AN EXEMPLAR-BASED MODEL OF CHAIN SHIFTS

Marc Ettlinger

University of California, Berkeley
marce@berkeley.edu

ABSTRACT

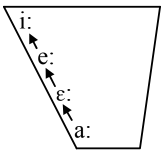
Explanations for historical chain shifts tend towards the teleological using abstract ideas like *balance* and *equilibrium* as the organizing principles of a language's sounds. This paper investigates whether there are more basic phonetic principles governing the behavior of sound categories with respect to one another. Using a computational simulation of agents communicating with each other, I show that vowel chain shifts fall out naturally from an exemplar-based model of sounds. This suggests that no overarching teleological mechanisms are required to account for chain shifts and that the self-organizing behavior of exemplar-based categories provides an adequate explanation.

Keywords: exemplars, computational modeling, vowel formants, self-organization, chain shifts

1. INTRODUCTION

Historical chain shifts are a type of sound change where the change of one sound triggers the change of another. For example, the Great Vowel Shift [6] consisted of a chain shift of the long vowels in English. Simplifying one part of the change, /e:/ moved up in the vowel space to /i:/, /ɛ:/ shifted to /e:/ and /a:/ shifted to /ɛ:/ (Fig. 1):

Figure 1: A portion of the Great Vowel Shift



The interconnectedness of these changes and those in other similar shifts have been used as evidence that the overall structure of the sounds of a language are regulated. However, the explanations for this regularization are often teleological including assertions that a disturbance in the balance of a system must be restored [5] or that "the system exerts a pull on sounds that aren't fully integrated." [13]

This begs the question of whether the movement of vowels is regulated by forces like *balance,* the *pull of a system* or *maximal dispersion* or whether there are other principles at work. The hypothesis explored here is that it is the nature of exemplar-based categories themselves that provides the basis for the self-regulating behavior exhibited by vowel systems.

Using a computational simulation of language-using agents, I show that an exemplar-based model of the production and perception of vowels predicts the shift of one vowel in the same direction as the shift of another without resorting to external constraints. This demonstrates that, in certain respects, phonological categories are *self*-regulating and that the appearance of a balanced system is epiphenomenal of more basic properties of exemplar categories.

2. MAXIMIZING THE VOWEL SPACE

The linked sound changes that make up a chain shift has been construed as being the result of maximizing the vowel space and that the maximally dispersed vowel space found cross-linguistically [14] is the result of maximizing acoustic distinctiveness [2, 11]. There are two problems with an account of this sort. First, there are a number of languages that do not maximize the vowel space, for example Abkhaz (ə, e, a; [19]) and Manobo (i, ɨ, ə, a, o, [15]). Second, these accounts describe a synchronic state of the inventory without explaining *how* the sounds become optimized. A more satisfactory account is one that explains the mechanisms by which these synchronic patterns emerge and can account for vowel systems that aren't maximally dispersed.

Two recent proposals have attempted to do that. One [7] uses a neural network to show that maximally dispersed vowel spaces are easier to learn, but they only looked at 5-vowel systems and showing that a system is easier to learn does not explain how a system can arise in the first place.

De Boer [1] proposes a model based on communicating agents involving a set of interactions between agents including imitation, feedback and success rate tracking. This mode of interaction adds a significant level of complexity, however, and may not have correlates in actual communication. Oudeyer [16] eliminated this by coupling the motor and perceptual maps through a neural network achieving similar results.

The model presented here differs from these previous accounts in two respects. First, it does not require complex interaction between agents and avoids the assumptions implicit when using neural networks. Instead, it simply posits that phonological categories exhibit properties of exemplar-based categories. Therefore, the results only reflect the self-organizing behavior of exemplar-based categories. Second, the previous models explain the well-attested balanced vowel systems, but do not have an explanation for unbalanced or non-optimal systems. This is the result of focusing on constraints on the final state of the vowel space instead of focusing on the diachronic processes of change leading to that final state. This model differs in that it seeks to model the actual process by which vowels change.

3. AN EXEMPLAR-BASED MODEL

The simulation presented here models chain shifts by simulating agents communicating with each other over time. Communication consists of the production of a random utterance by a random agent that is perceived and categorized by another. The representation of the agents' vowels and the processes of producing and perceiving these vowels are governed by exemplar theory.

First introduced in psychology as a theory of categorization [14], exemplar theory's application to linguistics is based on the idea that each of the sounds of the language represents a category which is represented by a cloud of the remembered tokens of that category. The next section presents a framework of how such a theory can be modeled computationally [17].

3.1. Vowel categories

Exemplar-models are motivated for use in Phonetics by experimental evidence that has shown that significant amounts of detailed information about heard speech is retained by listeners [8]. So, the category for the vowel /i/ consists of all of the tokens, or exemplars, of /i/ heard by the listener.

This does raise the problem that over the life span of a speaker, the sheer number of exemplars would grow intractably large for both brain and computer [8]. This is resolved in two ways [17].

First, experiments have shown that the effect of exemplars decay over time [4]. This is achieved by decrementing the strength of each exemplar exponentially according to the equation in (1) :

(1) $S_t = S_0 \cdot e^{(-t/\tau)}$,
where S_t=strength at time t, τ = decay constant

Eventually, an exemplar's strength decreases to a level where it has no effect on perception and is dropped from the exemplar cloud.

Second, implementation involves parameters for the exemplars that are discrete, rather than gradient. The use of discrete, rather than gradient, parameterization is based on evidence that listeners cannot distinguish between frequencies that differ below a certain threshold of 0.5% [18]. So, an exemplar for the vowel /i/ with an F1 of 285 Hz is stored along with one having an F1 of 286 Hz, which is within 0.5% of 285.

For vowels, the parameters that distinguish vowels can be simplified to be their formant frequencies. The chain shift discussed above can be further simplified to involve a change in just the F1 of the vowels in question, turning /e/ into /i/ and /a/ into /e/. Therefore, for the sake of this model, the one relevant piece of information for each exemplar is the value of F1. Future simulations will incorporate all relevant formant frequencies and cues such as length and formant transitions.

3.2. Production

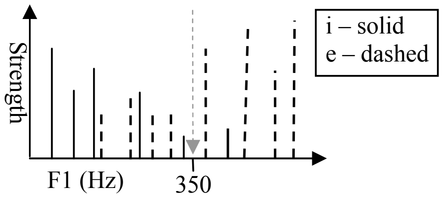
Because the focus of this model is on the behavior of vowels, each act of communication is simplified to the act of a random speaker saying a random vowel to a random hearer.

The production algorithm for a vowel involves the activation of a random exemplar within the cloud making up the vowel category, weighted by exemplar strength. The formant values of vowels for an individual speaker follow a Gaussian distribution, so variance with a standard deviation of ~45 Hz [20] is added to the frequency of the selected exemplar and the corresponding formant frequency is then "said" to the hearer.

3.3. Perception

The listener categorizes this new vowel token in one of its existing categories by looking at the nearby exemplars and assigning the new token to the category with the greatest net exemplar strength. For example, in Fig. 2, a token with an F1 of 350 Hz is heard and the nearest exemplars for the two closest vowels are shown. The figure shows more /e/ exemplars of greater strength near the new 350 Hz token than /i/ exemplars, so the new token is categorized as /e/ and added to the cloud of exemplars for the /e/ category.

Figure 2: Addition of a new exemplar



The equation for calculating the activation of each category is based on a fixed window of activation [10], and is shown in (2):

(2) $$\text{Activation}(x) = \sum_{i=x-w...x+w} s_i$$

where x = frequency (Hz), w = frequency window, s_i = exemplar strength of exemplars at frequency i

Identical results obtain if all exemplars are used, weighted in inverse proportion to their distance from the new exemplar.

3.4. Aging and cycles

After a parameterized number of token exchanges, all of the agents' exemplars are aged according to the equation in (1). This production-perception-aging sequence is repeated for a parameterized number of cycles. Vowel chain shifts are reported to occur with the course of a couple of generations [3], so the number of cycles is set to ~10,000,000, the approximate the number of times a person hears a vowel over the course of a lifetime. The results below show effects after 10,000 exchanges.

4. MODELING A CHAIN SHIFT

This section describes the successful implementation of the above model in simulating a chain shift whereby changing the quality of one vowel changes the quality of another. For the sake of brevity only two vowels are used, /i/ and /e/, but the result can be extrapolated to any number of vowels since the principles stay the same. In the model, /e/ shifts as /i/ shifts, and by the same mechanisms, /a/ also shifts as /e/ shifts.

4.1. Initial state

The description of a chain shift begins with a vowel system that is in some sort of temporary equilibrium, as was the case with English vowels some 500 years ago. A temporary equilibrium can be represented by a single speaker with a single F1 exemplar for each of the vowels (/i/ and /e/). Since the model seeks to investigate whether an incremental decrease in F1 for one vowel automatically decreases the F1 of the other vowel something needs to trigger the gradient decrease in F1 of one of the vowels. There are a number of potential articulatory and acoustic accounts of this, but that is not the subject of the investigation. Here, this decrease is set into motion by introducing another speaker with a lower F1 for /i/, but with the same F1 for /e/, and simulates a sociolinguistically driven pull chain (see Table 1):

Table 1: F1 at initial state (in Hz).

Vowel	Agent 1	Agent 2
/i/	285	215
/e/	570	570

4.2. Results

If only the /i/ vowel is considered, then the expectation is that after enough cycles, the F1 of the two agents will eventually converge to somewhere around (285+215)/2 = 250 Hz since the strength of each exemplar is the same and each agent has an equally likely chance of being a speaker or hearer. If the exemplar strength for the /i/ of one of the agents was higher than the other, we would except that the eventual F1 converged upon would be somewhere closer to the vowel of the agent with the higher strength.

Table 2: Convergence of F1 for /i/ (in Hz).

	Exchanges				
Agent	0	10	100	1000	10,000
Agent 1	285	285	254	247	247
Agent 2	215	214	226	233	242

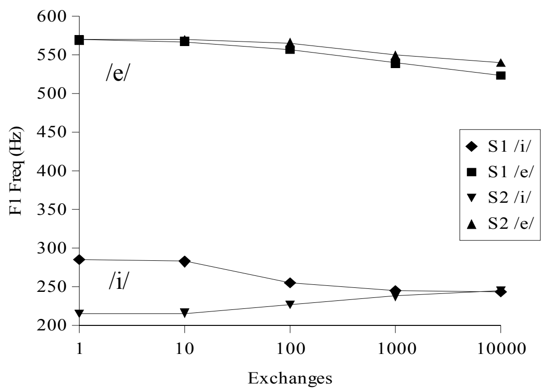
The results showing the average F1 of /i/ for both agents after n exchanges for 100 tests of the simulation are shown in Table 2. After ~10,000 exchanges, the vowels converge to around 250 Hz achieving the desired gradual decrement of F1 for agent 1's /i/. The data in Table 3 show the average of 100 runs for both vowels over the course of 10,000 exchanges and are graphed in Fig. 3.

The key number to look at is the formant frequency of the vowel /e/ for agent 1. As the table and graph show, there is an appreciable decrease in this value from its initial frequency (570 Hz) to the final average after 10,000 exchanges (524 Hz).

Table 3: F1 for /i/ and /e/ for both agents (in Hz).

Agent	vowel	Exchanges 0	10	100	1,000	10,000
agent 1	/i/	285	283	255	246	244
	/e/	**570**	**568**	**557**	**540**	**524**
agent 2	/i/	215	216	227	238	245
	/e/	570	571	566	550	541

Figure 3: F1 frequency for all speakers' vowels



This decrease is not be due to influence of the F1 frequency of the second agent because that agent starts with the same frequency for /e/ (570 Hz). Instead it's the decrease in F1 of the vowel /i/ that is causing the decrease in the F1 of /e/. This is precisely the behavior observed for chain shifts that this simulation attempts to model.

5. CONCLUSION

These simulations show that the nature of exemplar-based categories is such that when one vowel moves, another will fill the gap. The vowel space, therefore, is *self*-organizing in that its harmony and balance are maintained simply by virtue of how exemplar categories operate.

An examination of the behavior of each exchange suggests that the mechanism by which this happens is as suggested by Labov [9]. The process begins when the center of gravity of the exemplar cloud for /i/ (in Hz) begins to shift downward. This shifts the boundary – the frequency where the value for the equation in (2) is the same for both vowels – down as well. As this happens, tokens near that boundary begin to be categorized as /e/. The categorization of these low-F1 tokens as /e/ exemplars shifts /e/'s center of gravity, and therefore the boundary, down even further. As /i/ continues to change, so does /e/, and the well attested historical chain shift is observed.

This model also has the advantage of providing an explanation for why chain shifts sometimes happen and sometimes don't: vowels with a lower degree of variance do not engage in chain shifts.

Further modeling will hopefully capture the interaction of vowels in two dimensions and potentially provide insight into consonant chain shifts; why most, but not all, languages have maximally dispersed vowels; and how the intricacies of social interactions affect these phonological properties.

6. REFERENCES

[1] de Boer, B. 2000. Self organization in vowel systems, *J. of Phonetics* 28, 441–465.

[2] Flemming, E. 1995. *Auditory features in Phonology*. UCLA Ph.D. Dissertation.

[3] Gordon, M. 2001. *Small-town Values and Big-city Vowels: A Study of the Northern Cities Shift in Michigan*. Durham, N.C.: Duke University Press.

[4] Goldinger, S. D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *J. of Exp. Psyc.: Learning Mem. and Cog.* 22: 1166-1182.

[5] Jakobson, R 1931 *Über die phonologischen Sprachbünde* Travaux du Cercle Linguistique de Prague 4.

[6] Jespersen, Otto (1949). *A Modern English Grammar on Historical Principles.* Copenhagen: Ejnar Munksgaard.

[7] Joanisse, M., Seidenberg, M. 1997. [i e a u] and sometimes [o]: Perceptual and computational constraints on vowel inventories. *Proc. of the 15th Conf. of the Cog. Sci. Soc.* Stanford: CA.

[8] Johnson, K., Mullenix, J. 1997. *Talker Variability in Speech Processing*. San Diego: Academic Press.

[9] Labov, W. 1994. *Principles of linguistic change. Volume I: Internal Factors*. Oxford: Basil Blackwell.

[10] Lacerda, F. 1997. Distributed memory representations generate the perceptual-magnet effect. ms.

[11] Liljencrants, J., Lindblom B. 1972. Numerical simulation of vowel quality. systems: The role of perceptual contrast. *Language* 48, 839-862.

[12] Maddieson, I. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

[13] Martinet, A. 1955. *Economie des changements phonétiques*. Berne: Francke.

[14] Medin, D. L., Schaffer, M. M. 1978. Context theory of classification learning. *Psych. Review*, 85, 207–238.

[15] Meiklejohn, P., Meiklejohn, K. 1958. Accentuation in Sarangani Manobo. *Oceania Linguistic Monographs* 3, 1-5. Sydney: University of Sydney.

[16] Oudeyer, P. 2006. *Self-Organization in the Evolution of Speech.* Oxford University Press, 2006.

[17] Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J. and P. Hopper (Eds.) *Frequency effects and the emergence of linguistic structure*. John Benjamins, Amsterdam.

[18] Roederer, J. 1973. *Introduction to the Physics and Psychophysics of Music*. The English U.P.: London.

[19] Trubetzkoy, N. 1930. Nordkaukasische Wortgleichungen. Wiener Zeitschrift für die Kunde des Morgenlandes. Bd. XXXVII, Heft 2. Wien.

[20] Wells, J.C., 1962. *A study of the formants of the pure vowels of British English*. ms. University of London.