

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

Abby Kaplan  
*University of Utah*

## 1 Introduction

Recent years have seen a growing body of literature on models of phonology that can be classed under the umbrella of ‘exemplar theory’. What these models have in common is the hypothesis that linguistic units (words, segments, and so on) are represented by clouds of remembered exemplars rather than by abstract templates. Thus, in contrast to traditional approaches that seek to minimize the amount of information that must be stored, exemplar models take a ‘memory is cheap’ approach and posit a huge number of richly detailed traces of linguistic units.

Much of the impetus for this approach is the desire to account for phenomena that are left unexplained by traditional generative models, such as frequency effects or gradient well-formedness judgments (Bybee & McClelland, 2005). But it turns out that this approach is *also* capable of modeling phonological phenomena that were once thought to require an abstract, symbolic system; a number of studies in the literature are dedicated to showing how this is so:

- Pierrehumbert (2001) demonstrates that exemplar clouds coalesce into relatively stable categories as long as an entrenchment bias (a tendency for exemplars to move towards the center of gravity of their category’s cloud) is present in either production or perception.
- Wedel (2004) shows how a similar feedback loop, wherein speakers’ productions are biased to resemble existing exemplars, and those productions are subsequently stored as new exemplars, predicts the existence of (near-)categorical phonotactic patterns in a language’s lexicon.
- Oudeyer (2006) presents a simulation of interacting ‘babbling’ agents who evolve a combinatorial phonology that reuses a small set of segments – essentially, duality of patterning.
- Ettliger (2007) shows how the natural behavior of adjacent clouds of exemplars provides a non-teleological explanation of chain shifts.

The present paper continues this line of research by exploring whether it is possible for an exemplar-based approach to model positional neutralization. Studies like these are important for two reasons. First, they allow us to evaluate whether the insights of traditional generative phonology are preserved in exemplar models: the latter theory is much more attractive if we know that we can adopt it without losing what we already know. Second, the literature contains surprisingly few actual working models in the exemplar family (at least in linguistics); since the behavior of these systems is not always easy to predict, especially as they grow more complex, it is important to continue to test whether these models actually do what we think they can do.

As it turns out, the simulations presented below show that a basic exemplar model is *not* capable of replicating a positional neutralization pattern, one with alternations among morphologically related forms; instead, the basic model predicts is that the variant that is the result of neutralization will be propagated throughout the paradigm. However, positional neutralization *can* be modeled if we introduce asymmetric paradigm uniformity (Albright, 2002), such that non-neutralized members of the paradigm are under no

---

\* Thanks to Andy Wedel, Aaron Kaplan, and audiences at the Old World Conference in Phonology, the University of Utah, and AMP 2014 for comments at various stages of this work. The Java code used in the simulations described in this paper is available from the author on request.

pressure to resemble neutralized members. This result puts a non-trivial restriction on exemplar theory: to account for attested phonological patterns in natural language, an adequate theory must have asymmetric relationships within inflectional paradigms.

## 2 Neutralization and Surface-to-Surface Similarity in Exemplar Models

Although positional neutralization has not, to my knowledge, been modeled in the existing literature, the components of it have been. Context-free neutralization is addressed by Pierrehumbert (2001), who simulates a marked category collapsing into an unmarked one; by Tupper (2014), who explores which exemplar categorization algorithms prevent adjacent clouds from collapsing; and by Yu (2007), who shows that tone merger in Cantonese as a result of *pinjam* alternations is not complete, a not unexpected result under an exemplar approach.<sup>1</sup> In the simulations below, neutralization of voicing distinctions word-finally is modeled using a production bias against final voiced consonants.

To account for alternations among morphologically related forms, traditional generative phonology posits abstract underlying forms; members of a morphological paradigm are similar to each other because they are derived from the same UR and differ from it only when necessary (when changed by a phonological rule, or when faithfulness to the UR is overridden by some higher-ranking constraint). In exemplar theory, or any other approach without abstract underlying representations, similarities between related forms require a positive mechanism that actively encourages related surface forms to look alike.<sup>2</sup> Some existing models already include components that promote similarity between distinct exemplar clouds; Wedel (2004), as noted above, uses a bias towards patterns present across the entire exemplar map as part of a feedback loop that results in near-categorical phonotactic patterns. In the simulations below, I take the straightforward step of positing that a bias of this sort is particularly strong between morphologically related words.

As suggested in the prior literature, and confirmed by the simulations below, the exemplar models described in prior work are already amply suited to modeling neutralization and paradigm uniformity separately. The question at issue in this paper is what happens when those two things are combined in a single model: is the result a set of morphologically related forms that differ only when required by the bias towards final devoicing? As we will see, the answer is no; instead of a classic positional neutralization pattern, the result of a simple model is that neutralization in one form spreads through the entire paradigm.

## 3 Simulations

To explore whether it is possible to model positional neutralization in an exemplar-based approach, I conducted a series of simulations on a small artificial language mimicking final devoicing, the classic example of positional neutralization (Lombardi, 2001). All of the simulations presented here involved a single agent with a 10-word lexicon. Each word had the shape CVC and was realized in two cases: the ‘nominative’, with no suffix; and the ‘accusative’, with the suffix [-i]. Vowels were drawn from the set {i, a, u} and consonants from the set {p, t, k, b, d, g}. Consonants were represented with two features: place (labial, coronal, or dorsal) and voicing (voiced or voiceless). Vowels were represented atomically, not decomposed into features such as height or backness.

During each simulation, the agent repeatedly produced exemplars and then stored those productions back in his own map. Productions never perfectly matched existing exemplars; they were subject to analogical pressure from other members of the same exemplar cloud and to random noise. In addition, in some simulations, productions were subject to additional analogical pressure from inflectionally related words (i.e., paradigm uniformity) or to a bias towards devoicing of final consonants. As a result, exemplar clouds drifted over time and revealed the self-organizing tendencies of the system. The production-perception loop proceeded as follows:

<sup>1</sup> The phenomena discussed by Yu (2007) and Pierrehumbert (2001) do not actually involve context-free neutralization; *pinjam* is morphologically conditioned, and Pierrehumbert gives final devoicing as a real-world example of the kind of neutralization she models. But in both of these studies, the authors focus on the neutralizing behavior of the relevant patterns and leave the positional facts aside.

<sup>2</sup> Positive identity mechanisms like these are of course not exclusive to theories that lack URs; phonologists working in Optimality Theory, for example, have proposed constraints that penalize morphologically related forms for differing from one another (Benua, 1997; McCarthy, 2005).

1. Create a seed lexicon
  - (a) Each case for each word was initialized with a single exemplar consisting of a randomly chosen CVC sequence (plus [-i] in the accusative).
  - (b) The initial nominative and accusative seeds of a given word were *not* related. Thus, any similarity between the nominative and accusative forms of the same word was the result of the bias towards paradigm uniformity in step 3.
2. Choose one exemplar for production
  - (a) A single wordform (either the nominative or accusative form of some word) was randomly chosen for production.
  - (b) A single exemplar from that wordform's cloud was randomly chosen at the base of the production, and copied into a new exemplar.
3. Apply analogical pressure
  - (a) To model category entrenchment (Pierrehumbert, 2001), the production was probabilistically altered to resemble other exemplars in its cloud.
    - i. For each feature of each segment at each position (e.g., voicing of C1), the most frequent value of that feature in that position in the exemplar cloud was determined.
    - ii. (In the event of a tie, one of the most common feature values was chosen at random.)
    - iii. In the production, each feature of each segment at each position had a 70% chance of changing to match the most frequent value of that feature in that position.
  - (b) In the B and D simulations, paradigm uniformity was modeled by applying the same analogical pressure from the cloud of exemplars from the other case of the same word.
4. Add bias and noise
  - (a) To simulate production noise, each feature of each segment had a 15% chance of switching to a random value.
  - (b) In the C and D simulations, which included a production bias towards final devoicing, word-final consonants had a 20% chance of becoming voiceless.
  - (c) The [-i] suffix of the accusative was never subject to bias or noise.
5. Categorize the new production
  - (a) The production was compared to each exemplar cloud (of the same case).
    - i. For each cloud, for each feature of each segment at each position, the most frequent value of that feature in that position in the exemplar cloud was determined – cf. step 3(a)i.
    - ii. For each feature of each segment at each position, one 'match' was registered if the feature of the production matched the most frequent feature of the exemplar cloud.
  - (b) The production was categorized in the exemplar cloud with the most matches. (In the event of a tie, one of the clouds with the most matches was chosen at random.)
6. Store the new production
  - (a) The production was added to the cloud of its chosen category.
  - (b) If that cloud already contained 5 exemplars, one existing exemplar was replaced at random.
7. Go back to step 2 and repeat 10,000 times

A Simulations			B Simulations			C Simulations		
<b>C1</b>			<b>C1</b>			<b>C1</b>		
NOM			NOM			NOM		
ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]
[-voi]	250	265	[-voi]	502	6	[-voi]	251	274
[+voi]	250	235	[+voi]	8	484	[+voi]	233	242
<b>C2</b>			<b>C2</b>			<b>C2</b>		
NOM			NOM			NOM		
ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]
[-voi]	257	250	[-voi]	501	9	[-voi]	447	43
[+voi]	246	247	[+voi]	4	486	[+voi]	465	45

  

Sample A		Sample B		Sample C	
NOM	ACC	NOM	ACC	NOM	ACC
bid	kabi	kub	kubi	bap	digi
tup	gadi	dig	digi	tap	dudi
pak	bagi	tat	tati	tip	dadi
tid	padi	kut	kuti	pat	piki
gut	bigi	pup	pipi	kut	taki
kag	didi	tuk	tuki	tig	diki
gat	piki	tit	titi	dik	bigi
tad	kati	kib	kibi	pip	tuki
gag	tabi	pad	pati	dik	puki
buk	kuki	kig	kigi	gik	kibi

**Table 1:** Results from the A, B, and C simulations. Each simulation type was run 100 times. The tables at the top show the number of voiced and voiceless consonants in various positions across all simulations; below each table is the output of a single representative simulation.

## 4 Results

I ran several types of simulations, varying whether exemplars were subject to paradigm uniformity and to production bias. Each simulation type was run 100 times; Tables 1 and 2 show counts of voiced and voiceless consonants in various positions across all runs of each simulation type, plus the outcome of a single representative simulation. In these sample outcomes, each exemplar cloud is represented by a single wordform that shows the most common exemplar type in that cloud; in the event of a tie, a value is chosen at random.

The basic A simulations included neither paradigm uniformity nor a bias towards final devoicing. As shown in Table 1, these simulations produced lexicons in which the nominative and accusative forms of a given word were completely unrelated, and in which consonants in all positions were about equally likely to be voiced or voiceless. Adding paradigm uniformity in the B simulations yielded identical nominative and accusative stems for nearly all words: 98.7% of all C2s matched in voicing, much higher than chance ( $\chi^2 = 946.729$ ,  $p < 2.2 \times 10^{-16}$ ; compare this result to the 50.4% matching C2s in the A simulations,  $\chi^2 = 0.049$ ,  $p = 0.82$ ). Thus, biasing productions to look like morphologically related words succeeds in producing paradigm uniformity effects, even in a model with no abstract underlying forms from which all members of a paradigm are derived.

The C simulations included a bias towards final devoicing (but no paradigm uniformity). This bias successfully induced consistently voiceless final consonants: 91.2% of C2s in the nominative (where they were word-final) were voiceless ( $\chi^2 = 677.329$ ,  $p < 2.2 \times 10^{-16}$ ), compared to 50.3% of C2s in the nominative in the A simulations ( $\chi^2 = 0.025$ ,  $p = 0.87$ ).

D Simulations			E Simulations		
<b>C1</b>			<b>C1</b>		
	NOM			NOM	
ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]
[-voi]	503	8	[-voi]	427	47
[+voi]	9	480	[+voi]	39	487
<b>C2</b>			<b>C2</b>		
	NOM			NOM	
ACC	[-voi]	[+voi]	ACC	[-voi]	[+voi]
[-voi]	893	7	[-voi]	539	45
[+voi]	19	81	[+voi]	349	67

  

Sample D		Sample E	
NOM	ACC	NOM	ACC
bik	biki	gap	gapi
tud	tudi	gip	gibi
kak	kaki	kat	kadi
bat	bati	kut	kuti
dak	daki	dat	dati
pip	pipi	bip	bipi
kup	kupi	pug	pugi
gat	gati	dik	dugi
put	puki	pip	pibi
tip	tupi	tap	tapi

**Table 2:** Results from the D, and E simulations. Each simulation type was run 100 times. The tables at the top show the number of voiced and voiceless consonants in various positions across all simulations; below each table is the output of a single representative simulation.

Separately, then, both the production bias towards final devoicing and the bias towards paradigm uniformity produce the desired effects. If we implement both biases in the same simulation, we might expect that the result would be a classic final devoicing pattern: stem-final consonants are equally likely to be voiced or voiceless when they are not word-final (i.e., when they are protected by a suffix), but they are consistently voiceless when they are final. In other words, to model positional neutralization, a simulation should produce nominative C2s that are mostly voiceless, but accusative C2s that are equally likely to be voiced or voiceless.

Both of these biases were implemented in the D simulations; as shown in Table 2, the result was not positional neutralization. The vast majority of stems ended up with voiceless C2s in *both* the nominative *and* the accusative: in other words, final devoicing in the nominative resulted in voiceless-final stems across the board.

The fact that the simulation can produce this pattern is not, by itself, a problem. There are cases – for example, Māori verbs and Korean nouns – where paradigms appear to be undergoing restructuring on the basis of neutralization in part of the paradigm (Albright, 2008:19–21 and references therein). Rather, the problem is that the simulation produces *only* this pattern, and is apparently incapable of modeling classic positional neutralization. The solution implemented here draws on the proposal of Albright (2002) that the most informative members of a paradigm have privileged status. The idea is essentially that relationships among members of a paradigm are not symmetric: a paradigm cell that supports more contrasts has more influence than a cell that is neutralized.

Here, ‘informativeness’ was operationalized as entropy. When analogical pressure was applied to a production (step 3 above), the Shannon entropy of each feature of each segment at each position across all exemplars in the relevant cloud was calculated:

$$H_F = - \sum_{v \in V_F} p(v) \log p(v), \quad (1)$$

where  $V_F$  is the set of all possible values of feature  $F$ . A small value of  $H_F$  indicates that the value of feature  $F$  is highly predictable – for example, final consonants are usually voiceless – and thus not very informative; by contrast, a large value of  $H_F$  indicates that the feature is highly informative.  $H_F$  was used to scale the probability (70% in simulations A – D) that a given feature of the production would be changed to match the most common feature of the cloud it was being compared to. Specifically, the probability that feature  $F$  of the production would change was  $\frac{H_F}{H_{max_F}} \times .7$ , where  $H_{max_F}$  was the highest possible entropy given the number  $n$  of possible values of the feature in question (that is, the entropy if all  $n$  values are equally likely):

$$H_{max_F} = - \sum_{i=1}^n \frac{1}{n} \log \frac{1}{n} = - \log \frac{1}{n}. \quad (2)$$

Thus, a production was very likely to change to match an unpredictable feature (such as voicing in non-neutralized forms), but unlikely to change to match a predictable feature (such as voicing in final consonants).

This entropy-based modification of the paradigm uniformity bias was implemented in the E simulations, shown in Table 2. The result is robust positional neutralization: C2 was voiced in a substantial number of accusative forms (41.6%), but in the nominative, the majority of the C2s were voiceless (88.8%).

## 5 Discussion

The simulations presented here demonstrate that it is possible to model positional neutralization using an exemplar-based approach, but only under certain conditions. The two crucial components of this pattern – neutralization and analogical pressure – are already present in the literature, but simply combining them in a single model does not produce positional neutralization. Instead, the result of a ‘vanilla’ exemplar model with both positional bias and paradigm uniformity is a system in which neutralization in one cell is propagated throughout the entire paradigm.

The solution explored here is an asymmetric implementation of paradigm uniformity. Rather than allowing all members of a paradigm to influence each other equally, the direction of influence was manipulated: productions were highly likely to change to match unpredictable features, but not predictable ones. The result was a clear pattern of positional neutralization, with a robust voicing contrast in non-final positions and mostly voiceless consonants word-finally.

These results have two important implications for exemplar theory. First, they show that it is possible to model positional neutralization using an exemplar-based approach; this is good news for exemplar theory, because this type of pattern is very common and any adequate theory of phonology must be able to model it. But, second, these results show that exemplar theory is successful in replicating these patterns *only under certain conditions*. Thus, if an exemplar-type approach to phonology is correct, then these results teach us something substantive about how the model has to function: its treatment of paradigm uniformity effects cannot allow all paradigm members to influence one another equally.

A number of promising avenues for future research suggest themselves. First, the simulations presented here used categorical values for all features (e.g., [ $\pm$ voi]); it would be useful to replicate these results with continuously-valued features, at least for voicing. A particularly interesting question to explore along these lines is whether the neutralization produced in such a model is incomplete (Port & Crawford, 1989): is the precise realization of a given final C2 affected by whether the corresponding C2 in a different cell of the paradigm is voiced or voiceless?

A second question is raised by the results of the E simulations. Although a substantial number of accusative forms had voiced C2s, voiceless C2s were still a statistically significant majority (58.4%,  $p = 1.3 \times 10^{-7}$ ). In other words, the E simulations predict that if part of a paradigm is neutralized, there will be a bias towards words that have the neutralized form in all cells, although words with alternations may be present as well. To the extent that this prediction holds up in real examples of positional neutralization, the model is supported; to the extent that it does not, the model needs further refinement.

Finally, it would be useful to explore other methods of operationalizing the informativeness of different cells in a paradigm. The approach used here looks at informativeness on a feature-by-feature basis; nothing

would rule out a pattern in which, for example, the value of voicing in C1 is determined by one member of the paradigm and the value of voicing in C2 is determined by a different member. Thus, this approach is compatible with patterns that require ‘composite URs’ that combine information from several different surface forms (Kenstowicz & Kisseberth, 1979). By contrast, Albright (2002) argues on the basis of attested leveling patterns that inflectional paradigms refer to a single privileged base, the one that is most informative overall (although it may involve some neutralization) as measured by the degree to which it is possible to predict other members of the paradigm based on that form. Regardless of the specific form this asymmetric paradigm uniformity takes, these results provide independent support for the idea that informative surface forms are uniquely privileged in inflectional paradigms.

## References

- Albright, Adam (2002). *The Identification of Bases in Morphological Paradigms*. Ph.D. thesis, University of California, Los Angeles, Los Angeles, CA.
- Albright, Adam (2008). Explaining universal tendencies and language particulars in analogical change. Good, Jeff (ed.), *Linguistic Universals and Language Change*, Oxford University Press, Oxford, chap. 7, 144–181.
- Benua, Laura (1997). *Transderivational Identity: Phonological Relations between Words*. Ph.D. thesis, University of Massachusetts Amherst, Amherst, MA.
- Bybee, Joan & James L. McClelland (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22, 381–410.
- Ettliger, Marc (2007). Shifting categories: An exemplar-based computational model of chain shifts. *UC Berkeley Phonology Lab Annual Report* 177–182.
- Kenstowicz, Michael & Charles Kisseberth (1979). *Generative Phonology: Description and Theory*. Academic Press, New York, NY.
- Lombardi, Linda (2001). Why place and voice are different: Constraint-specific alternations in Optimality Theory. Lombardi, Linda (ed.), *Segmental Phonology in Optimality Theory: Constraints and Representations*, Cambridge University Press, Cambridge, chap. 1, 13–45.
- McCarthy, John J. (2005). Optimal paradigms. Downing, Laura J. & T. Alan Hall (eds.), *Paradigms in Phonological Theory*, Oxford Studies in Theoretical Linguistics, Oxford University Press, Oxford, chap. 8, 170–210.
- Oudeyer, Pierre-Yves (2006). *Self-Organization in the Evolution of Speech*. Studies in the Evolution of Language, Oxford University Press, Oxford.
- Pierrehumbert, Janet B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. Bybee, Joan L. & Paul J. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, no. 45 in *Typological Studies in Language*, John Benjamins Publishing Company, Amsterdam, 137–157.
- Port, Robert & Penny Crawford (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17:4, 257–282.
- Tupper, P. F. (2014). Exemplar dynamics and sound merger in language. Ms., Simon Fraser University.
- Wedel, Andrew (2004). *Self-Organization and Categorical Behavior in Phonology*. Ph.D. thesis, University of California, Santa Cruz.
- Yu, Alan C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24, 187–214.