen. Intraparadigmatic effects on the perception of voice. In Jeroen van de Weijer and Erik Jan van der Torre, editors, *Voicing in Dutch: (De)voicing – Phonology, Phonetics, and Psycholinguistics*, number 286 in Current Issues in Linguistic Theory, pages 151–172. John Benjamins Publishing Company, Amsterdam, 2007.

Marc Ettlinger. Shifting categories: An exemplar-based computational model of chain shifts. *UC Berkeley Phonology Lab Annual Report*, pages 177–182, 2007.

Paul Foulkes and Gerard Docherty. The social life of phonetics and phonology. *Journal of Phonetics*, 34(4):409–438, 2006.

Chip Gerfen and Kathleen Hall. Coda aspiration and incomplete neutralization in Eastern Andalusian Spanish. Ms., University of North Carolina at Chapel Hill, October 2001.

Ken Hale. Deep-surface canonical disparities in relation to analysis and change: An australian example. In *Current Trends in Linguistics*, volume II, pages 401–458. Mouton de Gruyter, Berlin, 1973.

Ken Hale. Remarks on g. sanders "Levelling in the history of Polynesian passive formations". *Journal of the Polynesian Society*, 100(1):99–101, 1991.

James R. Hurford, Michael Studdert-Kennedy, and Chris Knight, editors. *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press, Cambridge, 1998.

Keith Johnson. Speech perception without speaker normalization: An exemplar model. In Keith Johnson and John W. Mullenix, editors, *Talker Variability in Speech Processing*, chapter 8, pages 145–165. Academic Press, New York, NY, 1997.

Keith Johnson. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, 34:485–499, 2006.

Viktor Kharlamov. Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics*, 43:47–56, 2014.

James P. Kirby. *Cue Selection and Category Restructuring in Sound Change*. PhD thesis, University of Chicago, 2010.

Patricia K. Kuhl and Paul Iverson. Linguistic experience and the "perceptual magnet effect". In Winifred Strange, editor, *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, chapter 4, pages 121–154. York Press, Baltimore, MD, 1992.

Linda Lombardi. Why place and voice are different: Constraint-specific alternations in Optimality Theory. In Linda Lombardi, editor, *Segmental Phonology in Optimality Theory: Constraints and Representations*, chapter 1, pages 13–45. Cambridge University Press, Cambridge, 2001.

Samuel E. Martin. *A Reference Grammar of Korean*. Charles E. Tuttle, Rutland, VT, 1992.

Pierre-Yves Oudeyer. *Self-Organization in the Evolution of Speech*. Studies in the Evolution of Language. Oxford University Press, Oxford, 2006.

Pierre-Yves Oudeyer and Frédéric Kaplan. Language evolution as a Darwinian process: Computational studies. *Cognitive Processing*, 8(1):21–35, 2007.

Janet B. Pierrehumbert. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee and Paul J. Hopper, editors, *Frequency and the Emergence of Linguistic Structure*, number 45 in Typological Studies in Language, pages 137–157. John Benjamins Publishing Company, Amsterdam, 2001.

Janet B. Pierrehumbert. Word-specific phonetics. In Carlos Gussenhoven and Natasha Warner, editors, *Laboratory Phonology*, volume VII of *Phonology and Phonetics*, pages 101–139. Mouton de Gruyter, Berlin, 2002.

Janet B. Pierrehumbert. The next toolkit. *Journal of Phonetics*, 34(4):516–530, 2006.

Robert Port and Penny Crawford. Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17(4):257–282, 1989.

Robert Port and Michael O'Dell. Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13:455–471, 1985.

Miquel Simonet, Marcos Rohena-Madrazo, and Mercedes Paz. Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In Laura Colantoni and Jeffrey Steele, editors, *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, pages 72–86, Somerville, MA, 2008. Cascadilla Proceedings Project.

Louisa M. Slowiaczek and Daniel A. Dinnsen. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13(3):325–341, 1985.

Luc Steels. Synthesizing the origins of language and meaning using coevolution, self-organization and level formation. In Hurford et al. (1998), chapter 23, pages 384–404.

Donca Steriade. Paradigm uniformity and the phonetics-phonology boundary. In Michael B. Broe and Janet B. Pierrehumbert, editors, *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, chapter 22, pages 313–334. Cambridge University Press, Cambridge, 2000.

P. F. Tupper. Exemplar dynamics and sound merger in language. Ms., Simon Fraser University, December 2014.

Natasha Warner, Allard Jongman, Joan Sereno, and Rachèl Kemps. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32(2):251–276, 2004.

Andrew Wedel. *Self-Organization and Categorical Behavior in Phonology.* PhD thesis, University of California, Santa Cruz, 2004.

Andrew Wedel. Resolving pattern conflict: Variation and selection in phonology and morphology. In James P. Blevins and Juliette Blevins, editors, *Analogy in Grammar*, chapter 4, pages 83–100. Oxford University Press, Oxford, 2009.

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

Andrew B. Wedel. Feedback and regularity in the lexicon. *Phonology*, 24(1):147–185, 2007.

Alan C. L. Yu. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology*, 24:187–214, 2007.