Polarized Variation*

Kie Zuraw
University of California, Los Angeles
kie@ucla.edu

Received: February 25, 2016
Accepted: June 22, 2016

Abstract

In cases of exceptionality, there are usually many words that behave regularly, a smaller number that behave irregularly (the exceptions), and perhaps an even smaller number whose behavior varies. This paper presents several examples of exceptionality and variation that are polarized in this way: most items exhibit one behavior or the other consistently, with only a minority of items showing variation. The result is a U-shaped histogram of behavior rates. In some cases, this requires listing of surprisingly long units. There are, however, some cases of bell-shaped histograms, where most items show variation, and only a minority are consistent. Some simple simulations are presented to show how polarized variation can result when variation is between two categorical outcomes, and both types of variation can result when variation is along a phonetic continuum.

Keywords: exceptions; variation; corpus; diachronic change

Resum. *Variació polaritzada*

Quan parlem d'excepcionalitat, generalment hi ha moltes paraules amb un comportament regular, un grup menor amb comportament irregular (les excepcions), i potser un grup encara menor que varien. Aquest article presenta diversos exemples d'excepcionalitat i de variació que son polaritzats així: la majoria dels ítems exhibeixen un comportament o l'altre quasi constantment, amb només una minoria d'ítems que varien. Resulta un histograma en forma d'U, si tracem proporcions de comportament. Hi ha alguns casos que exigeixen allistar elements de llargada sorprenent. Tanmateix, existeixen casos d'histogrames en forma de campana: la majoria dels ítems exhibeixen variació, amb només una minoria que siguin constants. Es presenten simulacions senzilles que mostren de quina manera la variació pot resultar polaritzada quan és entre categories, i de quina manera les dues menes de variació son possibles quan la variació es troba en un contínuum.

Paraules clau: excepcions; variació; corpus; canvi diacrònic

Table of Contents

* Thanks to audiences at the UCLA phonology seminar, the 2015 Old World Conference on Phonology, and the University of Massachusetts for their comments and questions, especially Bruce Hayes, Brian Smith, and Joe Pater. Thanks particularly to the organizers of the workshop, Eulàlia Bonet, Joan Mascaró, Eduard Artés, and Teresa Cabré, for bringing this paper into existence, and to two anonymous reviewers for their careful and stimulating feedback.

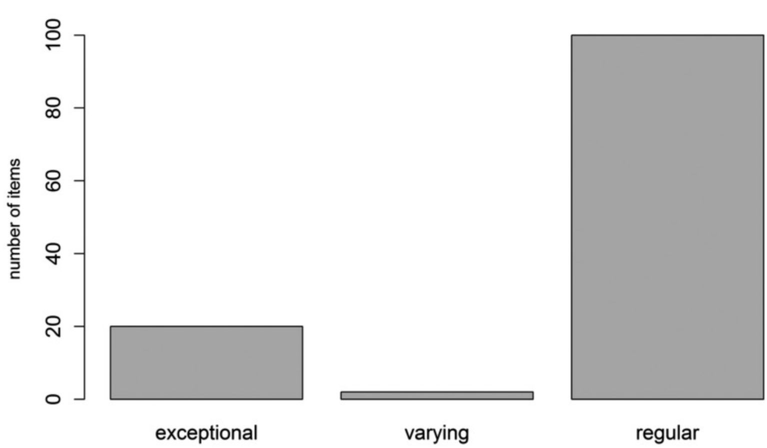
1. Introduction: free and lexical variation

We can define two extreme types of variation, and a continuum between them. At one extreme lies purely free variation, where each relevant word itself shows variation. For example, many American English speakers optionally reduce /nt/ to [ɾ̃] in roughly the same contexts where tapping applies (between two vowels, the second of which is unstressed): *spli[nt]er* varies with *spli[ɾ̃]er*, *pe[nt]atonic* with *pe[ɾ̃]atonic*. If the variation applied indiscriminately across words, with no words having an idiosyncratic tendency to favor one variant over the other, this would be a case of pure free variation. (In reality, there probably are some words with especially high or low tendencies to undergo the reduction, though this has yet to be documented as far as I know.)

At the other extreme lies purely lexical variation, where each word consistently shows one variant or the other. The English past tense is close to being of this type: for the most part, some verbs have a consistent regular past (*reach, reached*), and others a consistent irregular past (*teach*, *taught*). Here, variation is over types (different verbs, *reach* vs. *teach*), not over tokens (occasions of uttering *splinter*). In an experimental context, Richtsmeier (2008) uses the terms token variation and type variation to distinguish the two extremes of our continuum.

There do exist mixed cases that fall somewhere between free/token and lexical/type variation. Some English verbs do actually show free variation. For many speakers, *sneak* has an informal irregular past *snuck* alongside the prescriptively correct regular *sneaked*, and *dive* has interchangeable plurals *dove* and *dived* for many North American speakers. The real state of English past tense is thus close to being lexical variation, yielding a U-shaped distribution whose two arms are of uneven height, as shown in (1). If there were more words like *sneak* or *dive* (varying), at the expense of words like *teach* (exceptional) and *reach* (regular), the case would be closer to free variation.

(1) Schematic U-shaped distribution of exceptionality



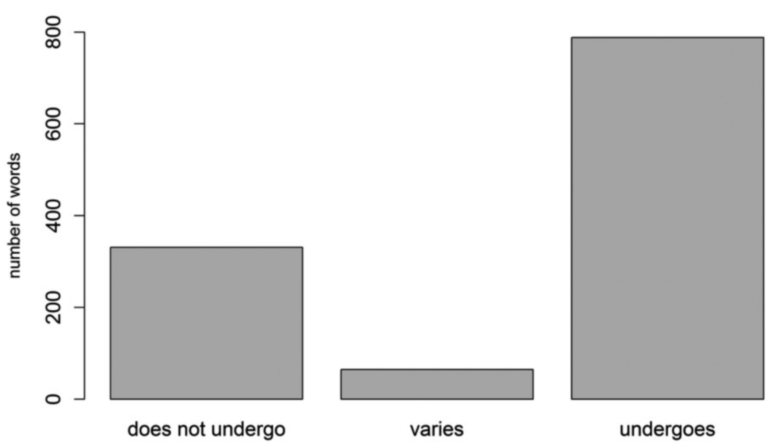
This paper has two parts. First, it presents several cases of U-shaped distributions, showing that they are common, that they involved both affixed words and longer phrases, and that they suggest listing of fairly large units is possible. In many cases it is not clear which variant should be labelled the exception, so the variants will be given more-neutral labels. The second part of the paper shows two ways that U-shaped distributions could arise: listing, and thresholding on a phonetic continuum.

2. Case studies of U-shaped distributions

2.1. Tagalog nasal substitution

Zuraw (2010) examines Tagalog nasal substitution, using data from English's 1986 dictionary. Undergoers of nasal substitution fuse a prefix nasal with a stem-initial obstruent, as in /maN+bigáj/ → [ma+migáj] 'distribute'. Non-undergoers simply assimilate the nasal in place: /paN+diníg/ → [pan+diníg] 'hearing'. (The non-undergoers could be labeled exceptions, but then again the rule applies only at the prefix-stem boundary, so it would also be reasonable to label the undergoers as exceptions.) If we tally up the number of words, across all phonological and morphological sub-cases, that undergo the rule, fail to undergo, or are listed with both variants in the dictionary, we derive a real U-shaped distribution, similar to the hypothetical one from (1):

(2) Tagalog nasal substitution data from Zuraw (2010), six most-common morphological constructions combined

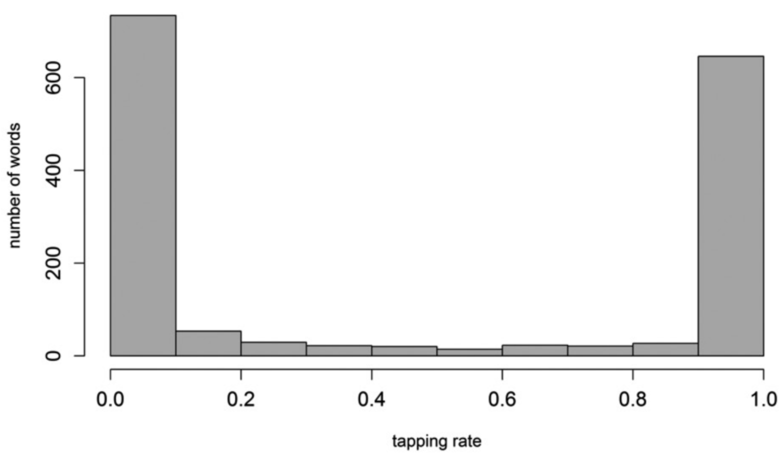


2.2. Tagalog tapping at prefix-stem boundary

Zuraw (2009) examines tapping in Tagalog, using data from the written corpus described in Zuraw 2006. Tapping occurs in a few environments, but here we will look at only the prefix-stem boundary. If a stem begins with /d/ and receives a vowel-final prefix, that /d/ is now in the environment V__V, and can undergo tapping: /ma+dumí/ → [ma+ɾumí] 'dirty'. Failure to undergo tapping is also common, as in /ma+dáhon/ → [ma+dáhon] 'leafy'. Because the data here are from a corpus, we can ask not just if a word undergoes tapping or not, but how often it undergoes tapping, and plot the number of words that show each rate of tapping. Very-low-frequency items must be avoided, because they will look artificially polarized. For example, if a word occurs only once, its observed rate of undergoing tapping must be 0% or 100%, no matter what its true underlying rate of undergoing might be. Because there are many low-frequency items (the lower the frequency, the more items there are with that frequency), this inflates the 0% and 100% counts. The data shown here are limited to items that occur at least 10 times.

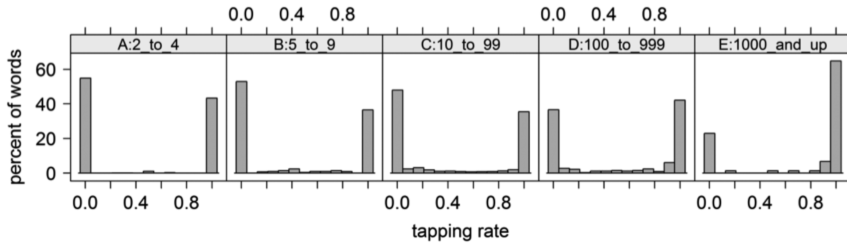
As the histogram in (3) shows, there are a large number of words that consistently fail to undergo tapping (on the far left bin, rates of 0% to 10% tapping), and a large number that consistently do undergo tapping (far right, 90% to 100% tapping rate)

(3) Histogram of tapping rates for prefixed words with frequency ≥ 10



Using corpus data also allows us to see how a word's frequency is related to its tapping rate. In (4), we see histograms for five groups of words: those with frequency 2 to 4 (Group A, not included in (3)); frequency 5 to 9 (Group B, also not included in (3)); frequency 10 to 99 (C); 100 to 999 (D); and 1000 and up (E). Because the groups contain different numbers of words, the vertical axis is the percentage of words in each group that fall into each tapping-rate bin, rather than the raw number.

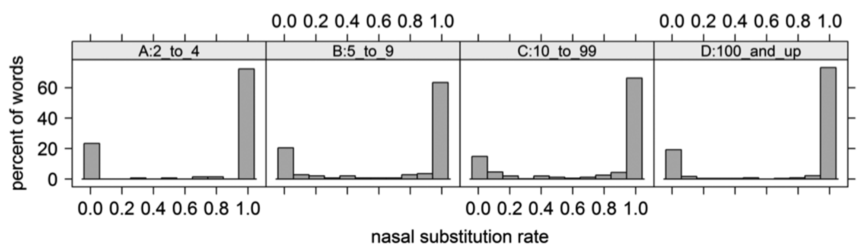
(4) Histograms of tapping rates, broken down by word frequency



We can see that the U shape remains in each frequency bin, although the balance of non-undergoers (left arm) and undergoers (right arm) changes from bin to bin, with more and more undergoers as frequency increases.

Returning briefly to Tagalog nasal substitution, the same corpus data can be used to get nasal-substitution rates for individual words (see Zuraw 2009 for limits on using written data for nasal substitution). The result, shown in (5), is similar: a U-shaped distribution in every frequency tranche, but with more undergoers at higher frequencies.

(5) Histograms of nasal substitution rates (all prefixes), broken down by word frequency



2.3. English aspiration

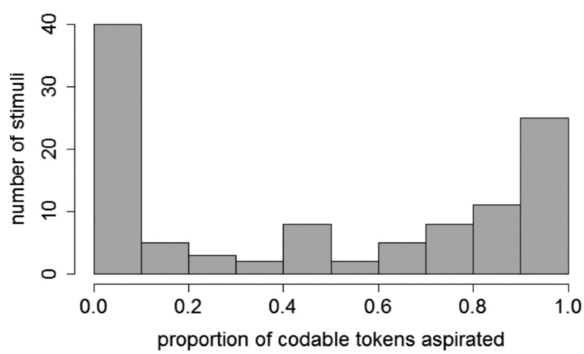
In English, word-initial voiceless stops are aspirated, regardless of stress, as in *[pʰ]ercéption*. The best way to describe the aspiration environment is probably <u>prosodic</u>-word-initial, because aspiration also applies after a productive prefix, again even if the stem's first syllable is unstressed: *mìs-[pʰ]ercéption* (Zuraw & Peperkamp 2015). If *misperception* were a single phonological word, then the stem-initial /p/ would not be eligible for aspiration, even if it is syllable-initial (*mis.per.cep.tion*), because the syllable it initiates is unstressed. But if the stem *perception* forms its own prosodic word, with the prefix *mis-* adjoined to form a larger prosodic word, then the /p/ remains eligible for aspiration.

Some words, however, act as monomorphemic, even if they are arguably still prefixed, such as *mistake*, with an unaspirated [t] (Ogden et al. 2000, Baker, Smith & Hawkins 2007, Smith, Baker & Hawkins 2012, Clayards, Hawkins & Gareth

2012). *Mistake* does contain the meaning of *mis*-, although not really the meaning of *take*, but nonetheless it is pronounced as though monomorphemic, with the /t/ unaspirated because it is not syllable-initial (*mi.stáke*, or perhaps an ambisyllabic *s*).

The studies just cited compare in detail a small number of words predicted to have a clear prefixed or a clear monomorphemic behavior. Zuraw & Peperkamp (2015) conducted a production study of a large number of words (110), looking for factors that affect the variation, and for items that vary. There were up to 16 tokens of each word (one for each participant, with data occasionally discarded for disfluency or spirantization), yielding a good picture of cross-speaker variation rates. As shown in (6), there were many words that were consistently unaspirated (far left: *disposed, disparities, mistakes, disturbing*, and 36 others), a smaller number that were consistently aspirated (far right: *miscalculations*, *misquote*, *mistrial*, *discontinuity*, and 22 others); in the middle, there were 44 items with variation (*displacement*, *discourteously*, *discolored*, and 41 others).[1]

(6) Histogram of aspiration rates in English prefixed words



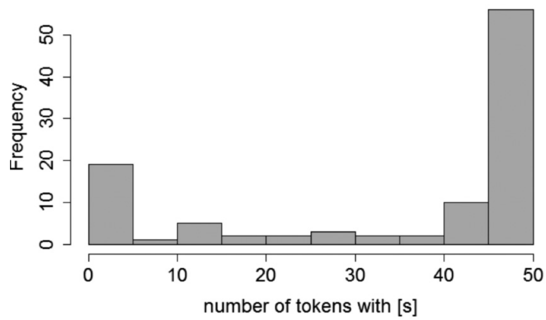
2.4. Baroni's Northern Italian s-voicing

Baroni (2001) examines the rule of Northern Italian intervocalic *s*-voicing. Within monomorphemes, the rule is straightforward: while *casa* 'house' would be [ˈkasa] in other dialects of Italian, for Northern speakers it is [ˈkaza]. The rule is typically blocked at a prefix-stem boundary, as in *a[s]immetrico* 'asymmetric'—that is, just as in English, stem-initial consonants behave as word-initial. But, also as in English, some prefixed words behave as though monomorphemic, and do undergo voicing: *pre[z]upposizione* 'presupposition'. Baroni tests many factors that could affect this variation, through a production study in which 10 Northern Italian speakers read aloud 102 real words in sentence context, 5 times each. Plotting the data from Baroni's appendix, we see in (7) that the largest group of words consistently

1. Percent aspirated here excludes tokens coded as "unsure", and thus are different from the percentages shown in the appendix of Zuraw & Peperkamp (2015).

resist voicing (rightmost column), the next-largest consistently undergo it (leftmost column), and there are a substantial number of words—27—in between.

(7) Histogram of [s] rate in Baroni's data

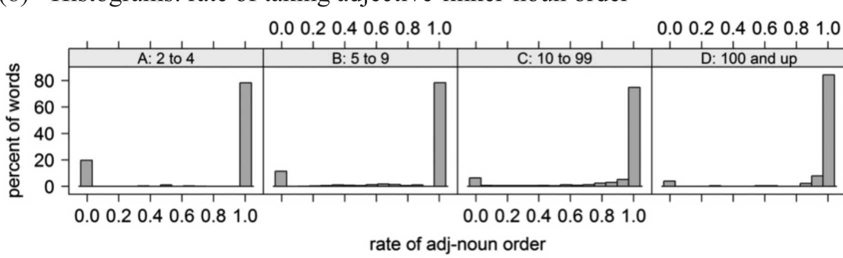


2.5. Tagalog noun/adjective order

The above cases have all involved prefixed words. Now, we move to larger, multi-word units.

In Tagalog, a noun and its modifying adjective can occur in either order, with a "linker" morpheme in between (Schachter & Otanes 1972): *maganda-ng babae* (beautiful-LINKER woman) and *babae-ng maganda* (woman-LINKER beautiful) both mean 'beautiful woman'. Shih & Zuraw (2014, submitted) examine phonological and other factors that influence the order variation, using data from the Zuraw (2006) corpus. The data are polarized: as shown in (8), at every frequency tranche most adjective/noun pairs consistently have the order adjective-linker-noun, with a smaller group of consistent noun-linker-adjective pairs (far left column in each histogram).

(8) Histograms: rate of taking adjective-linker-noun order

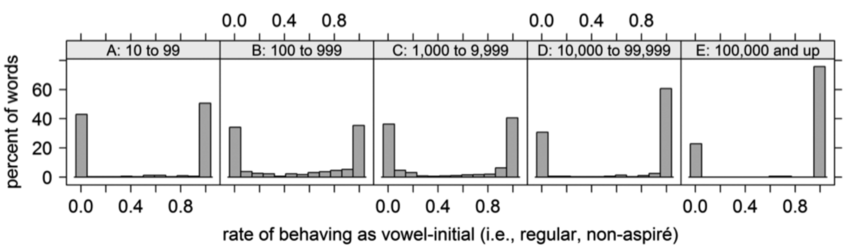


2.6. French h aspiré

In French, certain function words and adjectives undergo changes before a vowel-initial word, as in /lə eʁitaʒ/ → [l eʁitaʒ] 'the heritage', where the determiner loses its final vowel. There are some vowel-initial words that exceptionally block these changes, as in /lə eʁisõ/ → [lə eʁisõ] 'the hedgehog'; because most of these words are spelled with *h*, they are known as "h-aspiré" words. They have been subject to various analyses, but several phenomena can be covered at once if the words are treated as requiring a syllable boundary at their left edge (Frei 1929, Damourette & Pichon 1930, Malmberg 1943, Tranel 1981, Cornulier 1981, Tranel & Del Gobbo 2002, one possibility presented in Boersma 2007).

Zuraw & Hayes (submitted) examine variation in *h-aspiré*, using data from the Google n-grams corpus (Michel et al. 2011, Lin et al. 2012), which compiles frequencies of word sequences from digitized books. Zuraw and Hayes look only at words that are likely to show *h-aspiré* behavior: words that begin with a glide sound (*yacht* 'yacht'); words that begin with a vowel sound and are spelled with initial *h* (*héritage* or *hérisson*); and a few words that begin with a vowel sound and are spelled with a vowel letter, but are claimed in dictionaries and prescriptive sources to behave as *h-aspiré* (*uhlan* 'uhlan [type of soldier]'). Even within this group, the distribution is highly polarized at every frequency tranche, as shown in (9). The left arm of each U is words that are consistently *h-aspiré*, and the right arm is words that are consistently regular.

(9) Histogram of rate of regular (non-*h-aspiré*) behavior



2.7. French de *omission in coordination*

Just as in English, French *de* 'of' can be omitted in coordination: *morceaux de carottes et (de) tomates* 'pieces of carrots and (of) tomatoes' (Tseng 2003, Abeillé et al. 2004, Abeillé et al. 2005, Abeillé et al. 2006). Unlike in English, however, it is more common to retain the second *de*:

(10) Two options for French *de*

more common: de is repeated

a. morceaux [de carottes] et [de tomates]
b. pieces [of carrots] and [of tomatoes]

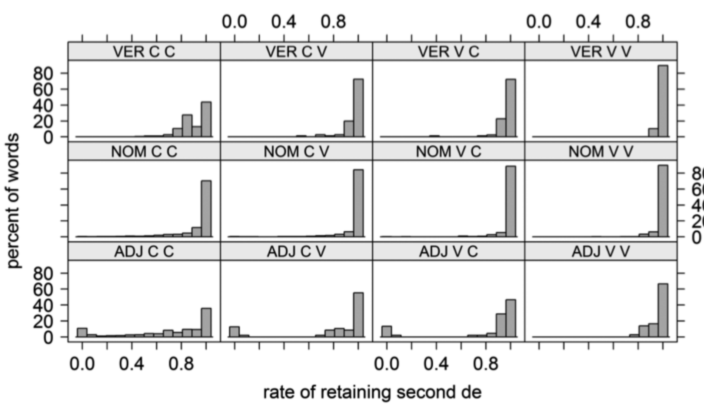
less common: first *de* takes wide scope / second *de* is omitted

c. morceaux de [carottes et tomates]
d. pieces of [carrots and tomatoes]

Zuraw (2015), again using data from Google n-grams, examines phonological factors that influence *de* omission, using simple cases of the form *de X et (de) Y* 'of X and (of) Y' and *de X ou (de) Y* 'of X or (of) Y', where X and Y are single words. The strongest factor is the initial sounds of X and Y: if either X or Y starts with a vowel (*h-aspiré* words are excluded here), the second *de* is more likely to appear. As the histograms in (11) show, most X-et-Y and X-ou-Y triples consistently retain the second *de*, and thus are counted in the rightmost column of their histogram. In the top two rows (VER = verbs, NOM = nouns), are no triples that consistently drop the second *de*. The effect of beginning with a consonant or vowel is seen mostly in the balance of variers vs. consistent retainers. If both X and Y begin with a consonant (C C), then there are more variers than if one begins with a vowel (C V or V C), or if both begin with vowels (V V). This is thus a one-sided polarization, with only one arm of the U present.

But in the bottom row (ADJ = adjective), where rates of *de* retention are lower over all, instead of just seeing more and more variers, we begin to see a small group of triplets that consistently omit the second *de* (the small columns in the far left of each ADJ histogram), outnumbering triples that retain at low rates.

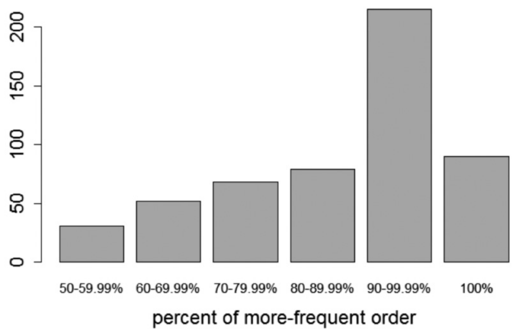
(11) Histograms of rate of retaining second *de*



2.8. English binomials

English coordinated pairs of words have become known as binomials. As Mollin (2014) discusses, some have a fixed order: *law and order* is common, and is also the name of a television program; ??*order and law*, even outside the context of television, sounds strange. But others vary freely, such as *knowledge and skills*/*skills and knowledge*. Mollin points out that cases of total irreversibility are rare; in her corpus data, it was more common for a pair to have a certain order 90% to 99.99% of the time than to have a certain order 100% of the time:

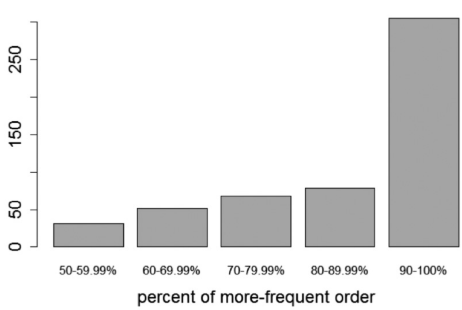
(12) Number of two-word pairs, adapted from Mollin's Figure 3.5, p. 45



(In the cases examined until now, there were two well-defined behaviors to examine: aspirated vs. unaspirated; noun-linker-adjective order vs. adjective-linker-noun order, etc. Here, there are no such labels available: does *law and order* belong in the same category as *knowledge and skills*, or as *skills and knowledge*? There seems to be no reasonable answer to this question. Therefore, Mollin simply asks, for each pair, how frequent the more-frequent order is. If the answer is 50%, the pair exhibits maximal variation; if 100%, minimal variation. The most polarized possible pattern, then, is one in which there is a tall bar at the right of the plot, but, unlike in the plots we have seen so far, there is not a tall bar at the left of the plot.)

The bar plot in (12) is not a typical histogram, however, because the horizontal axis is nonlinear: most of the bins cover 10 percentage points, but the rightmost covers only 0.01 percentage points. To make a histogram comparable to the ones used for the other cases here, we add the items on the far right to the adjacent bar:

(13) Histogram version of (12)



The result is strongly polarized: most of Mollin's pairs strongly favor one order, even if they do not show that order 100% of the time.

2.9. Other cases

The cases examined here have been chosen opportunistically, mainly by going through my own previous and ongoing research, looking for cases where big enough numbers were available.

I know of two other cases currently under investigation: Yanovich's (2014) work on Old High German ð and *d*, and Smith's (in preparation) reexamination of Hilpert's (2008) data on English comparatives. English adjectives can form their comparative as either *more X* or *Xer* (*more proud* vs. *prouder*, to use Hilpert's example). In Hilpert's data, lower-frequency adjectives (frequency < 100) show rather free variation, but higher-frequency adjectives (≥100) show a strong tendency to take *–er*, with a smaller group consistently taking *more*—this group only barely outnumbers the adjacent group of variers, though (Brian Smith, p.c.).

2.10. Within-speaker variation?

One worry for the data discussed here is whether we see any true free (token) variation at all, or just a mix of lexical (type) variation. Consider this scenario: suppose that for most items, speakers agreed on whether each item should have one behavior or the other; these items appear in the outer arms of the U-shaped histogram. Then suppose that for a few items, some speakers have one variant listed in their mental dictionaries, and some speakers the other; these are the items with intermediate observed rates of behavior. In that case, the only thing we need to explain is why there are few items with middle rates—that is, why speech communities are generally good at agreeing on one pronunciation.

In most of the cases here, we lack the within-speaker data we would need to assess this scenario, either because the data come from a multi-writer corpus, with individual authors not distinguished (the Tagalog and French cases), or because we have individual data but with only one token of each item from each speaker (English aspiration).

In the Northern Italian *s*-voicing case, Baroni elicited five tokens of each item from each speaker, but does not report data broken down by speaker in his appendix. He does, in discussing one item, *coseno* 'cosine', give a breakdown for that item though: 2 speakers consistently said *co[z]eno*, 4 consistently said *co[s]eno*, and 6 had variation.[2]

In the French case, although the corpus data offer no insight into individual variation, *h-aspiré* has attracted the attention of so many grammarians and linguists that the literature does have anecdotal examples of within-speaker or within-writer variation. Cohen (1963) mentions a flyer with *de un mois* and *d'un mois* 'of one month' occurring two paragraphs apart; *de un* treats *un* 'one' as *aspiré*, while *d'un* does not. Cohen also mentions a history manual that similarly varies between *de*

2. Baroni reports here the data for 12 speakers; the data shown in (7) are for just the 10 speakers whose results were strongly correlated with one another. The data remain similarly polarized if the results for all 12 speakers are plotted.

Henri IV and *d'Henri IV* 'of Henry IV'. Grevisse & Goosse (2011), a widely used prescriptive manual, mentions several examples of variation within French authors (pp. 57-63). Cornulier (1981), drawing on his own impressions, states that "a sentence containing both *de Hugo* and *d'Hugo* would not be remarkable in conversation" (p. 203; translation mine). He also remarks that "many people supposed to speak well seem to flip a coin each time" when it comes to pronouncing a final [n] in the article *un* 'a' in *un handicap* 'a handicap', *un HLM* 'a social housing project', or *un hameçon* 'a hook', "but, when questioned, decide in a definitive and sincere fashion, like the grammarians" (p. 203). Pronouncing *un* with a final [n] (liaison) means treating the vowel-initial following word as regular, while pronouncing with without [n] means treating it as *aspiré*. Gabriel & Meisenburg (2009) find intra-speaker variation in a reading-aloud task.

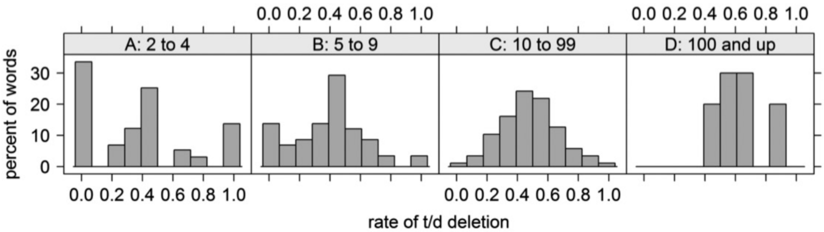
Although firm data are scant, then, it seems that within-speaker variation is real, at least in some cases. The next section of this paper asks how polarized (and non-polarized) patterns of variation can arise when there is within-speaker variation.

3. Modeling U-shaped (and non-U-shaped) distributions

3.1. Not all distributions are polarized: English t/d deletion, Tagalog o/u alternation

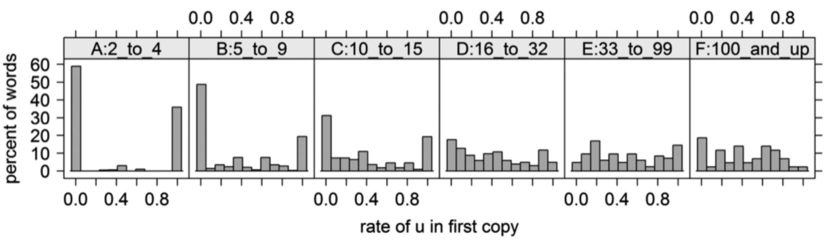
There are cases of variation that are not polarized. Coetzee & Kawahara (2013) examine English /t/ and /d/ deletion in words with final consonant clusters, such as *sexist* and *girlfriend*. Coetzee and Kawahara use spoken data from the Buckeye corpus (Pitt et al. 2007); their online appendix makes it possible to compile the rate of deletion for each word. The histograms in (14) show that the distribution of deletion rate, in each frequency tranche, is actually rather unimodal (except for the lowest-frequency items, which can show spurious U-shapedness, as discussed in Section 2.2), with the greatest number of words showing intermediate rates of deletion, and smaller numbers showing extremely high or extremely low rates.

(14) Histograms of English t/d deletion rates in Coetzee & Kawahara (2013) broken down by frequency



Data from Zuraw (2009) on Tagalog o/u alternations show a mixed pattern. Tagalog tends to have [o] in final syllables and [u] in non-final syllables, though contrast in loans and partial productivity of the alternation impose many complications. A simple example of an alternation would be [tákot] 'fear', [takút-in] 'to frighten-*patient focus*'. When a two-syllable word such as [hálo?] 'mixture' is reduplicated, the end of the first copy is pulled in two directions: on the one hand, it is not final in the syntactic word, and thus should have [u] rather than [o] ([hál**u**-hálo?] 'a dessert of mixed fruits, ice, and sweets']); on the other hand, it is arguably final in a prosodic word, and also corresponds to a word-final syllable, and on both counts should have [o] ([hál**o**-hálo?]). Both variants do tend to occur: [hál**o**-hálo?] ~ [hál**u**hálo?]. If we compile the rate of having (written) [u] in the first copy for each reduplicated word in the corpus, the pattern is U-shaped for lower frequencies and flat for higher frequencies:

(15) Histograms of Tagalog o-raising rates in reduplication, broken down by frequency



Not all distributions are polarized then. This section considers two mechanisms by which polarization could arise, then how non-polarized distributions could arise under the same mechanisms.

3.2. Polarization through lexical diffusion

Reali & Griffiths (2009) show how polarized distributions can be obtained if agents have a prior expectation of polarization. Can we obtain a U shape without such a prior bias? One possibility is to assume that what we observe in the polarized cases discussed above is a snapshot of the process of lexical diffusion (e.g., Wang 1969, Bybee 2001, Phillips 2006), a process by which words are moving, one by one, from one category to another, such as from exceptional to regular.

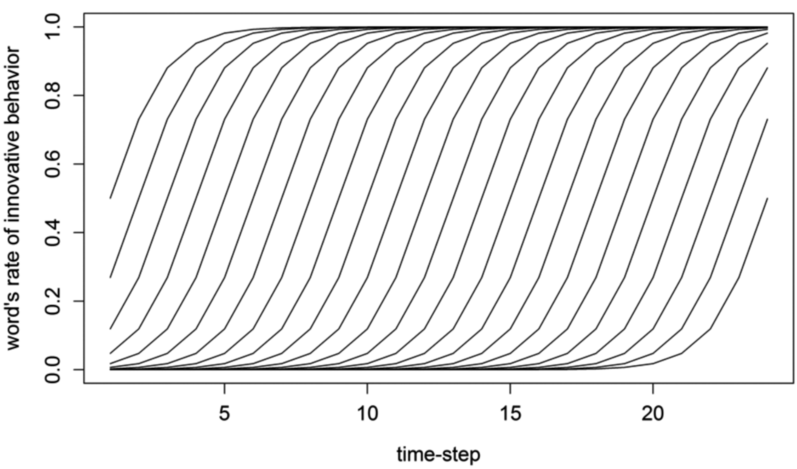
3.2.1. A simple model

We will begin with a very simple model and gradually build in more realism, though the final model will still be quite simplified.

Suppose that each word changes abruptly—that is, there is only a short transition period for each word during which it displays variation (we will return within this subsection to why the transition period should be short). Suppose further that different words transition at different times.

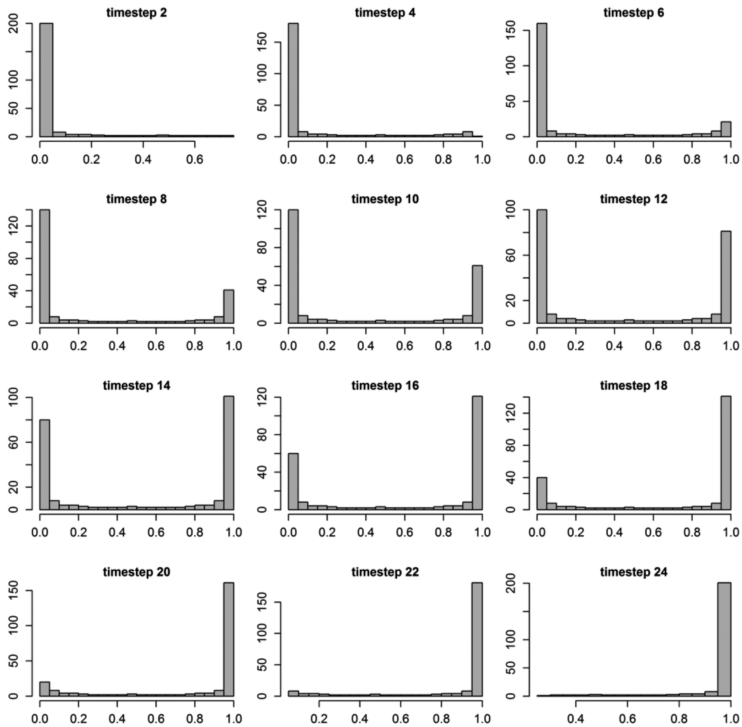
In our simple model there are 240 words, transitioning from a conservative variant to an innovative variant over the course of 24 time units, which could perhaps be thought of as decades for a slow change, or years for a fast change. Each word's rate of showing the innovative variant begins near zero, then quickly increases and tops out near 1 (using a logistic function, rate(time, intercept) = 1 / (1+e$^{-(k*time*intercept)}$)). The words begin their transitions at different times though (each word has its own value for *intercept*). The plot in (16) shows the behavior of individual words over time, but only for every 10th word, so that the individual traces can be distinguished.

(16) Increase in innovative behavior for each word over time (only every 10th word shown)



At each of the 24 time-steps, we can make a histogram like the ones in the case studies above: that is, we can compile the number of words that show each rate of innovative behavior. The histograms in (17), which plot every second time-step, show that at the beginning, nearly all the words consistently show the conservative behavior; over the time, more and more words consistently show the innovative behavior, until nearly all do. The key is that each word spends only a short time in the middle bins, moving quickly from the leftmost bin to the rightmost bin; thus, the middle bins are never very populated, and the distribution is always U-shaped, though the relative heights of the U's two arms change over time.

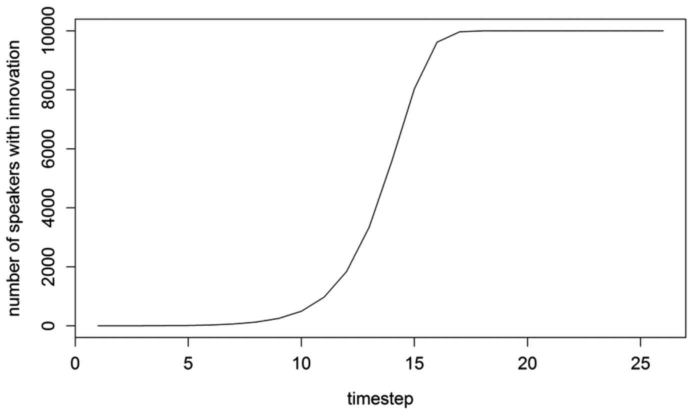
(17) Histograms across time of rate of innovative behavior



In the model above, we simply stipulated that each item changes abruptly (and thus spends little time in the middle bins of the histogram). Can we derive this behavior instead of stipulating it? Starting with Bloomfield (1933), linguists have noted that we can derive "S-shaped" change—change starts slowly, then is fast, then slows again, as shown for each word in (16)—if we adopt an epidemiological model. A speaker is infected by an innovation when she or he hears it from another speaker. Change is slow at first, because there are few innovators who can infect others, and change is slow at the end because almost everyone has already been infected.

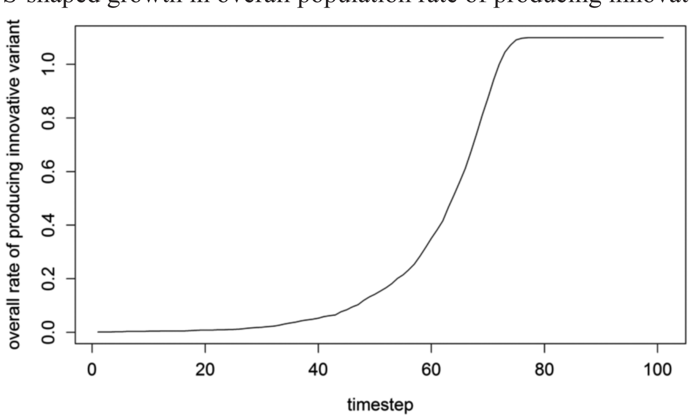
A simple simulation illustrates this. There are 10,000 speakers. At first, just one is innovative. At each time-step, speakers interact in randomly chosen pairs. If both are innovators or both are conservative, no change occurs; but if they mismatch, the innovator infects the non-innovator, who is henceforth an innovator too. The number of innovators in the population grows over time as shown in (18).

(18) S-shaped growth in number of innovating speakers for an item



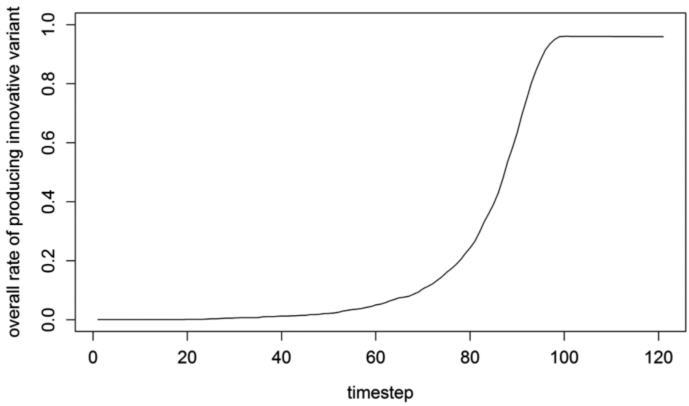
The above model is one where variation is only across speakers. To model the possibility of variation within each speaker, we allow each agent to have two listed forms (this is Baroni's 2001 proposal for Northern Italian *s*-voicing). Production depends on the strength of each form. Agents again meet in pairs. If one produces the innovative variant, the other increments its probability of producing the innovative form by 10 percentage points (unless it is already at 100%). The simulation begins with one agent at 10% innovative, and the rest at 0% innovative. Although it takes longer for change to spread, the spread is again (asymmetrically) S-shaped:

(19) S-shaped growth in overall population rate of producing innovative variant



The final adjustment to the model here is to allow not just the innovative form to be incremented, but also the conservative form. Specifically, an agent increments its innovative-form strength by 0.1 when it hears the innovative form, and increments the conservative form's strength by 0.0015 when it hears the conservative form. It is crucial that the conservative form be less amenable to strengthening; if the conservative increment is too large, change never takes off:

(20) Growth in overall population rate of producing innovative variant when both variants can be incremented



The model presented here was very crude, for expository purposes. More-serious models of language change have incorporated elements such as social network structure, or a more-sophisticated decision rule for speakers (Niyogi 2009, Pierrehumbert, Stonedahl & Daland 2014, and many others).

3.2.2. Why should an item change?

The model of change just developed stipulates an advantage for an innovative form: listeners are more likely to adopt, or increase their tendency to use, an innovative form they hear than they are to adopt a conservative form that they hear. Why should this be? Two answers have been proposed.

Sometimes one form has an inherent advantage over the other. Mollin (2014) argues that this is often the case for English binomials. One form could have a phonological advantage over the other, such as a better rhythm: *párt and párcel* has an alternating strong-weak-strong-weak rhythm, while *??párcel and párt* has a lapseful strong-weak-weak-strong rhythm (see also Shih 2014). Or one form could have a processing advantage, such as greater retrievability of the first coordinand because of frequency, saliency, or recent occurrence (Benor & Levy 2006). A linguistic or psycholinguistic advantage then helps a form become the frozen form, Mollin proposes. In terms of the model above, this could be implemented by having the probability of producing *part and parcel* vs. *parcel and part* be a function of both the strength of the listed entry for each order and the inherent probability of each order.

It is also possible, though, that innovation can be preferred for its own sake. As part of a series of simulations, Pierrehumbert et al. (2014) allow some speakers to be biased to prefer novelty, and others to shun novelty. The balance of such agents in the population will of course affect the spread of an innovation, but so can their distribution. Pierrehumbert et al. find that innovations are especially likely to spread if the agents socially connected to an innovator themselves prefer innovation. As Pierrehumbert et al. discuss, this may be linked to individual personality/cognitive traits and to tendencies for individuals with similar traits to affiliate socially (pp. 18-19). (See also Yu 2010, Stevens & Harrington 2014 on experimental findings with implications for the role of individual differences in sound change.)

3.2.3. What is the starting point for change from one category to another?

The starting point in the simulation above was that all items had the same behavior (the conservative variant); change then consisted in items' moving to the other category. This subsection considers starting points that seem plausible for the case studies presented above.

Tagalog tapping, English aspiration, and Northern Italian *s*-voicing (sections 2.2, 2.3, 2.4) seem to be in line with an all-conservative starting point. They involve an alternation that is blocked at the beginning of a prosodic word, including at the prefix-stem boundary. This is shown as Stage 1 in (21), using the Tagalog example. Some words, especially frequent words, could get reanalyzed as monomorphemic, or at least treated as not quite productively prefixed; such a reanalysis would both promote and be promoted by semantic drift from the unprefixed form (see Hay 2003).

(21) Genesis of exceptional tapping in Tagalog

Stage 1			*Stage 2*	
Unprefixed: tapping	[á**ɾ**aw]	'day'	*same as before*	[á**ɾ**aw]
Prefixed: no tapping	[ma+**d**áhon]	'leafy'	*same as before*	[ma+**d**áhon]
Prefixed: no tapping	[ma+**d**umí]	'dirty'	not quite prefixed: tapping	[ma**ɾ**umí]

The starting point then is of most items' being consistently conservative (tapping is blocked). Some items could even have already been reanalyzed as monomorphemic before tapping even arose, so that they were exceptions from the beginning. Individual words' analyses would change over time, crossing from prefixed to monomorphemic and perhaps the other direction too. These changes would have had to be abrupt, along the lines sketched in section 3.2.1, so that varying items are always rarer than consistent items.

In French *h-aspiré* (section 2.6), simplifying and idealizing somewhat, the starting point was that all items were regular. Consonant-initial words behaved as consonant-initial, and vowel-initial words as vowel-initial. This is shown as Stage 1 in (22), where consonant-initial words take the full article [la], and vowel-initial words take the form [l] to avoid a vowel-vowel sequence. The consonant /h/ then deleted, stranding some previously consonant-initial words as exceptional (Stage 2).

(22) Genesis of French *h-aspiré* (idealized), using the example of *la* 'the-feminine'

Stage 1			*Stage 2*	
C-initial: no deletion	[la tabl]	'the table'	*same as before*	[la tabl]
C-initial: no deletion	[la haʃ]	'the axe'	exception: no deletion	[la aʃ]
V-initial: deletion	[l arbr]	'the tree'	*same as before*	[l arbr]

French thus had a polarized starting state, with most vowel-initial words being consistently regular, and others (the *h-aspiré* words) being consistently exceptional. Vowel-initial items could then change category through regularization and hypercorrection, and new words entering the language would have to be assigned to a category.

English binomials present a third scenario: there is no clear starting point for an individual item, no reason why either *books and papers* or *papers and books* should have been the original order. The starting point here is probably free variation: any new coordination can occur in either order, though perhaps with a preference for one order if phonological, semantic, and processing factors favor it. Because these factors would be just as likely to conflict as to conspire to favor one variant strongly, it seems likely that most items would start out in the intermediate range. But, as a binomial begins to be lexicalized, it would move towards one arm of the U (this type of change was not implemented in section 3.2.1).

To summarize section 3.2, we can derive polarized, U-shaped distributions by assuming that individual items change category abruptly. Each item spends little time in the intermediate range, so that there are always few variers. In order for this model to work, most of the items in question must be listed. This is not too controversial for affixed words, but is more surprising for the longer units: Tagalog noun/linker/adjective vs. adjective/linker/noun (2.5), French coordinations with *de* (2.7), and English binomials (2.8), though the English binomial literature does typically assume that many items are listed (see Mollin 2014 for overview).

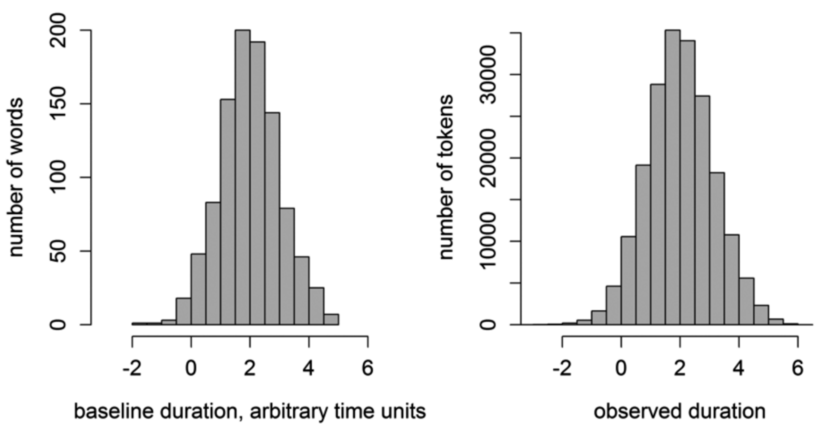
3.3. Polarization through thresholding

The question of how to derive a non-polarized distribution still has not been addressed. This section discusses how both polarized and non-polarized distributions could arise from variation along a phonetic continuum.

Suppose that Tagalog nasal substitution began (in a proto-language) as variable deletion. For example, the proto-language equivalent of /maN+bilí/, originally pronounced with simple nasal place assimilation as [mambilí], began to be pronounced as [mamilí]. It seems plausible that variation along these lines represents variation in gestural timing: the duration of the nasal gesture remains constant, but that of the labial closure shortens, producing a phonetic continuum from [mb] to [m]. When the duration of the [b] portion becomes short enough, /b/ is perceived as deleted, as the word could subsequently be lexicalized as /mamilí/. (Other phonetic changes would be relevant too: Beddor & Onsuwan 2003 show that for Ikalanga speakers differentiating between sequences like [a^{m}ba] and [ama], carryover nasalization is even more important as a cue than oral closure duration.)

We can simulate this situation with 1,000 words, each having a baseline preferred duration of non-nasal closure (that is, the [b] portion of [mb], the [t] portion of [nt], etc.). These baselines could be related to word frequency, stress, place of articulation, perhaps even other nearby segments, and other factors. These baseline preferred durations are normally distributed, as shown in (23)a, centered on the value 2 in arbitrary time units, with a standard deviation of 1. Each item is uttered 200 times, adding a small amount of Gaussian noise to the duration each time (mean 0, s.d. 1, divided by 0.5). The value of the noise can be positive or negative, so the oral stop duration can be longer or shorter than is usual for that word. The resulting observed durations, shown in (23)b, remain strongly unimodal.

(23) Simulation of variable stop deletion

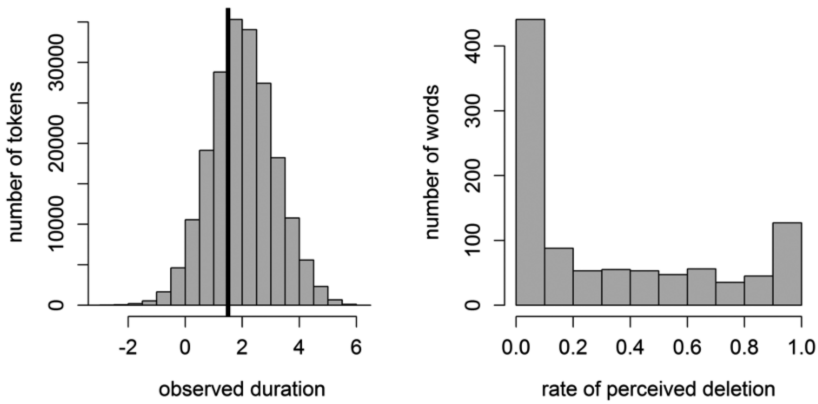


(a) Words' baseline preferences for oral-stop duration

(b) Tokens' observed oral-stop durations

To convert these observed durations into observations of deletion or non-deletion, we can impose a duration threshold under which the oral stop will be perceived as deleted, as shown in (24)a, which repeats (23)b, adding a line at 1.5 to mark the chosen threshold. The result in this case is that, although the tokens as a whole are distributed in a bell curve, most words' tokens are mostly on just one side of the threshold. For example, if a word's baseline preferred duration is 2.5, with the small amount of noise we added, only about 2% of its tokens will fall below the threshold of 1.5, and thus this word will contribute to the leftmost bin of the histogram in (24)b (consistent non-deletion).

(24) Thresholded perception of stop deletion



(a) (23)b, with threshold shown (b) Histogram of words' deletion rates

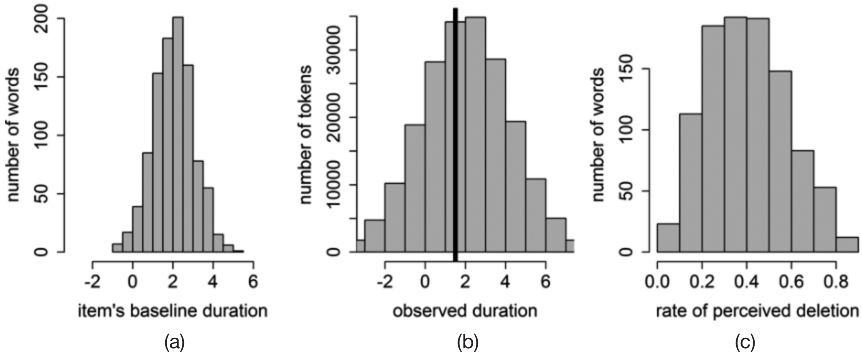
If listeners and learners had a threshold along these lines for deciding whether deletion had applied in words like [mambili] ~ [mamili], then each word would have a tendency to get lexicalized either with or without deletion, rather than as varying.

The scenario proposed for nasal substitution seems very much like English t/d deletion, though, which plausibly results from overlap by neighboring consonant gestures, and/or shortening of the t/d gesture itself.[3] Why is the distribution of t/d deletion in the Buckeye corpus unimodal then, instead of U-shaped?

The U shape in (24) depended on having the noise added to durations in production be small relative to how spread out the baseline duration preferences were. If noise is bigger, or if the baseline durations are clustered closer to the threshold, then things change. The plots in (25) show what happens if we repeat the simulation but multiply the Gaussian noise by 2 instead of dividing it by 2. In (25)a we have the distribution of baseline preferred duration (as for nasal substitution, mean is 2, standard deviation is 1); (25)b shows the distribution of observed durations, with the threshold again marked at 1.5. The distribution is more spread out in (25)b than it was in (23)b/(24)a, because we have added more noise to each token. Finally, (25)c shows the distribution of perceived deletion rate per word. It is unimodal, like Coetzee and Kawahara's (2013) t/d data, because most words have a good chance of having a healthy number of tokens on both sides of the threshold. To take again the example of a word with a baseline preferred duration of 2.5, now about 31% of its tokens will fall below the threshold.

3. Compare British English /t,d/, where Temple (2014) documents acoustic continua of release strength, constriction degree, and glottalization, among others, though duration was not examined; or Mitterer & Ernestus's (2006) similar findings for Dutch /t/ (gradience in amplitude reduction as compared to surrounding sounds).

(25) Larger noise creates a unimodal distribution



(There is at least one example of an English word whose *t*-deletion rate is so high that it has been lexicalized by many speakers without the *t*: [tɛks(t)] 'text', in the sense of 'mobile-phone text message' or 'to send a text message'. For these speakers, we can tell that the underlying form is /tɛks/ rather than standard /tɛkst/, because the plural or third-person singular is [tɛksɨz] rather than standard [tɛks(t)s], and the past tense is [tɛks(t)] rather than [tɛkstɨd].)

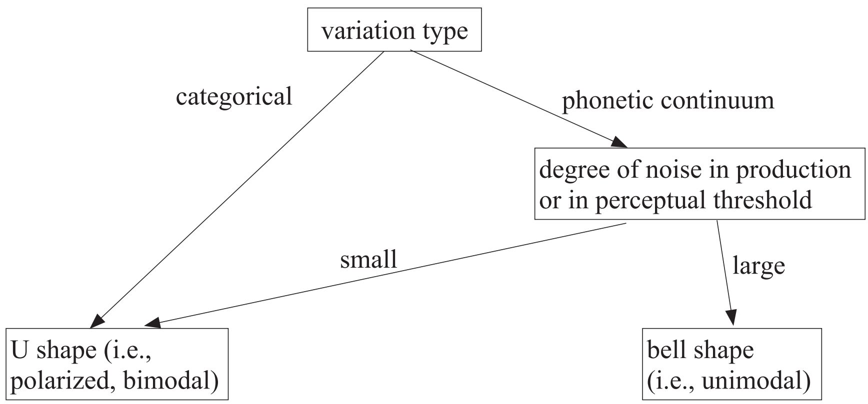
We can thus get either a unimodal or a bimodal distribution from the same basic scenario, by changing how much variation there is in production. We could get the very same result by varying how much noise there is in the perceiver or transcriber's threshold. That is, instead of holding the threshold constant at 1.5, we could allow it to vary from token to token; the more it varies, the more chance a word's tokens have of falling on the other side of the threshold, and the more unimodal the distribution will be.

3.4. Summary of models

This section has presented two ways to achieve a bimodal (polarized, U-shaped) distribution of item behavior, and one way to achieve a unimodal (bell-shaped) distribution. The decision tree in (26) schematizes the decision.

The first branch of the tree asks what the nature of the variation is. In some cases it is clearly categorical, such as in English binomials: there is no phonetic continuum from *fruits and vegetables* to *vegetables and fruits*; similarly for Tagalog noun/adjective order and French *de* omission (though in that last case there could in principle be a continuum). In other cases, the phonetics are unclear or unknown. As far as I know, Tagalog nasal substitution is categorical—the oral stop is either present or not—but this has not been studied rigorously. Kevin Ryan examined tapping in a corpus of Tagalog (Guevara et al. 2002); he found that there were two sharp categories along the duration dimension: [ɾ] (short duration) and [d] (long duration), with very few tokens in between, so Tagalog tapping is probably on the categorical branch (Zuraw & Ryan 2007). Baroni (1998) finds that, on both acoustic and electroglottographic grounds, Northern Italian [s] and [z] form two separate categories. As mentioned above, English t/d deletion probably belongs on the continuous branch. Tagalog o/u variation may also belong on this branch; anecdotally, speakers and transcribers are often unsure whether a Tagalog vowel is [o] or [u]. Detailed articulatory study may even reveal a mixed pattern: Ellis & Hardcastle (2002) found two English speakers who produced an articulatory continuum from no place assimilation (*re[d] coat*) to full assimilation (*re[g] coat*), but also two who produced only "full alveolars or complete assimilations" (p. 373).[4]

(26) Summary of models



4. Summary and conclusion

Section 2 presented several cases of U-shaped variation: a tendency for each item to show a consistent behavior rather than an intermediate rate of variation. That is, most of the variation here was type variation, though all cases did have at least some admixture of token variation.

Section 3 showed that listing of two options produces a U-shaped distribution, if each item moves from one variant to the other abruptly. When the two variants are categorically different, we saw a plausible mechanism for abrupt change. One consequence of this model was that it required us to posit listing for fairly long units, such as French *de X et/ou (de) Y* sequences or Tagalog adjective/noun pairs (or English binomials, where listing is widely acknowledged).

When variation is along a phonetic continuum, two outcomes were derived. If noise in producing tokens is small (compared to the variance of baselines across items), the result is U-shaped, but if token noise is large, the result is bell-shaped.

4. In Bybee's (2012) classification of diachronic changes as *sound change* or *phonological change*, changes that are articulatorily gradual will eventually become lexically regular, but while the change is in progress, lower-frequency words will tend to resist it. Changes that are articulatorily abrupt can leave behind high-frequency lexical exceptions. It is hard to say whether the cases reviewed here accord with this typology, because they involve the complicating factor of whether complex structures have become listed as units, which is more likely to occur for more-frequent items.

The general prediction that emerges is that when variation is categorical and is over listable units, U-shaped distribution will tend to result. When variation is continuous, either a U-shaped or a bell-shaped distribution is possible.

References

Abeillé, Anne, Bonami, Olivier, Godard, Danièle & Tseng, Jesse. 2004. The syntax of French de-N' phrases. In Müller, Stefan (ed.). *HPSG 2004: Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, 6-26. Stanford: CSLI Publications.

Abeillé, Anne, Bonami, Olivier, Godard, Danièle & Tseng, Jesse. 2005. Les syntagmes nominaux français de la forme de-N'. *Travaux de linguistique* 50: 79-95. <http://dx.doi.org/10.3917/tl.050.0079>

Abeillé, Anne, Bonami, Olivier, Godard, Danièle & Tseng, Jesse. 2006. The syntax of French à and de: an HPSG analysis. In Saint-Dizier, Patrick (ed.). *Syntax and semantics of prepositions*, 147-162. Dordrecht: Springer.

Baker, Rachel, Smith, Rachel & Hawkins, Sarah. 2007. Phonetic differences between *mis-* and *dis-* in English prefixed and pseudo-prefixed words. *Proceedings of ICPhS XVI*, 553-556. Saarbruecken.

Baroni, Marco. 1998. The phonetic nature of the Northern Italian allophones [s] and [z] in words with variable realization: electroglottographic and acoustic evidence. *UCLA Working Papers in Phonetics* 96: 166-174.

Baroni, Marco. 2001. The representation of prefixed forms in the Italian lexicon: Evidence from the distribution of intervocalic [s] and [z] in northern Italian. In Booij, Geert & van Marle, Jaap (eds.). *Yearbook of Morphology 1999*, 121-152. Dordrecht: Springer.

Beddor, Patrice Speeter & Chutamanee Onsuwan. 2003. Perception of prenasalized stops. In Solé, M. J., Recasens, D. & Romero, J. (eds.). *Proceedings of the 15th International Congress of the Phonetic Sciences*, vol. 1, 407-410. Barcelona: Universitat Autònoma de Barcelona.

Benor, Sarah & Roger Levy. 2006. The Chicken or the Egg? A Probabilistic Analysis of English Binomials. *Language* 82(2): 233-278. <http://dx.doi.org/10.1353/lan.2006.0077>

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117(12): 1989-2054. <http://dx.doi.org/10.1016/j.lingua.2006.11.004>

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Clayards, Meghan, Hawkins, Sarah & Gaskell, Gareth. 2012. Listeners decode acoustic-phonetic cues to morphological structure. Presentation at *AMLaP*. Riva del Garda.

Coetzee, Andries W. & Kawahara, Shigeto. 2013. Frequency biases in phonological variation. *Natural Language & Linguistic Theory* 31(1): 47-89. <http://dx.doi.org/10.1007/s11049-012-9179-z>

Cohen, Marcel. 1963. *Nouveaux regards sur la langue française*. Paris: Editions Sociales.

Cornulier, Benoît de. 1981. H-aspirée et la syllabation: Expressions disjonctives. In Goyvaerts, Didier L. (ed.). *Phonology in the 1980s*, 183-230. Ghent: E. Story-Scientia.

Damourette, Jacques & Pichon, Édouard. 1930. *Des mots à la pensée: essai de grammaire de la langue française*. Vol. 1. Paris: Éditions d'Artrey.

Ellis, Lucy & William J. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics* 30(3): 373-396.
<http://dx.doi.org/10.1006/jpho.2001.0162>

English, Leo. 1986. *Tagalog-English dictionary*. Manila: Congregation of the Most Holy Redeemer; distributed by Philippine National Book Store.

Frei, Henri. 1929. *La grammaire des fautes: introduction à la linguistique fonctionnelle, assimilation et différenciation, brièveté et invariabilité, expressivité*. Paris: Librairie Paul Geuthner.

Gabriel, Christoph & Meisenburg, Trudel. 2009. Silent onsets? An optimality-theoretic approach to French h aspiré words. In Féry, Caroline, Kügler, Frank & van de Vijver, Ruben (eds.). *Variation and gradience in phonetics and phonology*, 163-184. Berlin: Mouton de Gruyter.

Grevisse, Maurice & Goosse, André. 2011. *Le bon usage: grammaire française*. 15th ed. Brussels: De Boeck-Duculot.

Guevara, Rowena Cristina L., Co, Melvin, Espina, Evan, Garcia, Ian Dexter, Tan, Emerson, Ensomo, Ryan & Sagum, Ramil. 2002. Development of a Filipino speech corpus. *3rd National ECE Conference*. Philippines.

Hay, Jennifer. 2003. *Causes and consequences of word structure*. New York: Routledge.

Hilpert, Martin. 2008. The English comparative--language structure and language use. *English language and linguistics* 12(3): 395-417.
<http://dx.doi.org/10.1017/S1360674308002694>

Lin, Yuri, Michel, Jean-Baptiste, Aiden, Erez Lieberman, Orwant, Jon, Brockman, Will & Petrov, Slav. 2012. Syntactic Annotations for the Google Books Ngram Corpus. *Proceedings of the ACL 2012 System Demonstrations*, 169-174. (ACL '12). Stroudsburg, PA, USA: Association for Computational Linguistics.

Malmberg, Bertil. 1943. *Le système consonantique du français moderne: études de phonétique et de phonologie*. (Études Romanes de Lund VII). Lund: Alf Lombard.

Michel, Jean-Baptiste, Shen, Yuan Kui, Aiden, Aviva Presser, Veres, Adrian, Gray, Matthew K., Pickett, Joseph P., Hoiberg, Dale, Clancy, Dan, Norvig, Peter, Orwant, John, Pinker, Steven, Nowak, Martin A. & Aiden, Erez Lieberman. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331(6014): 176-182.
<http://dx.doi.org/10.1126/science.1199644>

Mitterer, Holger & Mirjam Ernestus. 2006. Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics* 34(1): 73-103.
<http://dx.doi.org/10.1016/j.wocn.2005.03.003>

Mollin, Sandra. 2014. *The (Ir)reversibility of English Binomials: Corpus, constraints, developments*. John Benjamins Publishing Company.

Niyogi, Partha. 2009. *The Computational Nature of Language Learning and Evolution*. The MIT Press.

Ogden, Richard, Hawkins, Sarah, House, Jill, Huckvale, Mark, Local, John, Carter, Paul, Dankovičová, Jana & Heid, Sebastian. 2000. ProSynth: an integrated prosodic approach to device-independent, natural-sounding speech synthesis. *Computer Speech & Language* 14(3): 177-210. <http://dx.doi.org/10.1006/csla.2000.0141>

Phillips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. Palgrave Macmillan.

Pierrehumbert, Janet B., Stonedahl, Forrest & Daland, Robert. 2014. A model of grassroots changes in linguistic systems. *ArXiv e-prints* 1408. 1985.

Pitt, Mark A, Dilley, Laura, Johnson, Keith, Kiesling, Scott, Raymond, William, Hume, Elizabeth & Eric Fosler-Lussier. 2007. *Buckeye corpus of conversational speech (2nd release)*. Columbus, OH: Department of Psychology, Ohio State University. <www.buckeyecorpus.osu.edu>.

Reali, Florencia & Griffiths, Thomas L. 2009. The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition* 111(3): 317-328. <http://dx.doi.org/10.1016/j.cognition.2009.02.012>

Richtsmeier, Peter T. 2008. From Perceptual Learning to Speech Production: Generalizing Phonotactic Probabilities in Language Acquisition. PhD dissertation, University of Arizona.

Schachter, Paul & Otanes, Fe T. 1972. *Tagalog Reference Grammar*. Berkeley, CA: University of California Press.

Shih, Stephanie. 2014. Towards optimal rhythm. PhD dissertation, Stanford University.

Shih, Stephanie & Zuraw, Kie. Submitted. Phonological conditions on variable adjective-noun word order in Tagalog.

Shih, Stephanie & Zuraw, Kie. 2014. Phonological factors in Tagalog adjective-noun word order variation. Presentation at the annual meeting of the Linguistic Society of America (LSA). Minneapolis.

Smith, Brian. In preparation. Modeling English comparative change. Manuscript, UCLA.

Smith, Rachel, Baker, Rachel & Hawkins, Sarah. 2012. Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics* 40(5): 689-705. <http://dx.doi.org/10.1016/j.wocn.2012.04.002>

Stevens, Mary & Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1). e003. <http://dx.doi.org/10.3989/loquens.2014.003>

Temple, Rosalind. 2014. Where and what is (t, d)? In Chiara Celata & Silvia Calamai (eds.), *Advances in Sociophonetics*, 97-136. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Tranel, Bernard. 1981. *Concreteness in Generative Phonology: Evidence from French*. Berkeley: University of California Press.

Tranel, Bernard & Del Gobbo, Francesca. 2002. Local conjunction in Italian and French phonology. *Romance phonology and variation: selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000*, 191-218. Amsterdam and Philadelphia: Benjamins.

Tseng, Jesse. 2003. Phrasal affixes and French morphosyntax. In Gerhard Jäger (ed.), *FGVienna: Proceedings of Formal Grammar 2003*, 177-188. Technische Universität Wien.

Wang, William S.-Y. 1969. Competing Changes as a Cause of Residue. *Language* 45(1): 9-25.
<http://dx.doi.org/10.2307/411748>

Yanovich, Igor. 2014. Frequency effects and jumps between grammars in Old High German. Presentation. Presentation at the UCLA phonology seminar, Los Angeles.

Yu, Alan C. L. 2010. Perceptual Compensation Is Correlated with Individuals' "Autistic" Traits: Implications for Models of Sound Change. *PLOS ONE* 5(8). e11950.
<http://dx.doi.org/10.1371/journal.pone.0011950>

Zuraw, Kie. 2006. Using the Web as a phonological corpus: a case study from Tagalog. *Proceedings of the 2nd International Workshop on Web As Corpus*, 59-66. (WAC '06). Stroudsburg, PA, USA: Association for Computational Linguistics.
<http://dx.doi.org/10.3115/1628297.1628306>

Zuraw, Kie. 2009. Frequency influences on rule application within and across words. *Proceedings of CLS (Chicago Linguistic Society) 43*.

Zuraw, Kie. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language and Linguistic Theory* 28(2): 417-472.
<http://dx.doi.org/10.1007/s11049-010-9095-z>

Zuraw, Kie. 2015. Allomorphs of French *de* and their ability to take wide scope in coordination. *Linguistics Vanguard* 1(1).
<http://dx.doi.org/10.1515/lingvan-2014-1017>

Zuraw, Kie & Hayes, Bruce. Submitted. Intersecting constraint families: an argument for Harmonic Grammar.

Zuraw, Kie & Peperkamp, Sharon. 2015. Aspiration and the gradient structure of English prefixed words. *Proceedings of ICPhS XVIII*. Glasgow.

Zuraw, Kie & Ryan, Kevin M. 2007. Frequency influences on phonological rule application within and across words. Presentation at Workshop on Variation, Gradience, and Frequency in Phonology held at LSA Institute, Stanford University.