

# Neutral Models of Sound Change

Andrea Ceolin

Supervisor: Don Ringe

Dissertation proposal committee: Charles Yang (chair), Robin Clark, Rolf Noyer

## 1 Introduction

A fundamental question regarding language is whether it has evolved through adaptation (Pinker & Bloom, 1990, Jackendoff & Pinker, 2005, Pinker & Jackendoff, 2005) or is a byproduct of a single genetic mutation (Hauser et al., 2002, Fitch et al., 2005, Chomsky & Berwick, 2015). The analogy between language evolution and biological evolution goes back to Darwin and has been applied frequently to this domain, in particular as an explanation of language change (e.g. Yang, 2000, Niyogi & Berwick, 2009). However, while there have been many attempts to build evolutionary models that model language change in terms of selection, fewer attempts have been made to formalize neutral models of evolution (cf. Baxter et al., 2009, Kauhanen, 2017, Newberry et al., 2017).

In this work, I address the question of whether sound change is better explained in terms of adaptation for communication purposes or in terms of stochastic processes. The starting points of my dissertation are the last three chapters of Labov (1994), where functionalist explanations of sound change are ruled out on the basis of empirical investigations: as the Neogrammarians put it, sound change is essentially a mechanical and phonetically determined process. The need to write a dissertation on a scientific debate which appeared settled after Labov’s contributions arises after the emergence of a new ‘functionalist’ wave, inspired by information theory and by mixed-effect regression models. The recent works of Piantadosi et al. (2011), Kaplan (2011), Cohen Priva (2012), Graff (2012), Wedel et al. (2013), Bouchard-Côté et al. (2013), Dautriche et al. (2017), and Mahowald et al. (2018) provide arguments in favor of functional pressures on language change, which in some cases (Sóskuthy, 2013, Cohen Priva, 2017) are presented as solutions to the actuation problem in Weinreich et al. (1968).

Two parallel strategies will be adopted to address this question. On the one hand, I will use historical data in every circumstance in which they are needed to determine the behavior of a model of sound change and to test the predictions of existing models. On the

---

<sup>0</sup>Thanks to Ryan Budnick, Ollie Sayeed, Jordan Kodner, Spencer Caplan, David Wilson, Aletheia Cui, Caitlin Richter, Don Ringe, Charles Yang, Meredith Tamminga, Robin Clark and Rolf Noyer for all their comments on various parts of this manuscript and for many inspiring discussions.

other hand, I will develop a null model of sound change, in order to establish a baseline against which claims and results can be evaluated.

An important point in which this work is different from the works cited above is that no theoretical argument or conclusion will be drawn based purely on the results of artificial simulations or statistical correlations. Any time a model produces an interesting line on a graph or a significant p-value, I will investigate the plausibility of the patterns that might have led to that result, and I will carefully check if they match linguistic observations.

This work is organized in the following sections:

- In chapter 2, I propose an artificial neutral model of sound change, to study the effect of regular sound change on a lexicon over time. Building a model of sound change allows us to see how languages would look if sound change were completely independent from any functional factor.
- In chapter 3, I address the question of whether mergers are more common than phonemic splits (Labov 1994:331), and I show that even though this appears to be true, there is another source of sound change, namely contraction, which is at least as widespread as mergers when one looks at historical data.
- In chapter 4, I show that the interaction between mergers, contractions and phonemic splits makes some interesting predictions about how phonemic dispersion in language vocabularies is expected to change over time. The predictions made by the artificial model are compatible with what we see looking at the historical development of Proto-Romance into Italian.
- In chapter 5, I discuss the Functional Load Hypothesis as a possible constraint on mergers, and I test the predictions of a Functional Load model versus a model in which mergers simply result from phonetic confusability.
- In chapter 6, I use historical records to estimate the impact of lexical borrowing and word formation on the lexica of English and Italian, and I discuss how the two phenomena can influence the distribution of phonemes in a language.

## 2 An Artificial Model of Sound Change

Scholars who are interested in testing the hypothesis that languages are optimized for communication typically focus on analyzing synchronic data from modern languages. For instance, Graff (2012) uses evidence from 60 language dictionaries to argue that probabilistic phonological patterns in the lexicon must be explained through communication efficiency. Cohen Priva (2012) uses speech data from American English, Spanish and Indonesian to argue in favor of a correlation between lenition and average phoneme informativity in the lexicon, suggesting that low informativity can trigger sound change. Dautriche et al. (2017) study the phonemic distribution of the lexicon in four languages (English, German, Dutch and French) and find that the phonemic dispersion in the lexicon is much lower than that predicted by an artificial language model trained on the

same lexicon, which suggests that words tend to re-use the same phonemic sequences even beyond the phonotactic constraints of the language.

It is not clear, in most of these works, whether the hypothesis that the observed patterns are the expected outcome of diachronic change has been seriously considered. This is unsurprising: with few notable exceptions (Kroch, 1989, Blevins, 2004), linguistic theory has proceeded for decades without considering the role of diachronic change in explaining synchronic patterns, mainly because the lack of availability of historical data in an accessible and uniform format makes it difficult to address this question.<sup>1</sup>

On the other hand, linguists have been able to formalize sound change explicitly since the 19th century. While the details of specific sound changes might appear intricate, the phenomenon of regular sound change is explicit from the mathematical viewpoint. For this reason, in this chapter I argue that linguists are in a good position to build formal models of sound change, and study their effect over time (see, among others, Morley, 2015, Beguš, 2018 and Kodner & Cerezo Falco, 2018).

**2.1 Types of Sound Change** Among different types of sound changes, one can distinguish changes that involve the pronunciation of specific sounds, and changes which instead affect the phonemic distribution, for instance by creating homonyms or adding some further dimension of contrasts in the phonemic inventory. The latter are the objects of investigation of this project, because they are the ones that have received the larger amount of attention in historical linguistics and the cognitive literature, since their interaction is what causes languages to diverge to the point of becoming mutually unintelligible. Following Hoenigswald (1960), we can divide them into three categories.

The first category is *unconditioned merger*. According to this type of sound change, two different phonemes in a language collapse into one, therefore causing homonymy between the words which could be distinguished only by that contrast. One widespread type of unconditioned merger is the one between the vowel of *lot*, /A/, and the vowel in *thought*, /0/, which affected English varieties spoken in Western and Northern America and in the British Isles. As a consequence of this merger, contrasts like *caught* and *cot* become indistinguishable.

The second category is *primary split*, or *conditioned merger*, where the phenomenon is limited to a specific environment, and therefore no phoneme is lost in the process. A case which can be ascribed to American English is the neutralization of the contrast between /t/ and /d/ between a stressed and an unstressed vowel: in words like *butter* and *wedding*, the two phonemes are perceived as an alveolar flap, /R/. However, the two phonemes contrast outside of this specific context, and therefore they are not involved in an unconditioned merger.

The third category is *secondary split*, or *phonemic split*, or simply *split*, and it describes a situation in which a sound which is originally in complementary distribution with another sound is reanalyzed as an independent phoneme, and therefore it starts appearing outside of its original context, and contrasting with the sound with which it was in alternation.

---

<sup>1</sup>I add to this short list the research program of the Parametric Comparison Method (Longobardi et al., 2013). See Longobardi (2012) for an insight on its implications for syntactic theory and a clear example of diachronic explanation.

A clear example of this phenomenon is the development of Latin /k/ before front vowels, which is the source of Italian /tʃ/ and Castilian Spanish /t/. A crucial property of splits is that they are always dependent on external factors: one of the most common sources of phonemic split is the loss of the conditioning environment due to other sound changes, but another source can be lexical borrowing, which can introduce words with alternations that are not predicted by the original environmental contrasts.

There are other cases of change that I will investigate in the following chapters, but for the moment these three cases will be the starting focus of our model. What I will not investigate is cases of phonetic change that have no consequences for the underlying grammar: cases in which phoneme X is perceived as a new phoneme Y across the board, when the latter was not already present in the language (otherwise it would be a case of unconditioned merger), and cases of ‘chain shifts’, namely cases in which phonemes {X, Y} become perceived as phonemes {Y, Z}, with X being lost, Y replacing X across the board, and Z appearing in all the contexts in which Y was present. If there are traces of X left in some context, this would qualify as a primary split. If Z was already present in the language, this would be a case of unconditioned merger.

Another type of change I will not address is lexical diffusion (Wang, 1969), which has been shown to be a plausible model to describe some sound changes in progress, but has found limited application in language reconstruction and in modeling long-term sound changes.

**2.2 An artificial model of sound change** A naive implementation of a sound change simulation would be a direct application of the Wright-Fisher model (Fisher, 1930, Wright, 1931) on a list of phonemes, which can be obtained just by concatenating the phonemes of a text, with each phoneme representing an haploid individual which reproduces asexually. In such a model, change over time is associated to a random sampling of the entire phoneme populations, which would lead to change in phoneme frequencies over time. Increase and decrease of phonemic frequencies would represent the outcome of mergers and splits (cf. Reali & Griffiths, 2009, Sindi & Dale, 2016, and Ceolin & Sayeed, 2019 for similar models).

A problem with such a model is that it would be hard to interpret, because it would be far from any linguistic intuition: in fact, the target of a regular sound change is not a change of a specific phoneme in a specific word, but a change affecting classes of sounds in specific phonetic environments. Phonemes do not normally change their frequencies over time because the words in which they are contained disappear from the lexicon: we know from linguistic reconstruction that the lexicon can be stable for centuries, while its phonetic properties can change abruptly.

In order to simulate sound change more realistically, we need a lexicon over which mergers and splits are applied. In theory, one could use a real-language lexicon as a starting point, like the English vocabulary. Limiting the model to simple CVC roots will make it simpler to implement and more interpretable. Notice that this would be sufficient to implement syllable structure considerations: we will distinguish between consonants and vowels, and among onsets, nuclei and codas.

Since the English lexicon is full of simple CVC roots, I decided to choose the 20 most frequent ones as the initial state:

- Lexicon: {bad, big, but, can, dad, for, get, god, him, her, job, let, lot, man, mom, not, put, sir, son, yes}

For the purpose of the neutral model, these words do not have any particular phonological representation: the only thing that matters is that different symbols represent different phonemes, and therefore words like ‘for’ and ‘job’ share the same vowel, but contrast in both the onset and the coda. Using English words rather than combinations of other symbols is motivated only as a way to make the identification of the patterns easier, but any other symbolic representation would have worked in the same way. Our alphabet contains 10 possible vowels and 18 possible consonants, for a total of a maximum of 28 symbols:

- C: {m, p, b, f, v, d, t, l, n, r, s, c, k, y, w, g, j, h}
- V: {i, I, e, E, a, @, O, o, U, u}

The next step is to apply mergers and splits to this lexicon in order to simulate sound change. In order to do that, I implemented an algorithm in Python, which applies the following function, MERGE, to the mini lexicon:

1. Select one position in the syllable (onset, nucleus, coda).
2. Select one segment in the inventory of the languages among those available for that syllable position (TARGET\_1).
3. Select a second segment in the inventory that is close in terms of featural distance to TARGET\_1 (TARGET\_2).
4. Segments in the phonetic environment are assigned to either the conservative or the conditioning environment depending on featural considerations.
5. TARGET\_1 becomes TARGET\_2 in the conditioning environment.

The function SPLIT works in a similar way. The only thing we have to do is to add the word NOT after ‘segment’ at point 3, as to obtain ‘a second segment NOT in the inventory’, because phonemic splits add segments to the inventory rather than merging existing segments.

Implementing the notion of feature in the system turned out to be the most delicate point. While generative phonology has provided feature vectors that can be implemented straightforwardly to represent phonemes (Hayes, 2011), the fact that feature hierarchies can change over time (Dresher, 2009) makes them difficult to model. As a starting point, I decided to simply project the consonant and the vowel sets on a arbitrary unidimensional feature space, assigning a single numeric value to each symbol. This is equivalent of projecting the symbols on a line, and will serve the simple purpose of capturing the intuition that some symbols are closer with respect to others. Appendix A contains more detailed

information about how the features are represented and how conditioning environments are defined.

Another delicate design choice is determining the frequency of the MERGE and SPLIT functions. Following Labov (1994: 331), we assume equiprobability between mergers and splits in the first place, even though the next section will discuss in detail whether the equiprobability assumption is plausible or not.

In Figure 1, we display a sample outcome of a first simulation, after a run with three sound changes, for illustrative purposes.

- The first sound change is a merger affecting ‘a’ and ‘o’. The conditioning environment is wide enough that this change affects all the words of the vocabulary with ‘a’ (*bad, dad, man* > *bod, dod, mon*), and therefore reduces the vowel inventory from five to four segments.
- The second sound change is a split of ‘l’ and ‘t’ in onsets before ‘i’. In this case, the change has no effect, because words such as *lip* are not present, otherwise this split would have had the consequence of introducing a new phoneme among the onsets; it would have not changed the phonemic inventory of the language, since the segment created is available in coda position.
- The third sound change is a merger of ‘n’ and ‘s’ in onsets (*not* > *sot*), and also in this case is unconditioned, because it targets all the vowels available. This change has the effect of increasing the number of minimal pairs in the language, by introducing the pairs *lot/sot* and *sot/son*.

Of course, a single run on a limited number of sound changes is not informative. However, by running several simulations and increasing the number of sound changes, we can keep track of the changes in the grammar and in the lexicon and determine their stability in time.

**2.3 Phonemic frequencies** A first and simple application of our model is the study of phonemic frequencies within a language. Tambovtsev & Martindale (2007) show that in 95 languages, within-language phonemic frequencies are distributed according to a power-law distribution. Specifically, the distribution proposed is the Yule-Simon distribution, a superset of the Zipf’s distribution that is found for word frequencies (Simon, 1955). This is an empirical fact that requires an explanation (cf. Sayeed & Ceolin, 2019).

Even though power-law distributions are usually associated with stochastic processes (Baayen, 1991, Banavar et al., 2004), one could also justify them as a byproduct of some functional pressure for optimizing communication (cf. the discussion in Piantadosi, 2014). For instance, Mandelbrot (1953), while reasoning on power law distributions, states that ‘[...] a quite statistical structure, entirely independent of meaning, appears, underlying meaningful written languages. This fact is to be considered as a very strong argument in favour of a thesis that language is a message intentionally if not consciously produced in order to be decoded word-by-word in the easiest possible fashion’. Or again, when describing specifically Yule’s distribution for phonemic frequencies, Tambovtsev & Martindale (2007) suggest that ‘It is interesting that the same equation describes the frequency distributions

```
#####!
#####Language Change is happening!
Conditional merger of /a/ in /o/ after [m p b f d l n s c g j]
Minimal Pairs:11
Phonemes:19
Conditional split of /l/ in /t/ in onsets before [i]
Minimal Pairs:11
Phonemes:19
Conditional merger of /n/ in /s/ in onsets before [i u e o]
Minimal Pairs:12
Phonemes:19
bad->bod
big->big
but->but
can->con
dad->dod
for->for
get->get
god->god
him->him
her->her
job->job
let->let
lot->lot
man->mon
mom->mom
not->sot
put->put
sir->sir
son->son
yes->yes
```

Figure 1: The output of a sound change simulation.

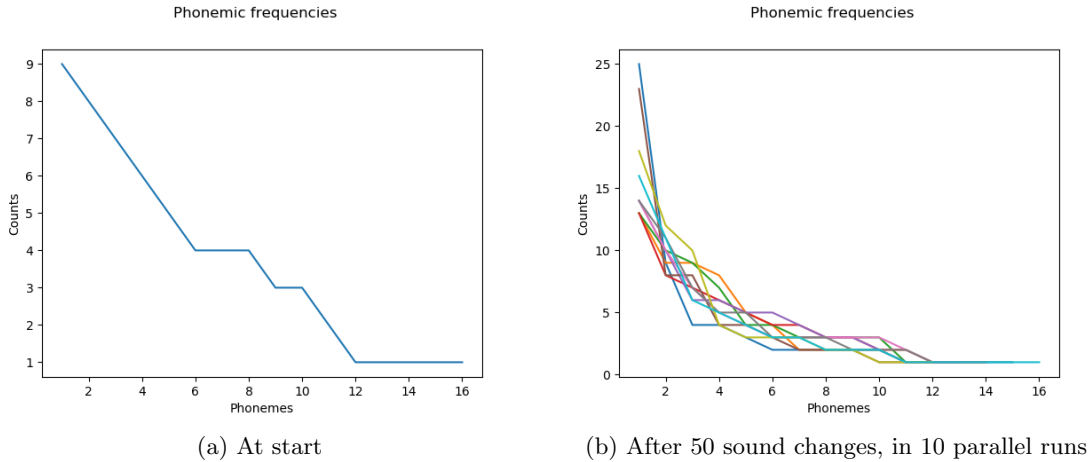


Figure 2: Phoneme frequencies

of DNA codons and phonemes. This may be purely a coincidence. However, it might imply a similarity between linguistic and genetic information transmission’.

With regard to phonemic frequencies, one should make sure that regular sound change, which is the main source of phoneme creation and deletion, is at least considered as a causal explanation, since it is the main process which shapes phonemic distributions over time, as we demonstrate later. For this reason, the relationship between diachronic change and Yule’s distribution is one of the first topics that can be explored through a neutral model of sound change.

Figure 2 summarizes a first experiment using our CVC lexicon and our artificial model of sound change. In Figure 2a, we plot the phonemes of the lexicon in the x-axis (represented by their rank) and their counts on their y-axis. Figure 2b shows that after the occurrence of 50 simulated sound changes, the distribution converges to a power law distribution. The plot displays ten parallel runs.

This is a first example that shows how a neutral model of sound change can derive patterns like Yule’s distribution without appealing to functional explanations.

**2.4 Size of phonemic inventories over time** Labov (1994:331) poses the following question: ‘most reports of phonemic change involve mergers [...] [this fact] would lead to the odd conclusion that most languages are steadily reducing their vowel inventory [...] it stands to reason that just as many phonemic splits must take place as mergers’. An artificial model of sound change can be used to address this question: how does phoneme creation/deletion relate to the relative frequency of mergers and splits?

Figure 3 tracks the number of phonemes in the mini lexicon if we apply 200 sound changes to the initial vocabulary. Surprisingly, we can notice that the equiprobability of mergers and splits is not enough to guarantee a stability of the inventory size: in fact, most of the runs end up with inventories containing fewer phonemes than the initial vocabulary. Chapter 4 will explain why this is the case.



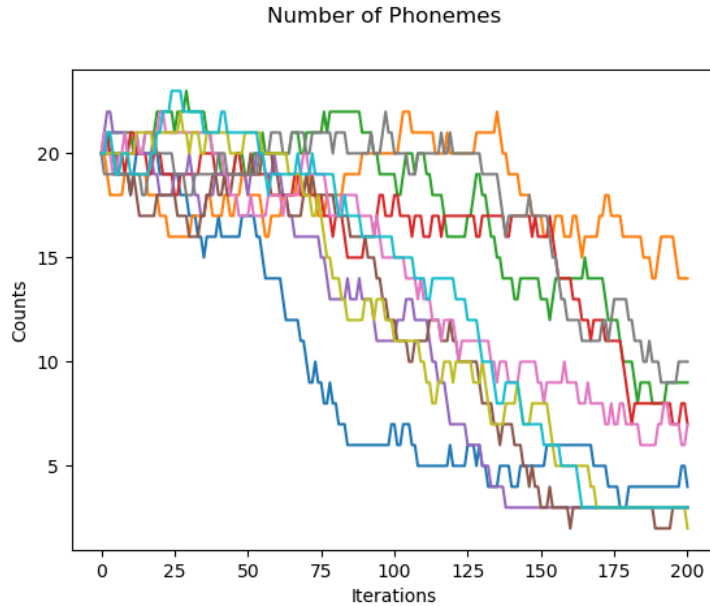


Figure 3: Phoneme inventories over time after 200 sound changes, in 10 parallel runs

**2.5 Summary** In this chapter, I presented an artificial model of sound change which includes all the factors that are known to affect sound change: the presence of mergers and splits, the presence of conditioning environments and the notion of feature to determine the outcome of a sound change. I argued that this model is useful to establish a baseline against which models of sound change can be evaluated. For instance, I showed that power law distributions can be generated by a very simple model of sound change and that the question of the stability of phonemic inventories over time requires some further exploration, given that a simple assumption of equiprobability between mergers and splits does not guarantee the desired outcome of a stability over time. The next chapter will focus on this very last point.

### 3 Mergers and Splits

This chapter addresses the question of how frequent mergers are with respect to splits. Since mergers are associated with phoneme loss while splits are associated with phoneme creation, it is clear that if one phenomenon is more frequent than the other, then a null model could be biased towards phoneme loss rather than phoneme creation, or vice versa. This research question was first addressed in Labov (1994), and returned to the center of the linguistic debate after Hay & Bauer (2007), which suggested a relationship between the size of phonemic inventories and the size of speaker populations, and led to the use of phonemic inventories to model population migrations (Atkinson, 2011, Creanza et al., 2015). In this chapter, I explore the relationship between sound change and the size of phonemic inventories using empirical evidence coming from reconstructed proto-families.

**3.1 The size of phonemic inventories** As mentioned in the preceding chapter, Labov (1994:331) states that even though most of the reported sound changes in progress are mergers, an equivalent number of splits must be in place in order to keep the size of phonemic inventories stable. Labov’s conjecture can be true: for instance, Tse (2016) proposes that splits are more likely to emerge as result of intense contact settings, and the reason why they are underreported is that sociolinguists have not focused on contact areas. Still, 25 years after Labov (1994), the majority of linguists would agree that mergers are still being reported at a higher rate than splits.

This fact does not necessarily imply that in absence of other forces, modern languages would have, on average, smaller phonemic inventories than historical ones: the real question is not whether mergers are more common than splits, but whether phoneme creation via mergers is more widespread than phoneme deletion via splits.<sup>2</sup>

While some readers will find this statement odd, others will note that there has actually been an animated debate about whether this hypothesis is plausible. A *Language* paper published in 2007 by Jennifer Hay and Laurie Bauer noticed that reports of phonemic inventories for languages with large populations contain on average a larger number of phonemes compared to those associated with smaller populations. Even though the claim has been rejected by several studies on statistical grounds (for instance in Donohue & Nichols, 2011 and Moran et al., 2012), the hypothesis was interesting enough to lead Quentin Atkinson to model the size of phonemic inventories using a serial founder effect model (Atkinson, 2011). According to the model, when a group of people leaves a population to migrate, they might develop sound changes that have the effect of reducing their phonemic inventory, in the same way as they might carry only a subset of the genes which characterize the population.

This scenario is surprisingly compatible with the common intuition that mergers, which are the main source of phonemic loss, are the main process active in sound change, and they tend to reduce the size of phonemic inventories over time.<sup>3</sup> In the next section we explore this question by looking at historical data.

**3.2 Estimating the rate of mergers and splits** The relationship between mergers and the size of phonemic inventories, and the possibility of an identifiable diachronic trend, has been tested against historical data only once, in Ringe (2011). In response to Atkinson (2011), Ringe (2011) reports the changes that affected the phonemic inventory of four Indo-European families, Indo-Iranian, Germanic, Greek and Latin, starting from Proto-Indo-European.

Interestingly, there is one process which turned out to have a prominent role in removing and creating segments in Don Ringe’s investigation: a phoneme with multiple places of articulation can be reanalyzed as a sequence of two phonemes (resolution), while a sequence of two phonemes can be reanalyzed as a single phoneme (contraction). Examples

---

<sup>2</sup>I thank Ryan Budnick for suggesting to make this point explicit.

<sup>3</sup>The finding is also compatible with the assumption that mergers spread only if a certain threshold of a population exhibits them (Yang, 2009), because thresholds might be easier to achieve in smaller populations, where individual grammars have a greater weight. Even though, ultimately, the only thing that matters is the exposure of the child to adult speech, a fact which is probably more influenced by social norms than by overall population sizes.

Lineage	Contractions	Borrowing	Splits	Resolutions	Mergers
Proto-Indo-European to Proto-Indo-Iranian	5		3		11
Proto-Indo-European to Proto-Germanic	3		2	4	8
Proto-Indo-European to Greek	14		1	2	12
Proto-Indo-European to Latin	6			6	8
Proto-Indo-Iranian to Sanskrit	1	5	3	2	4
Proto-Germanic to Old English	4		14		8
Proto-Uralic to Proto-Samoyed			2		3
Proto-Uralic to Proto-Finno-Ugric	5				
Proto-Samoyed to Proto-South-Samoyed					1
Proto-Samoyed to Proto-North-Samoyed					1
Proto-Finno-Ugric to Proto-Ugric			1		5
Proto-Finno-Ugric to Proto-Finno-Permic					1
Proto-Japanese to Old Japanese	8				
Proto-Korean to Middle Korean	4				
Proto-Tungusic to Manchu		1	1		2
Proto-Mongolic to Mongolian	13	1	1		
Proto-Turkic to Turkish	1	2			
<b>TOTAL</b>	<b>64</b>	<b>9</b>	<b>28</b>	<b>14</b>	<b>64</b>

Table 1: A summary of processes regulating phoneme creation and deletion.

of these types are the loss of labiovelar consonants in Latin ( $/k^w/$ ,  $/g^w/ > /kw/$ ,  $/gw/$ ) and the creation of vowel diphthongs in Latin and Ancient Greek, respectively. In Ringe’s pilot study, these phenomena were quite widespread. In particular, contractions, namely the reanalysis of particular clusters as independent phonemes, might constitute the force that balances the loss of phonemes caused by the frequency of mergers.

In order to corroborate this hypothesis, we can look at other language families for which historical data is available. A great source that can be consulted is Sammallahti (1988), which contains accurate reconstructions of the phonological systems of Uralic proto-languages, and therefore allows a replication of Ringe’s pilot study. Another source is Robbeets (2003), which reconstructs the phonemic inventories of five language families (Japanese, Korean, Tungusic, Mongolian and Turkic). A summary of the findings can be found in the Appendix B.

The counts for Indo-European (from Ringe, 2011), along with those for Uralic and Altaic, are summarized in Table 1. It is very interesting that the same patterns are confirmed: if anything, contractions are even more widespread in the changes that affected families different from Indo-European. Phonemes created by secondary splits or loanwords are clearly outnumbered by those created by contractions. Resolutions, instead, are not reported.

This empirical fact has some further implications. While I think it would be reasonable to link both secondary splits and phonemes introduced through loanwords with language contact (Tse, 2016), and thus population dynamics, I do not see a clear relation between the development of contractions and external factors. On the contrary, contractions are a prototypical case of sound change entirely driven by acoustic considerations (Ohala, 1983), with no clear connection to population dynamics or functional considerations, like ambiguity avoidance. Overall, if phoneme creation is mostly the result of contractions, it must be essentially a stochastic process, therefore making it harder to link the size of

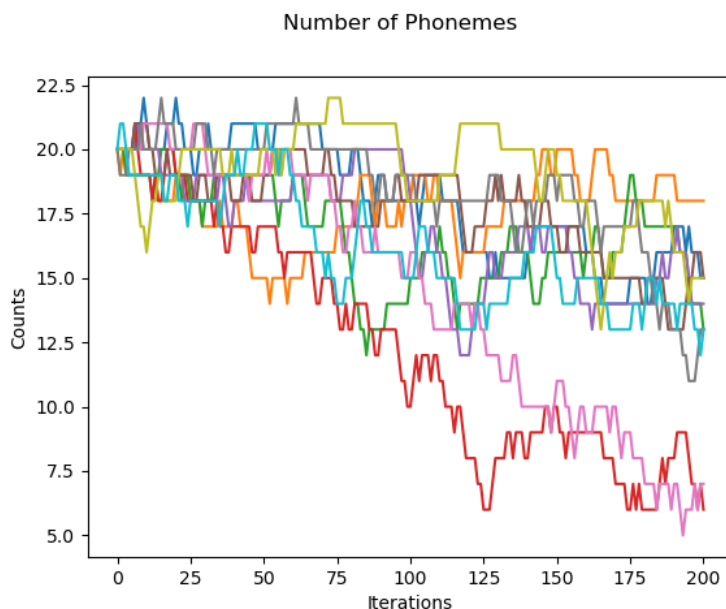


Figure 4: Phoneme inventories over time after 200 sound changes, in 10 parallel runs

phonemic inventories with considerations about the population.

**3.3 Implications for a null model of sound change** The data reported in this chapter show that a model in which mergers and splits are the only forces determining sound change over time would not be adequate to study evolutionary dynamics. A more realistic model needs to include contractions. See Appendix C for the way in which contractions have been implemented in the artificial model.

In Figure 4, we ran the simulation of the preceding chapter after having tuned our artificial model as to include mergers, splits and contractions with a relative frequency determined on the base of our historical investigation, that yields roughly a probability distribution of 0.4 of mergers, 0.4 of contractions and 0.2 of splits. Interestingly, the tendency toward a reduction of the number of phonemes over time does not go completely away: while some languages at the end of the run display a number of phonemes similar in size to the initial stage, the average tendency is still a reduction over time. The next chapter will focus in detail on the reasons why this happens, by addressing the relation between sound change and phonetic dispersion.

**3.4 Summary** In this chapter we corroborated Ringe (2011) in showing that contractions are the main source of phoneme creation, by looking at the historical reconstruction of some additional language families. This settles the question of how it is possible that, given the rarity of splits with respect to mergers, phonemic inventories appear quite stable over time.

## 4 Phonemic dispersion in the lexicon

In the preceding chapter, we determined that mergers and contractions are the main processes through which languages gain and lose phonemes, while secondary splits appear to be less common. In this chapter, we investigate the interaction among these different types of sound change and focus on the predictions that they make about phonemic distributions in the long run.

**4.1 The Production and Perception trade-off** There have been several attempts to explain phonological patterns in terms of communication efficiency (see Graff, 2012, among others). Arguments in favor of some functional pressures on the forces that shape phonological inventories and phonotactic grammars go in two opposite directions: they are either explained as the need to minimize effort on the part of the speaker (‘production-based’, Wright et al., 2004), or the need to minimize ambiguity on the part of the listener (‘perception-based’, Liljencrants & Lindblom, 1972, Flemming, 2004). These two explanations are clearly complementary: it is straightforward that languages with many phonemes distributed in a uniform way would be very efficient from the viewpoint of ambiguity avoidance, but would require a great cognitive effort for language learners and a great physical effort for adults when it comes to producing sounds. On the other hand, languages where phonemes are distributed unevenly, with sequences of sounds that are easy to produce being more frequent than sequences that require more articulatory effort, would be easier to produce, but more likely to produce misunderstanding.<sup>4</sup>

A question that we ask in this chapter is whether sound change is neutral with respect to these two pressures, or whether it interacts with such pressures in any significant, quantifiable way, with consequences for the question of the size of phonemic inventories. For reasons that will be clear in the next section, I will especially focus on minimal pairs.

**4.2 Minimal Pairs in time** Minimal pairs have a special role in both language acquisition and language change. They are needed to learn phonological categories (Cui, 2019) and they can be used to predict the spread of mergers (Yang, 2009, Djärv, 2017). For this reason, the functionalist literature has been using minimal pairs to try to test the hypothesis that the language lexicon is optimized for communication (Graff, 2012, Dautriche et al., 2017). Similar measures that are typically used to quantify phonetic dispersion include number of phonemes, average edit distance, and phonological neighborhood density (Munson & Solomon, 2004), but the use of minimal pairs is the most popular, because of its simplicity and its psychological implications. The argument is that low phonetic dispersion correlates with a high number of minimal pairs, while high phonetic dispersion

---

<sup>4</sup>The problem of ambiguity is overstated in this simplistic formulation: in the real world, speakers can rely on several strategies to disambiguate between alternative readings and retrieve the intended message. This point is usually not stressed enough in the functional literature, probably because it is very hard to quantify: speakers of English know that *cat* and *cap* are two different words, but it would be very difficult to isolate the contexts in which this phonetic distinction is useful at all to retrieve the meaning of an utterance. We will come back to this point in the following sections.

correlates with a low number of minimal pairs.<sup>5</sup>

In order to relate sound change and communication efficiency, Dautriche et al. (2017) is an excellent starting point. In this work, the authors are interested in the production/perception trade-off, and they design a way to tell apart these two pressures on the phonemic distribution of a lexicon.

Establishing a baseline for minimal pairs is not intuitive, because they are a function of the average word-length and the size of the vocabulary, which might include the presence of inflectional and derivational morphemes. However, Dautriche et al. decide to draw a baseline by taking data from the three languages in CELEX (Dutch, English and German, Baayen et al., 1993) plus French, from Lexique (New et al., 2004), and excluding all the words that contain derivational morphemes. Then, they train various language models using  $n$ -grams. These models are used to generate pseudo-words, using a window of  $n-1$  phonemes as a probabilistic conditional environment (which requires adding START and END symbols to initiate and terminate the sequences). The model that yielded the best results after the evaluation phase was a 5-gram model, with Laplace smoothing ( $c=0.01$ ). In a replication of their model that I coded, the model produced words like *snaked* even though they did not appear in the training data, as an effect of the presence of both *snake* and *naked*. This strategy is adopted to ‘teach’ phonotactics to the model. After the model is trained for each language, they select a subset of real words of a certain length, and they compare it with a subset of artificial words of the same length generated by the model, in order to control for word-length. The results are interesting: they find that the number of minimal pairs in real languages (among other measures) is much higher than that found in the pseudo-vocabulary they generated, and they derive from this that the pressure for phonemic ‘clumpiness’ might be stronger than the pressure for phonemic dispersion in real languages, concluding that languages might be optimized for learnability and production.<sup>6</sup>

It is not clear how to explain the results that Dautriche et al. obtained from a linguistic viewpoint. One possibility is that there are some phonological patterns that the model was not able to reproduce: for instance, by not being conditioned on the following segment, the model cannot tell whether the generated consonant is a word-final or a word-internal consonant. The reasoning can be similar for ‘fossil affixes’, namely suffixes and prefixes which were productive at early stages of the language. However,  $n$ -gram models are usually very effective at learning these patterns, and therefore it is possible that their results are the outcome of some non-trivial process that we investigate in the next section.

**4.3 Diachronic change and phonemic dispersion** Contrary to other works in the literature, Dautriche et al. (2017) acknowledge the possibility that diachronic change might be an explanation for their result: ‘[...] *it is likely that we may have a lot to learn from*

---

<sup>5</sup>One problem with using minimal pairs as a measure of phonemic dispersion is that this measure does not handle homonymy well: a lexicon in which all the words have the same phonemic shape would have no minimal pairs at all. When dealing with homonymy, an edit distance would yield a more interpretable result.

<sup>6</sup>‘*The space of wordforms for Dutch, English, German and French is clumpier than what would be expected by the best chance model, across a wide variety of measures: minimal pairs, average Levenshtein distance and several network properties. The strongest evidence comes from minimal pairs, for which the effect size was quite large.*’ Dautriche et al. (2017:144).

*diachronic data to observe how clumpiness evolves in the lexicon as new words appear in the language. While we discussed the possibility that there are pressures for clumpiness exerting on the lexicon, another possibility is that there are only pressures for dispersion and not clumpiness, but that word coining leads to clumpy initial states.'*

While the results of this work are not easy to interpret, the hypothesis that they present as a possible explanation can be tested.

It is intuitive that unconditioned mergers can introduce minimal pairs in the lexicon: by eliminating a dimension of contrast, they can cause some near-minimal pair to become a minimal pair. The only case in which they can remove minimal pairs is when a contrast is the only contrast between two words: in this case, a minimal pair becomes a pair of homophones. In cases of conditioned mergers, minimal pairs can be both created and lost, because conditioned mergers can remove a contrast in some pairs, but add a contrast in others.

Minimal pairs by definition cannot be created just as a consequence of a phonemic split. Splits are conditional on the phonetic environment: if a split introduces a contrast between two words, then it means that there is already a contrast in the two words to begin with. A split depending on a particular environment would affect the minimal pairs in which that environment is contrastive: if we take Latin, CANTUM [kantum] ('song') and CENTUM [kentum] ('hundred') form a minimal pair, but when word-initial /k/ is palatalized in front of /e/ and /i/ as a result of a split, the two words have a consonant contrast in addition to a vowel contrast, and therefore the minimal pair is lost.<sup>7</sup>

Finally, contractions are a special case of splits: like splits, they can produce minimal pair loss. For instance, minimal pairs involving /sk-/ and /sp-/ words would be lost in any language in which the sound change /sk-/ > /s-/ occurs. However, replacing two phonemes with a new one can increase the number of minimal pairs within the set of words of the resulting new length, similarly to what might happen with deletion. This is the process through which *shy/pie* became a minimal pair in English.

At this point, determining whether regular sound change has an effect of increasing or decreasing the number of minimal pairs in a language becomes an empirical question, given that there are forces which go either way.

However, before going into attested historical changes, a question we might ask is: what are the predictions of a neutral model on sound change in terms of the variation in the number of minimal pairs over time? Given what we have seen in the last section about the number of phonemes over time, we can make a prediction about the behavior of our

---

<sup>7</sup>There is a more intricate case, namely the case of words that contain the conditioning environment that contributed to the allophonic split: in this case, the loss of the environment and the split affect the minimal pair simultaneously. For instance, the contrast between English final /ð/ and /θ/ in words like *bath/bathe* and *breath/breathe* was generated after the loss of a final vowel in the second word of the pairs. This means that these words were potential candidates for becoming homonyms, like *wreath/wreathe*, had they preserved the same vowel. Even though the event that created the homonymy was the loss of the final vowel, this is a case in which the split was involved in the process. Since the artificial model used in this work treats splits and mergers as independent, this process cannot be captured properly, because the model does not differentiate between splits resulting from the loss the conditioning environment of two allophones and splits resulting from other sources, for instance borrowing, contractions, or mergers involving other segments. For this reason, in cases like this one, our model would interpret the minimal pair creation as a consequence of another sound change, namely a contraction or a merger.

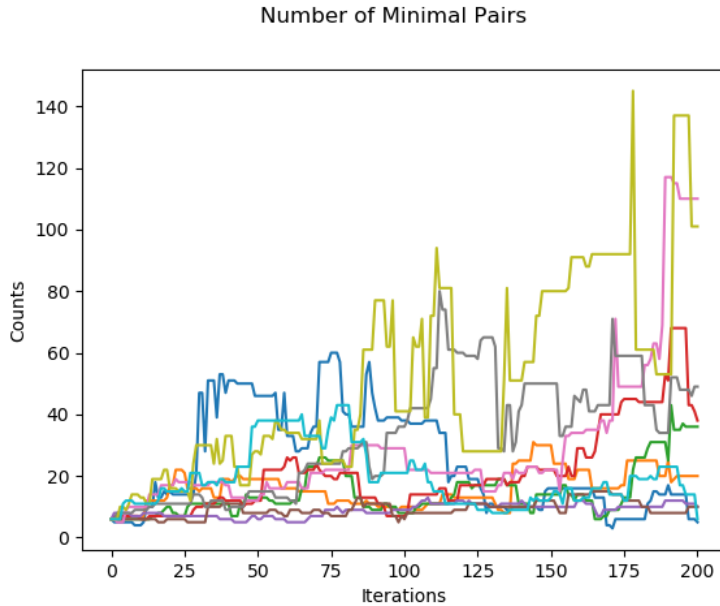


Figure 5: Minimal Pairs over time after 200 sound changes, in 10 parallel runs

artificial model of sound change: if the number of phonemes is reducing, the number of minimal pairs is necessarily increasing.

The simulations exhibit the pattern in Figure 5. We see that the number of minimal pairs tends to increase over time, a fact must be interacting with the decrease in the number of phonemes. Something interesting about this graph is that though on average there is an increase of minimal pairs, in several runs we see some drastic reduction of minimal pairs at certain specific points. This is associated with the creation of homophones: if a merger targets a contrast which is supported by many minimal pairs, and massive homonymy is created, then the number of minimal pairs is drastically reduced. Notice that other measures of phonemic dispersion, like a simple edit distance, would not have this problem, but since minimal pairs are the most typical measure of phonemic dispersion we will stick to them.

Of course, a first interpretation of this result is that we are not modeling sound change in the appropriate way, and therefore there is something in the way that sound change is applied to the lexicon or in the properties of the lexicon itself which is correlated with this behavior. However, the discussion in the next subsection shows that this behavior is the result of very well acknowledged properties of the sound changes described so far.

**4.4 The irreversibility of mergers** By taking a look at the output words of the artificial model, one can clearly identify the patterns which are causing the results of the simulations.

For instance, the pattern in Table 2 is frequent in the simulations. At a certain stage (A) there are no minimal pairs within a subset of the words, but every word shows a contrast of at least two phonemes. Then, a stage (B) is reached, where a merger of the





(A)		(B)	/e/ > /a/ +2 Minimal Pairs (MP)
sad		sad	
set		sat	
far		far	
her		har	
(A)		(B)	-2 MP complicated in a single step

Table 2: Minimal pair creation is not always reversible.

nucleus causes two minimal pairs to appear in the subset. Now, a property of this stage is that this result cannot be reverted in a single step: this is in fact a context where mergers are irreversible. A split of the nucleus conditioned on the onset, for instance, could potentially revert back the minimal pair ‘far’-‘har’ to its previous stage, but it would keep the minimal pair ‘sad’-‘sat’ in the language. The opposite is true for a split of the nucleus conditioned on codas. In the same way, a contraction would not be able to affect both minimal pairs. This is true also for other sound changes not included in our model so far, like deletion or epenthesis. Reverting the merger would require at least two steps. Notice that the same would not be true in case of phonemic splits: since splits are always conditioned on the environment, a conditioned merger can always revert them back to the previous stage.

The irreversibility of mergers is a well-known fact about sound change (Labov 1994:311). While there are historically attested cases of apparently reversed mergers (cf. Chomsky & Halle (1968) and Labov et al. (1991) on the *mate/meat* merger in Early Modern English), scholars agree that these are cases that do not simply involve a regular sound change, but the preservation of the contrast at the phonetic level, which Labov defined as ‘near-merger’. For the purpose of our problem, the irreversibility of mergers creates an asymmetry in the way sound change interacts with minimal pairs: in fact, mergers are expected to create minimal pairs at a rate which is faster than the rate of minimal pair deletion associated with splits and contractions. This has an indirect effect on the global number of phonemes: less variation in possible conditioning environments means less chances of a split.

Of course, minimal pairs can also be deleted after a merger, if the merger results in homonymy, as in Table 3. However, notice that the irreversibility of mergers causes some problems also in this second case. When a pair of words reaches stage (C), there aren’t strategies that can revert this stage back. As a consequence, the words are bound to the same outcomes if their phonemes are affected by sound change. This is another outcome that contributes to reduce the number of phonemes and increase the number of minimal pairs, and that ultimately can lead to a saturation of the whole vocabulary.

This scenario is unrealistic: languages have many ways to disambiguate in cases of massive homonymy, like word compounding, the development of suprasegmental features and the use of synonyms. However, for the purposes of our investigation, it is important to note that sound change is not a neutral force when it comes to the phonemic properties

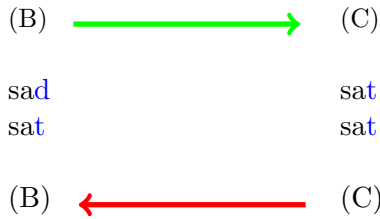


Table 3: Mergers and homonymy

of a lexicon; in fact, we might predict it to have the effect of reducing phonemic dispersion over time, exactly in the direction of the results obtained by Dautriche et al.

In the next section, we evaluate this hypothesis by looking at a case of historically attested language change.

**4.5 Mergers, Splits and Contractions in Proto-Romance** A simple way of testing whether the hypothesis that minimal pairs tend to increase over time is plausible is to collect several matched lists of cognate words at different stages in time for the same language and compare them. By limiting the investigation to cognate words, one can therefore study the effect of sound change on minimal pairs.

In order to have a compelling case study, one would need to find a language in which sound changes happened between two documented stages, and at the same time the language is conservative enough that the words between the two stages differ minimally, so that we can track the consequences of every single sound change in terms of minimal pairs.

The Romance data in the PLEDS database (PLEDS, 2017) are an excellent source: they are available in phonological representation and they contain cognate sets for each Romance language, which means that for each word in the modern languages we have the Proto-Romance reconstructed form. Italian is well-represented in the database, and therefore it is a good candidate for a comparative study. Table 4 contains a first summary of the entries.<sup>8</sup>

In order to make sure that the process behind minimal pair creation or loss was transparent, I focused on the set of cognates whose difference in terms of length between the Proto-Romance and the Italian stage is, at most, one segment. Interestingly, about two thirds of the Italian words (1966/2965) have a cognate in Latin of the same length, while if we also consider words whose cognates differ by one segment we obtain a total of 2676 words, which yield 691 minimal pairs in Proto-Romance and 823 minimal pairs in Italian.

---

<sup>8</sup>The main limitation of working with Romance is inflectional morphology. In PLEDS there are masculine and feminine alternations for some nouns, but only if they exhibit interesting semantic differences or they have different outcomes in other Romance languages. Otherwise, the masculine form is used as default. In this case I could not come out with a good strategy to address the issue: the only thing I did was to manually check the number of minimal pairs and to make sure that I was not overcounting the number of minimal pairs that were lost/created between the two stages just because of unreported, but attested gender alternations. For instance, a gender alternation of the kind of Lat. SENIÖRE/SENIÖRA (‘mister’/‘miss’), which constitutes a minimal pair, would be lost in a Romance languages, if by chance only one of the two words is reported, while actually both survived. I made sure that similar cases were not counted as cases of minimal pair lost from Proto-Romance to Italian.

Language	av. length	minimal pairs	words
Proto-Romance	6.91	1262	4189
Italian	6.61	953	2965

Table 4: Data summary

	Lost	Gained
Morphological Change	34	40
Split	25	0
Contraction	12	58
Merger	3	102
h-deletion	6	13
Total	83	213

Table 5: Minimal Pairs gains and losses between Proto-Romance and Italian.

This means that by controlling for vocabulary size, minimal pairs are indeed increasing.

Once obvious mistakes or irregular sound changes are removed, by comparing the number of minimal pairs that were lost from Proto-Romance to Italian to the number of those which were gained, we end up with Table 5. See Appendix D for a list of the specific sound changes involved and for some examples.

The first interesting pattern emerging from this table is that at least in Proto-Romance, morphological change is almost a zero-sum game: one can basically count a similar number of losses and gains. For instance, Italian feminine noun *rapa* ‘turnip’ derives from a Proto-Romance neuter noun RAPU(M), whose nominative plural suffix *-a*, being the same as the feminine suffix, triggered a gender reanalysis. The natural development would have been *\*rapo*, that in Italian would have formed a minimal pair with *ramo* (< RAMUM, ‘branch’), preserving the Proto-Romance minimal pair. However, the noun *rapa* acquired a new form in the regular Italian feminine plural *rape*, which means that now Italian has new minimal pairs formed by this noun and, for instance, nouns like *rame* ‘copper’. While this corpus, by not listing all the possible inflectional forms, is not accounting for all of these cases, the estimate is suggesting that we do not have any reason to believe that changes in one direction are more prominent than changes in the other one.

As expected, splits were the main factor behind minimal pair loss. The cases of minimal loss involved mainly two sound changes: the palatalization of /k/ in front of front vowels (e.g., CANTU [kantu] ‘song’ - CĒNTU [kĒntu] ‘hundred’ > It. *canto* [kanto] - It. *cento* [tʃĒnto]) and the diphthongization of lax mid vowels in stressed syllables (e.g., VĒTARE [vĒtare] - VŌTARE [votare] > It. *vietare* [vjetare] - It. *votare* [votare]).

When it comes to contractions, one can find, as expected, both minimal pair loss and minimal pair creation. PALIA [palja] ‘straw/hay’ and PALMA [palma] ‘palm’ are not minimal pairs in Italian, because the iod changed the preceding /l/ in the first word; however, It. *taglia* [taLLa] ‘size’ and It. *teglia* [teLLa] ‘pan’, which are minimal pairs in Italian, are the result of a contraction of two different consonant clusters (TALIA [talja] and TEGLA [tEgla]). We note that contractions were responsible for minimal pair gains

more than minimal pair losses, and this is not surprising, given that contractions lead to shorter words, for which minimal pairs are easier to find. In addition to consonant clusters with iod, the other case of contraction was the development of the diphthong /aw/ into /o/ (or /o/, depending on stress conditions), which ended up creating many more minimal pairs than those that were lost.

Mergers clearly had the effect of increasing the number of minimal pairs in the language, especially the mergers affecting mid-vowels (e.g., RĒTE [rete] ‘net’ - SĪTE [slte] ‘thirst’ > It. *rete* - *sete*) and consonant clusters (e.g., CATTŪ [kattu] ‘cat’ - PACTŪ [paktu] ‘pact’ > It. *gatto* - It. *patto*). A phenomenon that I decided to include is the deletion of Proto-Romance /h/ in all contexts, which can be considered a special case of merger (between /h/ and /∅/), since it affected many of the minimal pairs.

It is interesting to see that contrary to the intuition that contractions somehow balance mergers when it comes to phoneme loss and creation, when we analyze minimal pairs they instead behave like mergers in promoting minimal pair creation, at least in Proto-Romance: in other words, while they counter-balance the effect of mergers in terms of phoneme creation, they do not seem to counter-balance the effect of mergers in terms of minimal pair loss. Conversely, since they reduce the length of the words, they are more likely to create minimal pairs rather than removing them. For this reason, one might expect to detect an increase in minimal pairs over time even if the number of phonemes appears stable: overall, while the number of minimal pairs might have increased from Proto-Romance to Italian, the number of phonemes had not.<sup>9</sup> This suggests that an average reduction of phonemic inventory over time would be very slow, if anything.

Of course, these results cannot be generalized: one can imagine that languages in which splits outnumber both mergers and contractions and processes like compounding or derivational morphology increase the average length of the words, the number of losses would outnumber the number of gains. However, since in the preceding section we have seen that splits are relatively marked compared to mergers and contractions, it is reasonable to imagine that these languages would be typologically uncommon.

If this hypothesis turns out to be true, it has some implications for the scenario that Dautriche et al. (2017) described: it means that if mergers and splits are the main actors in minimal pair creation and deletion, and the effect of minimal pair creations caused by mergers is more prominent than the effect of minimal pair loss caused by splits, than one might expect to see the number of minimal pairs increasing just as a natural consequence of regular sound change, without appealing to functional explanations of any kind. In Dautriche et al.’s terms, diachronic change creates a pressure for phonemic clumpiness rather than a pressure for phonemic dispersion, but not for the reasons that are found in the functional literature: this just follows mechanically from a neutral model of sound change.

**4.6 Summary** In this chapter, we addressed the question of how sound change relates to the loss and the creation of minimal pairs.

With regard to the number of minimal pairs, we find a match among the results of

---

<sup>9</sup>Italian has 23 consonants and 7 vowels, if we exclude diphthongs, while Proto-Romance had 18 consonants and 9 vowels: actually, the number of phonemes had in fact increased in Italian.

Dautriche et al. (2017), the predictions made by our artificial model, and a first pilot study on Proto-Romance: all cases are compatible with a scenario in which, all things being equal, languages have easier strategies to create new minimal pairs than ways to lose them. If this turns out to be true, it would raise some doubts about the possibility of using low phonetic dispersion to make claims about a functional pressure to optimize languages for learnability and production: an increase of minimal pairs and of phonetic clumpiness is compatible with the predictions of a neutral model of sound change.

Note that this conjecture immediately raises the question of what are the strategies that languages use to avoid the increase of ambiguity and homonymy. In the next section we consider a common argument used to support functional pressures on sound change: namely the fact that mergers do not apply across the board, but are constrained by functional considerations.

**4.7 Future directions** In the development of my dissertation, I will:

- expand the investigation of the sources of phoneme creation by looking at other language families, possibly outside of Europe. Chinese would be a good candidate, because it is known for having developed a high rate of compounds between the Old Chinese and the Middle Chinese period (Sampson, 2015). It would be informative to estimate how this process affected the number of phonemes and the number of minimal pairs in the language.
- extend the analysis conducted on Proto-Romance to other languages. In this case the possible extensions are limited due to the requirements: we need phonological transcriptions, a variety of changes, and the preservation of a large part of the vocabulary. Old English can constitute a possible case.
- compare these findings with those in Graff (2012). In particular, Graff (2012:125) states that the distribution of minimal pairs in English and other languages cannot be derived from sound change only, because they do not simply follow from the overall phonemic distribution of the languages. I will show that a neutral model of sound change can indeed derive the patterns described by Graff.

## 5 The Functional Load Hypothesis

One of the most popular constraints on sound change is the Functional Load Hypothesis, namely the hypothesis that the likelihood of a merger between two phonemes is inversely proportional to the number of minimal pairs in which those two phonemes are contrastive. This hypothesis was formulated in the context of the Prague School (Mathesius, 1929, Jakobson, 1931, Martinet, 1955) and is meant to provide an empirical basis to the idea that sound change is functional: it is the mechanism through which learners remove non-productive distinctions in the lexicon. This hypothesis is coherent with the theoretical assumption that phonemes are defined in terms of their contrasting properties and with models of acquisition (Cui, 2019).

Some attempts have been made to test this hypothesis empirically. In particular, King (1967), citing Martinet, stressed the importance of investigating the hypothesis.<sup>10</sup>

King was the first one to attempt an empirical test of Functional Load. In King (1967), he derives a Functional Load measure through a formula that weights the number of contrasts between two phonemes and their frequency in a text, and applies the measure to some phonemic contrasts at different diachronic stages in four different Germanic languages, in order to compare the measure calculated for the pairs that merged with measures calculated for the pairs that did not merge. The results, however, did not support any Functional Load effect. Other attempts using information-theoretic models (Hockett, 1967, Wang & Bilger, 1967, Surendran & Niyogi, 2006) have confirmed King's findings. However, a number of recent works which looked at a wider sample of languages found a Functional Load effect in various language families (Silverman, 2010, Bouchard-Côté et al., 2013, Wedel et al., 2013, Babinski & Bower, 2018), in contrast with previous findings.

In the next section, we discuss the model and the results in Wedel et al. (2013).

**5.1 Wedel et al. (2013)** Wedel et al. (2013) is the most cited paper in favor of the Functional Load Hypothesis, and it has had great influence on subsequent linguistic work (Sóskuthy, 2013, Kiparsky, 2016, Babinski & Bower, 2018). In their study, Wedel et al. collected a large sample of mergers in different contemporary languages, and divided them according to their language and conditioning environment. As a result, they obtained seven different classes of mergers in English, three for Cantonese, two for Korean, two for Slovak and one each for French, German, Dutch and Spanish, for a total of 18 classes of mergers. There is variation not only in terms of languages, but also in terms of classes represented: the greatest class is the one that contains mergers of British English interdental fricatives (six instances, from Wells, 1982) while seven classes only have one instance (e.g., the *pin/pen* merger, from Labov et al., 2008).

The methodology of the paper is sound. The authors are interested in evaluating the interplay between two different factors, phonemic frequencies and number of minimal pairs, in predicting the realization or the absence of a merger between two phonemes. For this reason, they use a simple Mixed Effect Logistic Regression model.

The number of minimal pairs associated to each phonemic pair is the measure of Functional Load that they test. Choosing a measure for phonemic frequencies is more complicated: one can decide just to take the sum of the frequencies of the two phonemes of the pair, or other measures such as the product, or the maximum. Moreover, one can decide to count frequencies based on tokens in a corpus or types in a dictionary. The way they choose the measure to use is empirical: after different experiments, they found that the frequency of the most common of the two phonemes in the pair is the measure that better fits the model. They found no significant differences in using token or type frequencies.

Since they also note that many pairs show no minimal pairs at all, they also add to the model a dummy variable to encode for the absence of minimal pairs. This is equivalent of

---

<sup>10</sup> *Martinet has cogently and persistently argued that functional load has rich yet unexplored possibilities for the linguist who attempts to plumb the causality of sound change* (King 1967:833).

	Estimate	Std. Error	z value	Pr(>—z—)
(Intercept)	-3.33	0.43	-7.82	<0.001*
MinPairs	-0.34	0.92	-3.65	<0.001*
Phoneme Probability	0.40	0.23	1.77	0.076
NoMP	-0.51	0.50	-1.02	0.306
Phon. prob. by NoMP	1.53	0.72	2.11	0.035*

Table 6: Wedel et al’s Mixed Effect model.

keeping the pairs for which no minimal pairs are present separated in the analysis.

The setup of a Mixed Effect model allows them to add the merger class as a random intercept. The merger classes differ both in terms of the corpora/languages that were used to collect the numbers and the contexts in which they apply: for instance, it is natural to have large numbers by comparing vowel contrasts in a whole corpus and low numbers by comparing only the vowel contrasts that appear before /l/ or /r/. The purpose of having a random effect for merger class is to take into account the variability among classes, making sure that the differences across classes are not influencing the final result.

Of course, one of the main challenges of testing the Functional Load Hypothesis is to prove that any functional load effect is independent from measures of phonetic similarities: the fact that two phonemes which are closer phonetically are more likely to merge than two phonemes which are not is uncontroversial. In order to control for phonetic similarity, Wedel et al. tried to limit the amount of comparable phonemic pairs for each merger class. They basically use two criteria: i) consonant pairs are compared with consonants pairs only, and the same is done for vowels; ii) the distance in terms of phonemic features is constant within the whole class. For instance, in the class that contains the /f/-/T/ merger, we only find contrasts between phonemes that differ for one feature: indeed, it might be informative to ask whether /f/-/v/ and /T/-/D/ are also reported as mergers, but it would not be informative to extend the same question to contrasts like /f/-/k/, because since they contain two different contrasts (continuant and place), we might already expect a merger between them to be unlikely.

The model shows that the number of minimal pairs is the best predictor of a merger (Table 6). Phoneme probability is marginally significant, and only in the absence of minimal pairs: this means that in all the cases in which minimal pairs are not present, a phonemic contrast which involves a frequent phoneme is more likely to merge than a contrast that involves less common phonemes.

This result supports Functional Load. In the next sections, we discuss a couple of potential confounds of the model.

**5.2 Input data** One potential confound of this and other papers is that they require written corpora to estimate counts. One of the most popular sources of texts is the aforementioned CELEX database (Baayen et al., 1993), which contains about 160.000 English wordforms, reducible to about 50.000 lemmas. However, since the Functional Load Hypothesis links language change to acquisition (i.e., mergers result from failing to acquire a contrast), it seems very unlikely that children have access to words like

*abjuration*, *kilohertz* and *phlegmatic* in the construction of their phonemic inventories. For this reason, a question one might ask is whether the effect would still be robust if child directed speech data is used instead.

A natural choice would be to use the list of 551 words employed in Cui (2019), and derived from Carlson et al. (2014), which according to the acquisition literature provides a good estimate of the vocabulary of a 4-year-old English-speaking child. The main problem with this list is that contrary to what is done in Cui (2019), some of the mergers that we are interested in occur in specific conditioning environments, and therefore a short list will create a sparsity problem. For this reason, I decided to replicate the study by using a wider sample of English words extracted from the Providence Corpus (Demuth et al., 2006) in CHILDES (MacWhinney, 2014).

By selecting the words in the mother speech that appear >50 times and removing some repetitions and some proper nouns, I ended up with a list of 1603 words. Wedel et al. reported that the use of token versus type frequencies did not have an impact on the model. However, one can easily see how the choice would have an impact of the evaluation of pairs containing /D/: a token count would skew the frequency of the phoneme because of the presence of function words like *the*, *this* and *that*, while switching to a type count would make it low. Their solution is to remove all function words, but notice that this strategy would basically remove all instances of the phoneme in the analysis, since the phoneme only appears in function words if one focuses on the most frequent words only (this is clear in the analysis in Cui, 2019). As a consequence, I decided to keep the function words in the list, but to consider the type frequency counts. For consistency with Wedel et al.’s analysis, the data were phonemicized using the classic CMU dictionary, developed at the Carnegie Mellon University.

I also decided to replicate the study on German and Dutch. Unfortunately, while it would be appropriate to use CHILDES data for consistency, the CHILDES data for those languages are not phonologically transcribed. As a proxy, I decided to stick to CELEX data (and CELEX transcriptions), but I applied a threshold on the wordlist as to work with the most frequent words only. Applying a filter of >50 occurrences, I obtained a list of 6412 for German. While this is approximately four times bigger than the English list, I decided to keep the whole list, since I expect the database to contain some noise and some redundancy. The Dutch wordlist for CELEX contains higher counts, so that in order to obtain a similar size, one has to raise the threshold to >500. By doing so, I obtained a list of 5280 words for Dutch.

**5.3 Phonetic control** The way in which Wedel et al. control for phonetic similarity, by simply limiting the investigation of potential mergers to phonemes which differ by only one feature, is controversial. The assumption of equal weight among features is not reasonable: while /t/ and /f/ are articulated in the front part of the mouth, /h/ is articulated in the far back, and therefore the assumption that /t/-/f/ and /t/-/h/ share the same phonetic distance is not motivated. Moreover, when testing potential mergers that involve the interdental fricatives, the set of unmerged pairs contains also pairs which do not involve dental fricatives, like /p/-/k/. Technically, both /t/-/f/ and /p/-/k/ differ in only one



feature, but in practice they have different degrees of confusability.<sup>11</sup>

In the absence of robust phonetic metrics, a confusion matrix would be more indicative of phonetic confusability, a point which has also been clearly made in Hualde & Prieto (2014) and Eychenne & Jang (2018). For this reason, I decided to replicate the study by adding a confusion index to the predictors. The confusion matrix employed is the one in Weber & Smits (2003), and the choice was motivated by the fact that the authors tested all the phonemic contrasts using the same conditions. The analysis was replicated using the classic Wang & Bilger (1967) and Hillenbrand et al. (1995), but note that this forced the exclusion of the phonemic pairs that were not tested in these works. For German and Dutch, I use the confusion matrices in Jouvet et al. (2015) and Smits et al. (2003), respectively.

For each merger, combining the confusability indices of a pair of phonemes by simply summing them ( $p$  of perceiving X as Y plus  $p$  of perceiving Y as X) turned out to yield the best model, and so I used their sum in all the conditions.<sup>12</sup>

**5.4 Reported mergers** In order to replicate Wedel et al.’s experiment, I focused on the subset of the data employed which covers mergers in British English, Dutch, American English and German. This accounts for about 50% of the mergers investigated. The mergers are reported in Table 7.

The consonant mergers reported for British English are all taken from the classic Wells (1982), and they involve the two interdental fricatives. The merger of the fricatives with /f/ and /v/, also known as th-fronting, is reported in Wells (1982) as a characteristic of the Cockney dialect. The merger with /t/ and /d/, also known as th-stopping, is reported for Irish English, specifically in the urban dialects of Cork and Dublin. While mergers of the fricatives with /s/ and /z/ are also listed as present in the same source, Wells does not report any location where *think-sink* and *mouth-mouse* are consistently produced as homophones, even though he suggests that in Gaelic English, speakers might use the same pronunciation for the two words. From the discussion, it is clear that this is a substratum effect from Gaelic rather than a regular sound change of an English dialect. This merger is also problematic because it is reported as /D-/ts/ in Wells, with an affricate rather than an alveolar fricative, but as /D-/z/ in Wedel et al.’s paper. For this reason, I decided to exclude it from the investigation.

The vowel mergers for British English were all reported in Wells (1982), and are uncontroversial. The three consonant mergers studied for Dutch are reported in Kissine et al. (2003).

The mergers reported for American English vowels are problematic. The traditional *cot-caught* merger suffers from the problem that according to the phonetic transcription of the CMU dictionary, it contrasts also before /r/, while in this context we know that /O/ is realized as /o/, and then escapes the merger. Since vowel contrasts before /r/ require an independent analysis, I decided to excluded all of them for the minimal pair counts. The other two unconditioned mergers, attributed to Labov et al. (2008), are not reported in the Atlas of North American English. As for the mergers before /l/ and /n/, they

---

<sup>11</sup>Cf. Ceolin (2019) for some alternative phonetic measures.

<sup>12</sup>I thank Ryan Budnick for some useful discussions on possible confusability indices.

are correct, with the exception of a transcription mistake relative to the HULL-HALL merger. The only merger that I temporarily excluded is intervocalic flapping, because the conditioning environment is uncommon in the wordlist I employed.

Finally, the *bären/beeren* merger is another well attested merger that affected German (Wiese, 2000).

British Eng.		American Eng.	
T-f		A-0	LOT-THOUGHT
D-v		*æ-A	
T-t		*0I-Ä	
D-d		ɹ	
*T-s		i-l	HILL-HELL
*D-z		U-u	PULL-POOL
aI-0I	PRICE-CHOICE	U-oU	BULL-BOWL
U@-0:	CURE-THOUGHT	2-A	HULL-HALL (2-0)
l@-E@	NEAR-SQUARE	U-Ä	BULL-HULL
3:-E@	NURSE-SQUARE		
		ɹ	
Dutch		l-E	PIN-PEN
		German	
s-z		e-E:	
f-v			
X-Y			

Table 7: Mergers in Wedel et al. 2013

**5.5 Results** The results of the first replication, using the lme4 package in R, are in Table 8. As we see from the results, the coefficient related to the number of minimal pairs is marginally statistically significant. We might have expected the coefficient to be less significant than that in Wedel et al.’s analysis because we are using only a subset of their data, and therefore the result is consistent with expectations. For the same reason, it is not a surprise that the marginal coefficients associated with the phoneme probability predictors are not significant.

Very interestingly, if we add the confusability index predictor (Table 9), namely the sum of the reciprocal confusability, the results do not change. As expected, the presence of a phonetic control improves the overall analysis (with the AIC decreasing from 130.0 to 119.8), but did not affect the minimal pair coefficient. I also tried to include additional information about phoneme frequencies, by adding as a predictor the probability of the less frequent phoneme of the pair, and the analysis did not change. Changing the English confusability matrices to Wang & Bilger (1967) and Hillenbrand et al. (1995) led to an increase of the AIC score and the p-values, but to a similar pattern. These facts confirm a functional load effect, which is detectable beyond both phoneme frequencies and phonetic

	Estimate	Std. Error	z value	Pr(>—z—)
(Intercept)	-1.48466	0.57321	-2.590	0.00959*
MinPairs	-0.35859	0.17520	-2.047	0.04068*
Phoneme Probability	0.37262	0.36551	1.019	0.30798
NoMP	-0.70144	0.70198	-0.999	0.31769
Phon. prob. by NoMP	-0.09863	0.61097	-0.161	0.87175

Table 8: Mixed-effect Logistic Regression, with Merger Set as Random effect. AIC 130.0

	Estimate	Std. Error	z value	Pr(>—z—)
(Intercept)	-2.38577	0.65193	-3.660	0.000253*
MinPairs	-0.35312	0.16343	-2.161	0.030720*
Phoneme Probability	0.18824	0.46323	0.406	0.684473
NoMP	-0.64017	0.72233	-0.886	0.375479
CI	0.06586	0.01740	3.784	0.000154*
Phon. prob. by NoMP	-0.35478	0.69297	-0.512	0.608677

Table 9: Mixed-effect Logistic Regression, with Merger Set as Random effect. AIC 119.8

confusability.

In order to identify the source of the signal, I decided to rerun the analysis by considering vowels and consonants separately. In both cases, a model with the minimal pair and the confusability coefficients turned out to be the best model in terms of AIC. The results are in Table 10 and Table 11. The comparison between the two tables shows that even though the consonant class is limited to two sets, the two relevant coefficients are still significant. The situation is different for the vowel table: here, the minimal pair coefficient is not significant anymore.

An inspection of the vowel sets shows that several of them are not compatible with the prediction of the Functional Load Hypothesis. For instance, the PRICE-CHOICE merger for British English has four minimal pairs (*tie/toy*, *boy/bye*, *boy/by* and *boy/buy*), which is quite a high number given that we are only calculating them on a restricted subset of 1603 words. Also the NURSE-SQUARE merger is well represented (*her/hair*, *where/wear* and *where/were*). Similarly, one would expect that finding minimal pairs in the HILL-HELL class for American English would be difficult, since we are counting them on a specific phonetic environment and in a restricted list; still, one finds *feel/fill* and *wheel/will*, while for the majority of the vowel contrasts before /l/ minimal pairs are absent. The situation

	Estimate	Std. Error	z value	Pr(>—z—)
(Intercept)	-2.97430	0.84338	-3.527	0.000421*
MinPairs	-0.37884	0.18045	-2.099	0.035782*
CI	0.07983	0.03012	2.651	0.008028*

Table 10: Mixed-effect Logistic Regression, with Merger Set as Random effect, for Consonants

	Estimate	Std. Error	z value	Pr(>—z—)
(Intercept)	-2.97224	0.52261	-5.687	<0.000001*
MinPairs	-0.15511	0.13727	-1.130	0.2585
CI	0.06617	0.02403	2.754	0.0059*

Table 11: Mixed-effect Logistic Regression, with Merger Set as Random effect, for Vowels

is the same for the German merger: five minimal pairs are found, which is exactly the average number of minimal pairs identifiable including all the vowel contrasts.

One can argue that only minimal pairs where the part of speech is the same are relevant, since the others can all be disambiguated from the context. However, there are various problems with this observation: first, this is true also for the number of minimal pairs calculated for the *unmerged* pairs, which represent our baseline, and therefore repeating the analysis and counting the number of minimal pairs only among words of the same syntactic category does not guarantee a better result; second, experimental studies have showed that homophones can influence lexical access even if the ambiguity is between two words of different syntactic categories (Boland & Blodgett, 2001); third, the argument of contextual disambiguation is so general that can be used against the whole functional load literature: in presence of a large context, information-theoretic models and functional load accounts simply do not work, because ambiguity is almost non-existent when the speaker can access contextual information.

For these reasons, we can temporarily conclude that functional load does not really explain the patterns we see for vowels in any clear way. On the other hand, the situation with the consonant mergers is interesting (Table 12). British English is difficult to evaluate: in the case of th-fronting, we have only one homophone, *three/free*, but the contrast is also the most confusable one. However, the second most confusable contrast is /p-/f/, but interestingly this is also the contrast with the highest number of minimal pairs among the pairs reported. A merger between these two sounds has not been reported. The other two cases of high confusability (/D-/d/ and /D-/v/) are instead reported. This looks like a potential case in which functional considerations might be at work.

If the British English case is still unclear, the Dutch case is compelling. Here we are dealing with voicing contrasts, and excluding for a moment the /X-/Y/ contrast, whose phonological reality is not clear, the other four cases of voicing contrasts we are left with are those in the table. On confusability grounds, we would have no reason to disfavor the /t-/d/ or the /p-/b/ contrast, but still they are preserved, while the other voicing contrasts, those involving fricatives, are lost. In particular, the /t-/d/ contrast showed, by far, the highest confusability index according to different combinations of the confusion indices in Smits et al. (2003). For the advocates of Functional Load, this is a great case of minimal pair effect.

As expected, it is hard to point to clear cases in which functional considerations are needed, but the approach of controlling for phonetic confusability and breaking down the data is a step in the right direction. In particular, we have seen that while vowel contrasts do not reveal any clear pattern, in both of the consonant classes there were cases where minimal pairs might have prevented a merger that would have been plausible on phonetic

Language	Merger	MP	CI
English	T-f	1	32.3
English	D-v	0	21.9
English	T-t	4	17.1
English	D-d	2	23.5
English	T-s	4	9.8
English	D-z	0	20.4
English	T-D	0	20.4
English	p-f	6	24.6
English	b-v	0	20.8
Dutch	f-v	1	58.14
Dutch	s-z	0	39.96
Dutch	p-b	8	49.25
Dutch	t-d	15	90.34

Table 12: Consonant mergers

grounds. More empirical evidence is needed in order to formulate a theory, especially given that other case studies reported in the literature (Hualde & Prieto, 2014 and Eychenne & Jang, 2018) have found contradicting evidence.

**5.6 Functional Load and Minimal Pairs** We started this chapter by considering whether possible constraints on mergers would influence the question that we left open at the end of the previous chapter, namely the fact that the number of minimal pairs and the number of homophones might tend to increase naturally over time as a consequence of sound change. The question of the existence of Functional Load was complex, and we concluded that the evidence in favor of it is rather unclear.

However, let us take a step back and assume that Functional Load is indeed true. What implications would it have for a model of language change? In order to answer this question, I modified the code of the artificial model so as to limit the number of mergers to segments which exhibit the lowest number of minimal pairs. This means that so long as there is any contrast in onset, nucleus or coda position for which minimal pairs do not exist, the model will select those two segments rather than two segments whose merger would cause homonymy. This is a deterministic implementation of Functional Load. We can then check if anything substantially changes once mergers are constrained.

Figure 6 shows that this is not the case. Both figures show that in the initial stage, languages behave in a similar way: the number of phonemes is fluctuating around the initial value, while minimal pairs have a slower and regular increase. This is a consequence of the fact that mergers do not produce homonymy at the beginning, because Functional Load blocks mergers between phonemes which have minimal pairs. However, in the long run, they converge to the results that we have seen in the last section: a reduced number of phonemes and a higher number of minimal pairs. This means that even if the Functional Load Hypothesis were true, it would not drastically change the sound change patterns identified in the preceding section: therefore, its effect on the lexicon over time would be

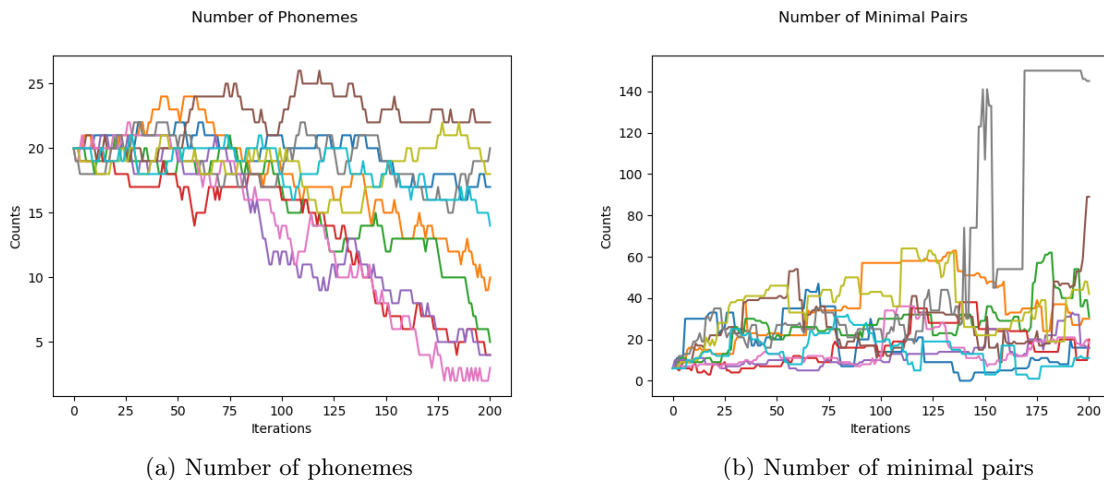


Figure 6: Number of phonemes and number of minimal pairs after 200 iterations, in five different parallel runs.

limited. This also follows from the considerations about the irreversibility of mergers.

**5.7 Summary** In this chapter, we investigated the Functional Load Hypothesis, and we discussed the fact that literature findings do not agree on a single account of the hypothesis. We focused on the model presented in Wedel et al. (2013), which so far represented the best evidence in favor of Functional Load. Replicating the model revealed that the results presented in the paper are valid even if the hypothesis is tested on smaller corpora, which validates the hypothesis. However, we were unable to provide sufficient evidence for the fact that the Functional Load Hypothesis is something unrelated to simple mechanisms like phoneme confusability, especially when it comes to sound changes affecting vowels. Instead, the hypothesis turned out to make plausible predictions on consonant mergers, and more investigation is needed to determine exactly why that would be the case.

Finally, we ended with the suggestion that even if the Functional Load Hypothesis were true, it is not clear under which viewpoint it would make languages more optimal. While its effect in the short run are clear, the effects on the long run might be negligible.

In the next chapter, we focus on two other potential sources that have the effect of changing the lexicon over time: borrowing and word-formation.

**5.8 Future directions** In the development of my dissertation, I will:

- extend the analysis presented in this chapter to a wider number of languages for which confusion matrices are available. In particular, the data already collected for this dissertation would make the extension to the Romance family quite straightforward.
- analyze the mergers in British and American English using the variational model in Yang (2000). The adaption of the model to the spread of mergers in Yang (2009)

and Djärv (2017) is promising for two reasons: first, the model combines together considerations of minimal pairs and phonemic confusability, and therefore can be used to address the study of the consonant mergers we described, since these are cases where both forces seem to be at work; second, contrary to probabilistic Functional Load models, the variational model can be used to determine an upper bound on the spread of mergers. It would be interesting to see if the model can predict, under certain conditions, the impossibility of a merger: previous works have tried to quantify the amount of homonyms tolerated by languages (Köng, 2012, also summarized in Behr, 2015), and even though the variational model was not devised for this purpose, it can be adapted for this task. This would constitute an improvement over classic functional load models, which are stochastic in nature, and therefore unfalsifiable (Sampson, 2015).

- replicate the analysis in Cohen Priva (2017), which presents a correlation between average informativity and lenition as a solution to the actuation problem.

## 6 Borrowing and Word-formation

In this chapter, we focus on two other phenomena which exhibit some interaction with sound change.

First, we focus on borrowing. We address the questions posed by Martin (2007) and Graff (2012) about the cognitive mechanisms behind word borrowing, and test them looking at diachronic data. We conclude that borrowing has a positive effect in terms of increasing phonemic dispersion in the lexicon, but we do not find support for the hypothesis that borrowing is influenced by functional or other considerations, for instance the presence of rare or unfamiliar phonemes in the potential loanword.

Second, we focus on word-formation. It is well known that compounding is a strategy that speakers can use to avoid ambiguity: in regions affected by the *pin/pen* merger, speakers can refer to the second word as *ink pen* in order to disambiguate, while a similar case has been made for Chinese; see Köng (2012), Sampson (2015) and Behr (2015) for a brief discussion on homophony, compounding and tonogenesis. However, I raise some doubts on word-formation as a regular strategy to avoid ambiguity, because an analysis of neologisms in English and Italian shows limited presence of neologisms among the most frequent words, and in general a limited contribution to the phonemic distribution in the vocabulary. For this reason, word-formation appears more as a last resort strategy to avoid ambiguity, rather than a regular property of lexical change.

**6.1 Borrowing** The impact of loanwords on the phonemic distribution of a lexicon appears as an intractable problem, given the nature of the phenomenon: historical events can lead some languages to radically change a great portion of their vocabulary in a short time, or to be virtually immune from loanwords for geographical and sociopolitical reasons. However, borrowing is not conceptually different from population migration, a phenomenon that biologists have extensively studied. For instance, it has been proven that a very low amount of migration can prevent genetic drift between two populations

(Frankham et al., 2002).<sup>13</sup> In this section, we will draw a parallel between the process of migration and borrowing.

There is at least one work which has tried to study the relationship between sound change and lexical borrowing. Martin (2007) aims at building an artificial model of language change, similarly to what we have done here. Rather than focusing on sound change, Martin proposes a model in which lexical change is driven by synonymy resolution: given that speakers are often faced with several options to refer to objects and concepts, there might be a pressure for selecting some words over others. His model follows the spreading activation model in Dell (1986), according to which lexical access involves the activation of the phonemic representation of the word. According to this model, retrieving the meaning of words which contain more frequent phonemes is easier than retrieving the meaning of words containing rare phoneme sequences. For this reason, speakers are inclined to resolve synonymy by selecting words whose phonemic sequences are more common, and remove from use words that contain less frequent phonemes.

This model resembles a classic rich-get-richer model (Simon, 1955), a model through which frequent elements tend to increase their frequency over time at the expense of less frequent elements, which slowly lose probability mass. Therefore, it is not surprising to see that Martin replicates all the results produced by our artificial model: his model can give Yule's phonemic distributions over time (Chapter 2) and also predicts homonymy in the long run (Chapter 4). For this reason, Martin is also interested in possible constraints on sound change. One source of stability he proposes is an arbitrary lexical weight, which causes some words to be more resistant to change in spite of their phonological representation; this is a technical implementation of the notion of basic vocabulary. A second source of stability is introduced through a phonetic bias: since some sounds are attractive because of phonetic reasons, if they happen to have a low frequency in a language, speakers will favor them whenever they have a chance to do so.

This model makes two different predictions when it comes to borrowing. On the one hand, it predicts that speakers should be biased towards borrowing words which contain familiar phonemes. On the other hand, they should be also biased towards borrowing words which contain phonetically unmarked phonemes, i.e., phonemes which are easy to produce and perceive. The most striking example in support for this view is the fact that Latin borrowed a considerable number of 'b-' words from Greek (35) with respect to, for instance, the number of 'd-' words (15), even though in Greek 'd' was twice as frequent as 'b' as a word-initial consonant. The trend continued in the Romance languages: in particular, French retained a large part of Latin's 'b-' words compared to words starting in 'd-', and neologisms in French derived from 'b-' words were more frequent than those derived from 'd-' words. Martin's motivation for this statistical trend is a bias towards 'b' as a phoneme in loanwords, word-retention and word-formation. This prediction is tested in the following sections.

**6.2 English and Italian loanwords** A way to unveil the potential factors influencing borrowing is to take a sample of the vocabulary of a modern language, label the words as native or borrowed, and count their phonemes. In this way, patterns like the Romance 'b-'

---

<sup>13</sup>I'm thankful to Bob Berwick for having pointed this reference to me.



ones are expected to emerge cross-linguistically, if borrowing is biased towards phonetically unmarked sounds. A first experiment is designed for English and Italian.<sup>14</sup> For English, we use a 1000-word list from the Corpus of Contemporary American English (COCA, Davies, 2009), while for Italian we used a 2000-word list from the database SUBTLEX. A bigger list is necessary because, as we will see later, Italian words are redundant for morphological reasons.

One might argue that wordlists of this size cannot be used to properly estimate the contribution of loanwords and word-formation to the vocabulary, because frequent words are more conservative (Pagel et al., 2007). However, since we know that word frequency follows Zipf’s law, a wordlist of the most frequent words will give us a reasonable estimate of the amount of phonemes that children are exposed to in the course of language acquisition. We might expect that a mental representation of a phonological grammar, that includes phonotactic rules and can regulate word-formation at later stages in life, should be fully available at a very young age, and therefore should be detectable by looking at frequent words only.

From the methodological viewpoint, determining what counts as an innovation is a delicate decision. As a starting point, I decided to count as an innovation every word which cannot be tracked to either an Old English (or Latin, for Italian) or a foreign root, results from compounding, or is derived from an existing root through a morphological process.<sup>15</sup> The full list of innovations is available in Appendix E. The results for English are plotted in Figure 7.

A first interesting property that emerges from the pie chart is that the numbers of words which can be tracked back to Old English (red) and the number words which were instead borrowed after the Old English stage (blue) are roughly equal. Even though it is known that the English vocabulary has been highly influenced by historical contacts with Scandinavian, French and Latin populations, here we see that the massive presence of loanwords is also detectable by looking at just the most frequent words. On the contrary, the number of innovations (green) is limited.

The scatter plot projects normalized phonemic frequencies for native words (red dots) and the frequency of those phonemes in loanwords (blue dots), so that we can immediately see if there are any discrepancies.

The highest positive skew is represented by /k/ and /p/, which are not among the most common segments in English. Their presence in borrowed words is related to Grimm’s law: loanwords from other Indo-European families, where voiceless stops remained unchanged, had the effect of increasing their frequency. Interestingly, this did not affect /t/ in the same way. The positive skew in /s/ is a surprise, but can be explained by the fact that it is the most frequent consonant in French (Yamaguchi, 2007), because it derives from both Latin /s/ and from the palatalization of Latin /k/.

In terms of negative skews, they seem to affect many different phonemes, especially those which are absent or rare in many donor languages (/w/, /h/, /ʈ/ and /D/).

On the other hand, we do not see a /b/ bias: even though /b/ is not among the

---

<sup>14</sup>Phonemic transcriptions and labeling were performed with the help of Don Ringe.

<sup>15</sup>The most frequent case in both languages was deverbative nouns: the phenomenon has been productive in both English and Italian.

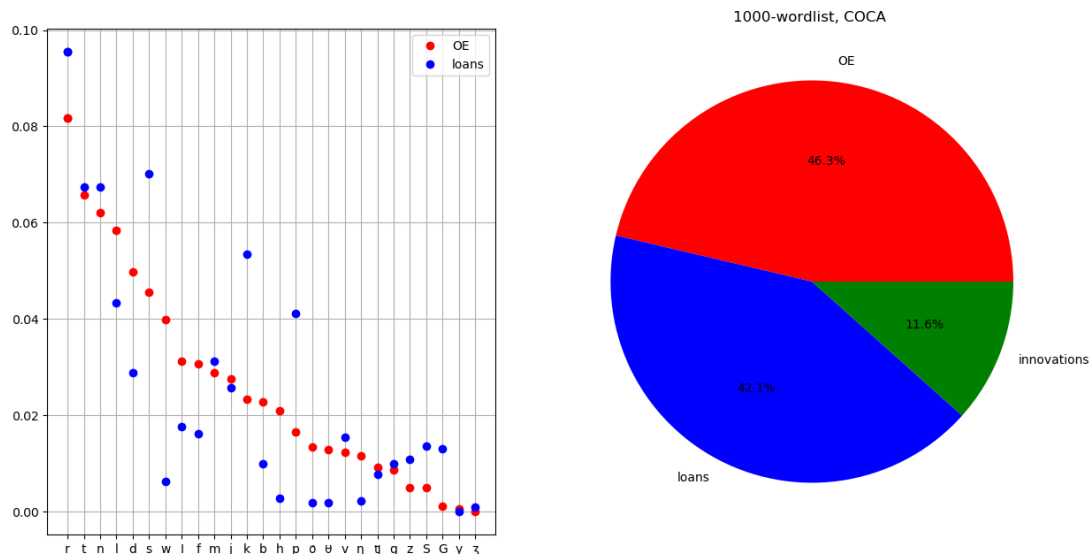


Figure 7: Consonant frequencies for the English COCA 1000-wordlist.

common phonemes in English, we do not find it in many loans, even though we know that Romance languages borrowed ‘b-’ words from their neighbors. A final mention needs to be made for loans containing /z/, /ʒ/ and /ʃ/. Even though these phonemes were rare in the vocabulary, nothing prevented an increase of their presence through loanwords.

So far, none of these patterns seem to require a non-historical explanation, but they can all be predicted from the phonological properties of the donor languages.

We can now check the Italian list. Two caveats with regard to Italian. First, morphology makes the most frequent words redundant in Italian, and since in the case of verbs and nouns we cannot simply remove all the morphological information, the words were reported with their most frequent morpheme in the case of nouns, and in the third person singular in the case of verbs. Second, there were two sources of high redundancy in the Italian wordlist: prepositions combined with articles, and adverbs with the suffix ‘-mente’, from Latin MENS ‘mind’. For these reasons, half of the words in the wordlist were discarded.

Looking at the pie chart, we see that in this case Latin roots account for most of the words, while the rest are equally divided between loanwords and innovations. In the scatter plot, the bias towards /b/ noticed in Martin (2007) is evident. There is not a single source of /b/: they come from both Greek (e.g. *bagno* ‘bathroom’, *ballo* ‘dance’, *barca* ‘boat’) and Germanic (e.g. *banca* ‘bank’, *bianco* ‘white’, *bisogno* ‘need’). We note that also for this case /g/ and /ts/, relatively uncommon phonemes, are less uncommon in loanwords. The case of /g/, in particular, is very interesting: the sources for them are all from Germanic /w-/ words (*guerra* [g<sup>w</sup>Erra] < PG \*werra ‘war’), but instead of borrowing them as they were, Italian reanalyzed them as /g<sup>w</sup>-/ words. This matches the intuition

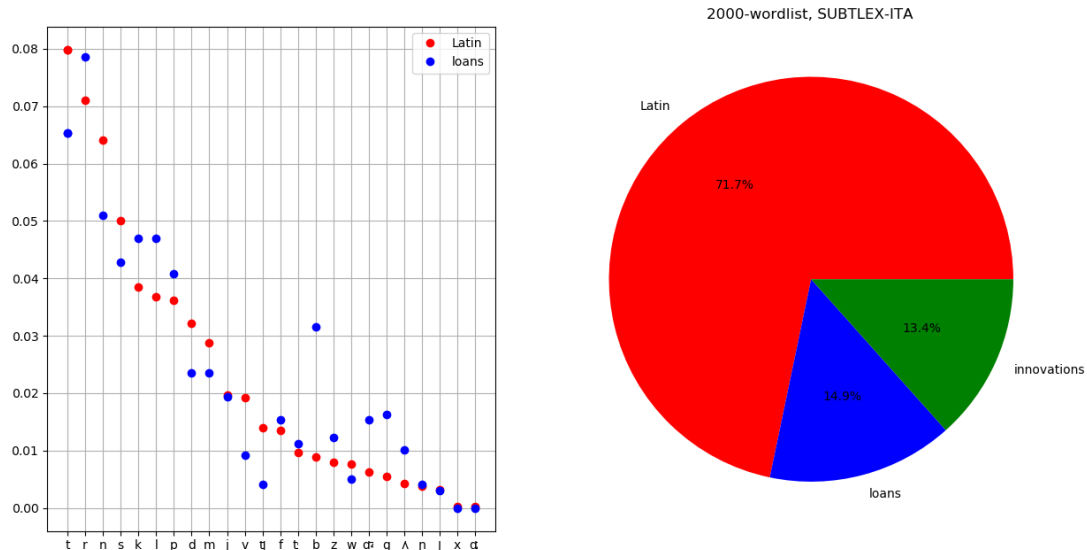


Figure 8: Consonant frequencies for the Italian SUBTLEX 2000-wordlist.

that there is no pressure against borrowing words in the extreme case in which they contain non-native phonemes, because languages have the option of reanalyzing them.

Now, while this first analysis shows that the pattern identified by Martin (2007) is true, namely Italian shows a number of ‘b’ that is clearly above average, it is not clear how much the case can be generalized to other phonemes. It stands out that English borrowed ‘b’ below average, even though the Italian wordlist shows that the phoneme was available in other Germanic languages. All the cases in which English showed a higher presence of certain phonemes than expected are cases that we can explain given the history and the properties of the donor languages. In particular, we do not see any ‘frequency’ pattern: if anything, most of the frequent phonemes were either equally represented or underrepresented in loanwords. On the contrary, in both cases we have seen an increase of relatively uncommon phonemes.

In conclusion, we did not detect any bias in loanwords. Even though this is a preliminary study, the methodology looks like the correct one to investigate potential biases: we might expect different patterns to emerge once more languages are investigated and the donor languages are taken into proper account.

**6.3 Word formation in English and Italian** The analysis proposed in the previous section can be replicated in the same way for word formation. Martin’s model assumes that word formation is constrained in the same way as borrowing. Unfortunately, since word formation is a marginal process in languages, at least when one looks at frequency lists, the data collected so far constitute a small sample. However, they reveal some interesting patterns.

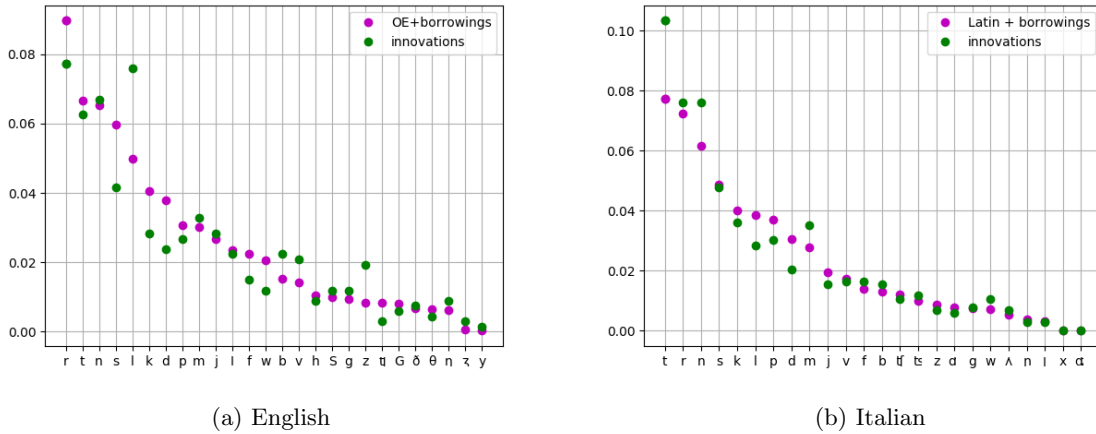


Figure 9: Innovations in English and Italian

In Figure 9 we present two graphs, for English and Italian respectively. The purple dots represent the normalized frequencies derived from the sum of the phonemes projected in the previous figures, namely those belonging to native words and loanwords, while the green dots represent phonemes belonging to innovations.

A first observation is that the differences between the matched dots are quite low, a fact that is surprising, given that the low number of innovations should have led to a higher variance just as a sampling effect.

Even the apparent exceptions have a clear explanation. In the case of English ‘l’, this is the result of two different morphemes, the ‘all-’ prefix (like in *always*, *already*, *although*) and the ‘-ly’ suffix (like in *really*, *actually*, *probably*). Among the phonemes which are underrepresented in innovations we find ‘s’, but this can be easily explained by the high number of recent French loanwords. As for Italian, the productivity of ‘t’ is linked to the use of the past participle suffix *-ato/-ata* in deverbative nouns: as we said earlier, deverbative nouns are one of the main sources of word formation, and therefore we might expect derivational morphemes to be overrepresented in new words.

The ‘b’ pattern detected for French in Martin (2007) is not visible here, which means that the productivity of the phoneme is compatible with its frequency. An inspection of the wordlist, though, shows that the innovations in ‘b’ display at least two clear patterns. One is onomatopoeia (*bambino* ‘baby’ and *bomba* ‘bomb’), and the other is words derived from *buono* ‘good’ and *bello* ‘beautiful’, two adjectives which are particularly productive in terms of word-formation (*bellezza* ‘beauty’, *benvenuto* ‘welcome’ and *buonanotte* ‘good night’). If this small sample is representative of what Martin found, then it could be that the fact that ‘b’ is productive in onomatopoeic words and that, among the words inherited from Latin, at least two were particularly productive in terms of word formation, was sufficient to increase the frequency of the phoneme in a language where it was uncommon. This would explain the productivity of ‘b’ in word formation, but would still leave its productivity in word retention and early Greek loanwords unexplained.

In summary, the data presented in this first pilot study seem to at least point in the same direction: borrowing and word formation do not yield the same patterns. In particular, while borrowing is generally more frequent than word formation and it generally ignores the actual productivity of the phonemes in the language, word formation not only appears to be less frequent, but has a neutral impact on the phonemic distribution of the language, since it is mainly a process of recombination of existing material.

**6.4 Neutral borrowing as a source of stability** One of the motivations behind the model proposed in Martin (2007), and in particular behind the presence of phonetic biases in loanwords, is grounded on the fact that his model of lexical evolution is not stable over time. Since this is also a concern with our neutral model of sound change, we can test how the presence of borrowing interacts with the stability of the system over time, and with the phonemic inventory.

Implementing borrowing in an artificial model is not hard. We can simply implement lexical replacement by removing one word at each iteration of sound change, adding it to a list of ‘potential loanwords’, and then replace the word with a new word, randomly sampled from the ‘potential loanwords’ list. Since this list of ‘potential loanwords’ increases at each iteration, it will be more unlikely over time for the language to borrow a word that was present at the very first iterations. The analogy would be that in a modern language like English, it would not be inconceivable to see words which resemble Germanic forms, which could have just survived in other lineages and entered the language later as loanwords; or, in case some phonemes were particularly frequent in Proto-Indo-European but were lost in English, their presence in neighbor languages would make it more likely for them to be borrowed, rather than phonemes that were never present at earlier stages. Note that this model does sample from a limited domain, in the sense that it would never borrow a word or a phoneme that was never present in the history of the language. While this is not what happens in the real world, this constraint is needed to avoid a scenario in which, by sampling from the entire space of combinations, we have an explosion of phoneme combinations which are completely different from everything the language has ever seen.

The main problem with this phenomenon is that, as we have seen in chapter 3, phoneme creation via borrowing is uncommon. Even though we have some clear cases in which it historically happened, and we have a good understanding of how it can happen (Gylfadóttir, 2018), it would be implausible to design a model of change where borrowing is a frequent source of phoneme creation. If we want to include it, it must be marked.

For this reason, I modified the model described in chapter 4 by including borrowing, and I estimate its frequency of occurrence following the rates in chapter 3. The results of the experiment are in Figure 10.

The results are interesting. After 500 iterations, most of the runs converge on the same number of phonemes, and the number of minimal pairs does not follow a monotonic increase as we have seen in the previous models. Even though the exact behavior of the simulation is hard to interpret, given the presence of different types of sound change interacting with lexical borrowing, it shows us that borrowing can serve as a balancing force. By potentially decreasing the frequency of certain phonemes or increasing the frequency of others, and especially introducing phoneme sequences that were lost in the

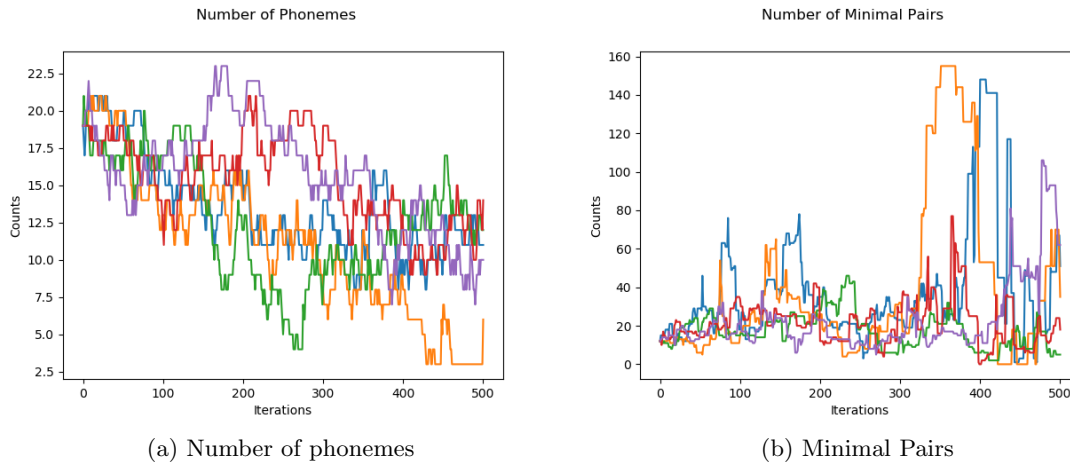


Figure 10: The size of the phonemic inventory and number of minimal pairs after 500 iterations, for five parallel runs.

language, borrowing can guarantee a level of phonemic dispersion that a model which is only based on mergers, splits and contractions cannot guarantee. This fact has a direct analogy in population genetics: low amounts of migration can have a tangible effect on genetic variability over time (Frankham et al., 2002).

This consideration has also some implications for rich-get-richer models like Martin’s, where frequent phonemic sequences tend to increase over time: it means that we do not need to postulate any cognitive bias regulating borrowing in order to avoid lexical saturation, but a very simple, neutral model of borrowing would achieve the same result. This observation provides a computational argument, in addition to those of the preceding section, against the necessity of functional pressures regulating processes like borrowing and word-formation.

Of course, this hypothesis can raise further speculations. For instance, as previously mentioned, it is well known that word compounding is one of the main strategies that can be employed to avoid ambiguity, and this has been proven to be the optimal strategy to increase the amount of information transmitted by languages (Nowak et al., 1999, Plotkin & Nowak, 2000). If what we are arguing here is true, namely that a modest presence of borrowing might be beneficial for the phonemic inventory, then it would follow that languages which are surrounded by different neighbors from which they occasionally borrow words should be less subjected to the development of suprasegmental features and word compounding than languages which are isolated. I will leave this claim here as a speculation, since it is not clear to me how this hypothesis can be tested, but I will report that Trudgill (2002) independently suggested that isolation can be a factor affecting phonological structures. I will leave this hypothesis to further investigation.

**6.5 Summary** In this section, we attempted to model borrowing and word formation and study their relationship with sound change.

The starting point of the analysis was Martin (2007), which represents one of the few attempts to discuss a general model of lexical change. Looking at the English and the Italian vocabulary, we have not found any evidence for clear phonetic biases or clear frequency effects in word borrowing and word formation. However, we think that gathering empirical evidence is the right path in order to address this question.

Moreover, by including borrowing in our artificial model, we showed that a simple model of sound change that includes lexical replacement reaches stability over time with respect to its phonemic inventory and the phonemic dispersion of its lexicon. This suggests that we do not need functional considerations, at least when it comes to borrowing and word formation, to develop plausible models of language change.

**6.6 Future directions** In the development of my dissertation, I will:

- gather more data on borrowing and word formation by expanding the English and Italian wordlists.
- further investigate the patterns of /b/ loanwords in Romance, since no explanation was provided by examining the most frequent words.
- extend the analysis on borrowing and word formation to other languages. Especially, it would be informative to see what happens in languages where word formation is more productive, and see what role compounding plays in resolving ambiguity (for instance, taking a look at Chinese and Polynesian languages) and shaping phonemic distributions over time.

## 7 Conclusion

In this work, I address the question of whether sound change is better explained in terms of functional adaptation or in terms of stochastic processes. Here is a brief summary of the findings:

- In chapter 2, I proposed a neutral model of sound change that can be used to test hypotheses about phoneme and lexical change over time. This model applies mergers and phonemic splits to a miniature lexicon, and can derive attested patterns like Yule's distribution of phoneme frequencies.
- In chapter 3, I try to solve Labov (1994)'s puzzle about the asymmetry in reports of mergers and splits in the literature. I support the thesis that contractions are the main source of phoneme creation, and since they represent a prototypical case of phonetically driven sound change (Ohala, 1983), they cannot be linked to extra-linguistic factors.
- In chapter 4, I address the question of whether we should expect minimal pairs to increase or decrease over time, and I argue that given both evidence from the artificial model and from diachronic data, it is not implausible to expect an increase of minimal pairs over time. This raises some skepticism towards models that link measures of

Date	Work
Summer 2019	Implement a feature geometry for the neutral model
September 2019	Write Chapter 2 and 3
Fall 2019	Expand the analyses in Chapter 4, 5 and 6
February-March 2020	Write Chapter 4, 5 and 6
April 2020	Write Chapter 1 and Conclusions
May 2020	Defend Dissertation

Table 13: Timetable

phonemic dispersion in the lexicon with some pressures for communication efficiency, because this phenomenon can just be derived from a stochastic model.

- In chapter 5, I discuss the Functional Load Hypothesis. I propose that one possible explanation for some contradictory findings in the literature is the lack of control for phonetic confusability. By replicating Wedel et al. (2013), I show that some Functional Load effect is indeed visible in the data they used, but it appears to be limited to consonant mergers; in fact, no Functional Load effect is detectable in vowel mergers. The findings require some more investigation. I end the chapter by suggesting that Functional Load has a negligible impact on the development of phonemic inventories over time.
- In chapter 6, I discuss borrowing and word formation and their effect on language change. First, I raise some skepticism regarding the claim of a cognitive bias towards loanwords, and a preliminary examination of two vocabulary samples from English and Italian does not identify any clear evidence for the claim. Then, a first analysis on word formation suggests that the processes underlying word formation are very different from those underlying borrowing. However, a neutral model of borrowing solves the problem of lexical saturation that emerged both in Martin (2007) and in our findings in chapter 3. Contrary to the solutions advanced by Martin (2007) and Graff (2012), this model does not assume any cognitive bias at any level. This finding suggests that language contact might be a factor which influences phonemic dispersion, a fact that the functional literature has not taken into account so far.

The temporary conclusion is that a neutral model of sound change which is based on mergers, splits, contractions and borrowings can achieve stability in time without the need of postulating any functional factor, and can derive patterns like power law distributions over time and levels of phonemic dispersion which have been attested in real languages just as an effect of regular sound change. Table 13 contains a timetable for the dissertation.



## References

- Atkinson, Quentin D. 2011. Phonemic diversity supports a serial founder effect model of language expansion from africa. *Science* 332(6027). 346–349.
- Baayen, Harald. 1991. A stochastic process for word frequency distributions. In *Proceedings of the 29th annual meeting on association for computational linguistics*, 271–278. Association for Computational Linguistics.
- Baayen, R Harald, Richard Piepenbrock & Rijn van H. 1993. *The CELEX lexical data base*. Linguistic Data Consortium.
- Babinski, Sarah & Claire Bowern. 2018. Mergers in bardi: Contextual probability and predictors of sound change. *Linguistics Vanguard* 4(s2).
- Banavar, Jayanth R, Paolo De Los Rios, Alessandro Flammini, Neal S Holter & Amos Maritan. 2004. Scale-free behavior and universality in random fragmentation and aggregation. *Physical Review E* 69(3). 036123.
- Baxter, Gareth J, Richard A Blythe, William Croft & Alan J McKane. 2009. Modeling language change: an evaluation of Trudgill’s theory of the emergence of New Zealand English. *Language Variation and Change* 21(2). 257–296.
- Beguš, Gašper. 2018. *Unnatural phonology: A synchrony-diachrony interface approach*: Harvard University dissertation.
- Behr, Wolfgang. 2015. Discussion 6: G. Sampson, ‘A Chinese phonological enigma: four comments’. *Journal of Chinese Linguistics* 43(2). 719–732.
- Benzing, Johannes. 1955. *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik*. Mainz.
- Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.
- Boland, Julie E & Allison Blodgett. 2001. Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements. *Journal of Memory and Language* 45(3). 391–411.
- Bouchard-Côté, Alexandre, David Hall, Thomas L Griffiths & Dan Klein. 2013. Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences* 110(11). 4224–4229.
- Carlson, Matthew T, Morgan Sonderegger & Max Bane. 2014. How children explore the phonological network in child-directed speech: A survival analysis of children’s first word productions. *Journal of memory and language* 75. 159–180.
- Ceolin, Andrea. 2019. Significant testing of the Altaic family. *Diachronica* 36:3. 299–336. <https://doi.org/10.1075/dia.17007.ceo>.
- Ceolin, Andrea & Ollie Sayeed. 2019. Modeling markedness with a split-and-merger model of sound change. In Nina Tahmasebi, Lars Borin, Adam Jatowt & Yang Xu (eds.), *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, ACL.
- Chomsky, Noam & Robert Berwick. 2015. *Why only us*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. ERIC.
- Cohen Priva, Uriel. 2012. *Sign and signal: Deriving linguistic generalizations from information utility*: Stanford University dissertation.

- Cohen Priva, Uriel. 2017. Informativity and the actuation of lenition. *Language* 93(3). 569–597.
- Creanza, Nicole, Merritt Ruhlen, Trevor J Pemberton, Noah A Rosenberg, Marcus W Feldman & Sohini Ramachandran. 2015. A comparison of worldwide phonemic and genetic variation in human populations. *Proceedings of the National Academy of Sciences* 112(5). 1265–1272.
- Cui, Aletheia. 2019. *The emergence of phonological categories*: University of Pennsylvania dissertation.
- Dautriche, Isabelle, Kyle Mahowald, Edward Gibson, Anne Christophe & Steven T Piantadosi. 2017. Words cluster phonetically beyond phonotactic regularities. *Cognition* 163. 128–145.
- Davies, Mark. 2009. The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International journal of corpus linguistics* 14(2). 159–190.
- Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological review* 93(3). 283.
- Demuth, Katherine, Jennifer Culbertson & Jennifer Alter. 2006. Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech* 49(2). 137–173.
- Djäv, Kajsa. 2017. Predicting mergers in new dialect formation. *University of Pennsylvania Working Papers in Linguistics* 23(1). 39–47.
- Donohue, Mark & Johanna Nichols. 2011. Does phoneme inventory size correlate with population size? *Linguistic Typology* 15(2). 161–170.
- Dresher, B Elan. 2009. *The contrastive hierarchy in phonology*, vol. 121. Cambridge University Press.
- Eychenne, Julien & Tae-Yeoub Jang. 2018. On the role of contrastivity in the development of the /e/-/E/ merger in Korean. *Journal of East Asian Linguistics* 27. 111–139.
- Fisher, Ronald Aylmer. 1930. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press.
- Fitch, W Tecumseh, Marc D Hauser & Noam Chomsky. 2005. The evolution of the language faculty: clarifications and implications. *Cognition* 97(2). 179–210.
- Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, 232–276. Cambridge University Press.
- Frankham, Richard, David A Briscoe & Jonathan D Ballou. 2002. *Introduction to conservation genetics*. Cambridge University Press.
- Graff, Peter Nepomuk Herwig Maria. 2012. *Communicative efficiency in the lexicon*: Massachusetts Institute of Technology dissertation.
- Gylfadóttir, Duna. 2018. *The effective borrowing of a phonemic contrast*: University of Pennsylvania dissertation.
- Hauser, Marc D, Noam Chomsky & W Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579.
- Hay, Jennifer & Laurie Bauer. 2007. Phoneme inventory size and population size. *Language* 83(2). 388–400.

- Hayes, Bruce. 2011. *Introductory phonology*, vol. 32. John Wiley & Sons.
- Hillenbrand, James, Laura A Getty, Michael J Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of american english vowels. *The Journal of the Acoustical society of America* 97(5). 3099–3111.
- Hockett, Charles F. 1967. The quantification of functional load. *Word* 23(1-3). 300–320.
- Hoeningswald, Henry M. 1960. *Language change and linguistic reconstruction*. University of Chicago Press.
- Hualde, José Ignacio & Pilar Prieto. 2014. Lenition of intervocalic alveolar fricatives in catalan and spanish. *Phonetica* 71(2). 109–127.
- Jackendoff, Ray & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition* 97(2). 211–225.
- Jakobson, Roman. 1931. Prinzipien der historischen phonologie. *Travaux du Cercle Linguistique de Prague* 4. 247–267.
- Jouvet, Denis, Anne Bonneau, Jürgen Trouvain, Frank Zimmerer, Yves Laprie & Bernd Möbius. 2015. Analysis of phone confusion matrices in a manually annotated French-German learner corpus. *Proceedings SLATE 2015, Workshop on Speech and Language Technology in Education*. Leipzig, Germany.
- Kaplan, Abby. 2011. How much homophony is normal? *Journal of linguistics* 47(3). 631–671.
- Kauhanen, Henri. 2017. Neutral change 1. *Journal of Linguistics* 53(2). 327–358.
- King, Robert D. 1967. Functional load and sound change. *Language* 831–852.
- Kiparsky, Paul. 2016. Labov, sound change, and phonological theory. *Journal of Sociolinguistics* 20(4). 464–488.
- Kissine, Mikhail, Hans Van de Velde & Roeland Van Hout. 2003. An acoustic study of standard dutch /v/ , /f/, /z/ and /s/. *Linguistics in the Netherlands* 20(1). 93–104.
- Kodner, Jordan & Christopher Cerezo Falco. 2018. A framework for representing language acquisition in a population setting. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)*, vol. 1, 1149–1159.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language variation and change* 1(3). 199–244.
- Kǒng, Jiāngpíng. 2012. A Study on the origin of Tibetan tones by homonym rate. In Wāng Fēng (ed.), *Yǔyánxué luncóng*, vol. 45, 112–127.
- Labov, William. 1994. *Principles of language change: Internal factors*. Oxford: Blackwell.
- Labov, William, Sharon Ash & Charles Boberg. 2008. *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.
- Labov, William, Mark Karen & Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language variation and change* 3(1). 33–74.
- Liljencrants, Johan & Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4). 839–862.
- Longobardi, Giuseppe. 2012. Homoplasmy or principled explanation? *Parameter Theory and Linguistic Change* 2. 304.
- Longobardi, Giuseppe, Cristina Guardiano, Giuseppina Silvestri, Alessio Boattini & Andrea Ceolin. 2013. Toward a syntactic phylogeny of modern Indo-European languages.

- Journal of Historical Linguistics* 3(1). 122–152.
- MacWhinney, Brian. 2014. *The CHILDES project: Tools for analyzing talk, Volume II: The database*. Psychology Press.
- Mahowald, Kyle, Isabelle Dautriche, Edward Gibson & Steven T Piantadosi. 2018. Word forms are structured for efficient use. *Cognitive science* 42(8). 3116–3134.
- Mandelbrot, Benoit. 1953. An informational theory of the statistical structure of language. *Communication theory* 84. 486–502.
- Martin, Andrew Thomas. 2007. *The evolving lexicon*: University of California, Los Angeles dissertation.
- Martinet, André. 1955. Economie des changements phonétiques. *Berne: Francke* .
- Mathesius, Vilém. 1929. La structure phonologique du lexique du tchèque moderne. *Travaux du Cercle linguistique de Prague* 1. 67–84.
- Moran, Steven, Daniel McCloy & Richard Wright. 2012. Revisiting the population vs phoneme-inventory correlation. In *LSA Annual Meeting Extended Abstracts*, .
- Morley, Rebecca L. 2015. Can phonological universals be emergent?: Modeling the space of sound change, lexical distribution, and hypothesis selection. *Language* 91(2). e40–e70.
- Munson, Benjamin & Nancy Pearl Solomon. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of speech, language, and hearing research* 47.5.
- New, Boris, Christophe Pallier, Marc Brysbaert & Ludovic Ferrand. 2004. Lexique 2: A new french lexical database. *Behavior Research Methods, Instruments, & Computers* 36(3). 516–524.
- Newberry, Mitchell G, Christopher A Ahern, Robin Clark & Joshua B Plotkin. 2017. Detecting evolutionary forces in language change. *Nature* 551(7679). 223.
- Niyogi, Partha & Robert C Berwick. 2009. The proper treatment of language acquisition and change in a population setting. *Proceedings of the National Academy of Sciences* 106(25). 10124–10129.
- Nowak, Martin A, David C Krakauer & Andreas Dress. 1999. An error limit for the evolution of language. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 266(1433). 2131–2136.
- Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech*, 189–216. Springer.
- Pagel, Mark, Quentin D Atkinson & Andrew Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout indo-european history. *Nature* 449(7163). 717.
- Piantadosi, Steven T. 2014. Zipf’s word frequency law in natural language: A critical review and future directions. *Psychonomic bulletin & review* 21(5). 1112–1130.
- Piantadosi, Steven T, Harry Tily & Edward Gibson. 2011. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* 108(9). 3526–3529.
- Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and brain sciences* 13(4). 707–727.
- Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: what’s special about it? *Cognition* 95(2). 201–236.
- PLEDS. 2017. *Etymological Database of Indo-European Languages*. n411.fmphost.com/fmi/webdPLEDS\_ONLINE.

- Plotkin, Joshua B & Martin A Nowak. 2000. Language evolution and information theory. *Journal of Theoretical Biology* 205(1). 147–159.
- Poppe, Nicholas. 1954. Grammar of Written Mongolian. *Porta Linguarum Orientalium, Neue Serie, I* 192(498). 665.
- Reali, Florencia & Thomas L Griffiths. 2009. Words as alleles: connecting language evolution with Bayesian learners to models of genetic drift. *Proceedings of the Royal Society B: Biological Sciences* 277(1680). 429–436.
- Ringe, Don. 2011. A pilot study for an investigation into Atkinson’s hypothesis. *Linguistic Typology* 15(2). 223–231.
- Robbeets, Martine Irma. 2003. *Is Japanese related to the Altaic Languages?*: Leiden University dissertation.
- Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. *The Uralic Languages: Description, History and Foreign Influences. Leiden* 478–554.
- Sampson, Geoffrey. 2015. A Chinese phonological enigma. *Journal of Chinese Linguistics* 43(2). 679–691.
- Sayeed, Ollie & Andrea Ceolin. 2019. ‘Markedness’ is an epiphenomenon of phonetically grounded sound change. In Katherine Hout, Anna Mai, Adam McCollum, Rose Sharon & Matt Zaslansky (eds.), *Proceedings of the Annual Meetings on Phonology*, Washington DC: Linguistic Society of America.
- Silverman, Daniel. 2010. Neutralization and anti-homophony in Korean. *Journal of linguistics* 46(2). 453–482.
- Simon, Herbert A. 1955. On a class of skew distribution functions. *Biometrika* 42(3/4). 425–440.
- Sindi, Suzanne S & Rick Dale. 2016. Culturomics as a data playground for tests of selection: Mathematical approaches to detecting selection in word use. *Journal of theoretical biology* 405. 140–149.
- Smits, Roel, Natasha Warner, James M McQueen & Anne Cutler. 2003. Unfolding of phonetic information over time: A database of Dutch diphone perception. *The Journal of the Acoustical Society of America* 113(1). 563–574.
- Sóskuthy, Marton. 2013. *Phonetic biases and systemic effects in the actuation of sound change*: University of Edinburgh dissertation.
- Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Amsterdam studies in the theory and the history of linguistic science* 4 279. 43.
- Svantesson, Jan-Olof, Anna Tsendina, Anastasia Karlsson & Vivan Franzén. 2005. *The phonology of mongolian*. Oxford University Press.
- Tambovtsev, Yuri & Colin Martindale. 2007. Phoneme frequencies follow a yule distribution. *SKASE Journal of Theoretical Linguistics* 4(2). 1–11.
- Trudgill, Peter. 2002. *Sociolinguistic Variation and Change*. ERIC.
- Tse, Holman. 2016. Contact-induced splits in Toronto Heritage Cantonese mid-vowels. *Linguistica Atlantica* 35(2).
- Wang, William. 1969. Competing changes as a cause of residue. *Language* 45(1). 9–25.
- Wang, WS-Y & Robert Bilger. 1967. The measurement of functional load. *Phonetica*

- 16(1). 36–54.
- Weber, Andrea & Roel Smits. 2003. Consonant and vowel confusion patterns by American English listeners. In *Proceedings of the 15th international congress of phonetic sciences*, 1437–1440. Barcelona, Spain.
- Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2). 179–186.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. *Empirical foundations for a theory of language change*. University of Texas Press Austin.
- Wells, John C. 1982. *Accent in English: An Introduction*. Melbourne: Cambridge University .
- Wiese, Richard. 2000. *The phonology of German*. Oxford University Press on Demand.
- Wright, Richard, Bruce Hayes, Robert Kirchner & Donca Steriade. 2004. *Phonetically based phonology*. Cambridge University Press.
- Wright, Sewall. 1931. Evolution in Mendelian populations. *Genetics* 16(2). 97.
- Yamaguchi, Naomi. 2007. Markedness, frequency: Can we predict the order of acquisition of consonants? In *Lingo 2007, the second oxford postgraduate conference in linguistics*, 236–243.
- Yang, Charles. 2000. Internal and external forces in language change. *Language variation and change* 12(3). 231–250.
- Yang, Charles. 2009. Population structure and language change. Ms. University of Pennsylvania .

# Appendices

## A An Artificial Model of Sound Change

**Featural representation** In order to make sound changes more realistic, we need to implement the level of representation of phonological features in the system. A first simple way of achieving this goal is to assign an index to each symbol, by creating a unidimensional feature vector.

m	P	b	f	v	d	t	l	n	r	s	c	k	y	w	g	j	h
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
i	I	e	E	a	@	O	o	U	u								
0	1	2	3	4	5	6	7	8	9								

Instead of ordering the symbols arbitrarily, I decided to order them in a way to partially capture plausible distinctions of English phoneme, for instance in terms of place of articulation and voicing. Of course, since we are currently working with abstract symbols that do not have a real interpretation, any alternative order would do. In the development of my dissertation, I will design a multidimensional feature system with the goal of approximating the contrasts that distinguish real phonemes.

When a function like MERGER is applied, we can force the requirement that the two segments are adjacent when ordered according to their index: in this case, in an inventory like [‘b’, ‘t’, ‘k’], mergers like ‘b’ (2) and ‘k’ (12) are avoided, as long as other segments in the range 3-13 are available as potential targets. However, ‘t’ can merge with either of them, as long as other closer segments are not present in the inventory. This is compatible with the fact that contrastive hierarchies can be different depending on the number and the type of phonemes in the grammar (Dresher, 2009). Reducing the dimensions of variation to a single one is certainly making the model unrealistic, but it should be sufficient for a first study. The main loss of this representation is that we cannot model featural changes, and therefore we are not accounting for changes affecting multiple segments that share a specific feature.

**Conditioning environment** In order to implement conditional changes, we need to implement the notion of context. Again, while a multidimensional feature vector representation would make the notion of context more realistic, a first approximation of context can be determined by randomly generating a number and using it to create a conditioning environment and a conservative environment. The environments are determined by the direction of the sound change: a change from low indices to high indices will have ‘high’ indices as its conditioning environment, while in the opposite case the ‘low’ indices will act as conditioning environment. As a consequence, the symbols at the borders will be more active as conditioning environments, which is not incompatible with linguistic intuitions. In the case of consonants change, the conditioning environment is represented by the nucleus, while in the case of vowel changes either the onset or the nucleus is selected

as the context.

As an example, a sound change that targets ‘c’ and forces a split into ‘k’ in onsets will work as follows. Let us assume that a random generator picks the number 2, and therefore all the vowels which are mapped to a number less or equal than 2 (‘i’, ‘I’, ‘e’) are selected as one environment, while the vowels which are mapped to a higher number (‘o’, ‘U’, ‘u’) are the second environment. The directionality of the change forces the ‘high’ indeces to be the conditioning environment, while ‘low’ indeces are the conservative environment. This would cause the following changes:

- cen, cet, cot > cen, cet, **kot**

From the formal viewpoint this is the reverse of the process through which Latin /k/ is palatalized in Romance languages. Again, a more precise representation of the features and the implementation of mergers as feature assimilation would lead these functions to mimic realistic phonemic change, but for the purposes of a neutral model of change we will temporarily abstract away from this, and eventually refine the model in a second moment.

**Relative Frequencies of Mergers and Splits and a first simulation** A harder problem is to determine the relative frequencies through which mergers and splits should apply. Sociolinguistic investigations report that the number of mergers is higher than the number of splits (Labov, 1994), and therefore splits should be more marked than mergers, but this assumption is problematic for reasons that are discussed in the paper. As a first tentative model, I implemented the following eight functions:

- Conditioned mergers in onsets (conditioned on nuclei)
- Conditioned mergers in nuclei (conditioned on onsets)
- Conditioned mergers in nuclei (conditioned on codas)
- Conditioned mergers in codas (conditioned on nuclei)
- Conditioned splits in onsets (conditioned on nuclei)
- Conditioned splits in nuclei (conditioned on onsets)
- Conditioned splits in nuclei (conditioned on codas)
- Conditioned splits in codas (conditioned on nuclei)

Ultimately, we would like our model to be robust with respect to these specific parameters, and therefore we will try to determine to what extent varying the relative frequency of mergers and splits and their conditioning environment would cause a change in the predictions. As a starting point, we assume that these functions are equiprobable. The assumption that a change in the nucleus is twice as likely as a change in the onset (or the coda) is needed to avoid a saturation of the conditioning environments: if we condition vowel changes just on codas, it follows that a couple of coda mergers can quickly reduce the variation needed for vowel splits to apply. Note that unconditioned mergers are not



directly modeled, but they constitute an extreme case in which all the present symbols are considered conditioning environment, which is in line with the fact that unconditioned mergers are more common than conditional mergers.

## B Mergers and Splits in Uralic and Altaic languages

**Uralic** The following data summarizes the historical phonological development of Uralic Sammallahti (1988). For the Proto-Finno-Ugric branch, the report stops at the level of Proto-Ugric and Proto-Finno-Permic, because further developments are quite complicated to track.

- **Proto-Uralic to Proto-Samoyed, consonant inventory:** for Proto-Uralic 17 consonants can be reconstructed, but internal reconstruction from Samoyed languages points to only 13 segments. Two phonemes which were probably representing fricatives, /d/ and /d'/, merge in all positions with /r/ and /j/. The same happens with the phoneme /s'/, which merges to the unmarked /s/. The status of Proto-Uralic /x/ is less clear, and therefore we leave it out. Overall, it is clear that at least three phonemes were lost as a consequence of mergers.
- **Proto-Uralic to Proto-Samoyed, vowel inventory:** Proto-Uralic had a eight vowel system, with four high vowels. Two new vowels, /ě/ and /ɘ/, resulted from high vowels splits. So we have splits adding two phonemes to the inventory. The presence of a tentative phoneme /ö/ is also reported, but its sources are not clear.
- **Proto-Uralic to Proto-Finno-Ugric, consonant inventory:** the inventory is stable. Internal reconstruction point to three new phonemes, /š/, /c'/ and /l'/, whose presence is postulated only for words that have no cognates with other branches, and therefore their origin is not clear.
- **Proto-Uralic to Proto-Finno-Ugric, vowel inventory:** contractions are reported from the interaction of Proto-Uralic /x/ with five different vowels, so that the vowel inventory of Proto-Finno-Ugric is reconstructed with 13 total vowels. In this case, contractions resulted in five new phonemes.
- **Proto-Samoyed to Proto-South-Samoyed, consonant inventory:** no changes.
- **Proto-Samoyed to Proto-South-Samoyed, vowel inventory:** a merger affected the two back vowels /ä/ and /å/, therefore yielding only one back vowel for Proto-South-Samoyed. A second merger is reported for /ö/, but since the status of the vowel in Proto-Samoyed is not clear, this merger is discarded. We then consider only one merger for this branch.
- **Proto-Samoyed to Proto-North-Samoyed, consonant inventory:** /c/ and /t/ merge in all environments.
- **Proto-Samoyed to Proto-North-Samoyed, vowel inventory:** no changes.

- **Proto-Finno-Ugric to Proto-Ugric, consonant inventory:** /s/ and /š/ merge into /ʃ/.
- **Proto-Finno-Ugric to Proto-Ugric, vowel inventory:** four of the long vowels (/uu/, /oo/, /11/, /ee/) are merged. Interestingly, the merger between /uu/ and /u/ has the consequence that the allophonic split between /u/ and /ũ/ is phonemicized. Here, we have four mergers and one split in total.
- **Proto-Finno-Ugric to Proto-Finno-Permic, consonant inventory:** no changes.
- **Proto-Finno-Ugric to Proto-Finno-Permic, vowel inventory:** /11/ merge into /oo/.

**Altaic** The following data summarizes the historical phonological development of Altaic Robbeets (2003).

- **Proto-Japanese to Old Japanese, consonant inventory:** The main difference between the consonant inventory of Proto-Japanese and Old Japanese is that internal comparison would not allow to reconstruct voiced obstruents for Proto-Japanese. The fact that in Old Japanese we find the voiced obstruents /b/, /d/, /z/ and /g/, but they do not appear in word-initial position, can be explained through a development of consonant clusters with nasals (\*np, \*nt, \*ns, \*ng) into voiced obstruents. In this case, contractions resulted in four new phonemes.
- **Proto-Japanese to Old-Japanese, vowel inventory:** While Old Japanese has eight vowels, scholars agree that only four vowels (/a/, /o/, /i/, /u/) should be reconstructed for Proto-Japanese. Old Japanese developed one high central vowel and three new mid vowels as a consequence of vowel hiatus. Also in this case, contractions resulted in four new phonemes.
- **Proto-Korean to Middle Korean, consonant inventory** The Middle Korean alphabet encodes many phonemes that cannot be reconstructed for Proto-Korean. While Robbeets acknowledges that the three phonemes associated to voiced fricatives might just be allophones of the obstruents, there is solid evidence that the aspirated obstruents /p<sup>h</sup>/, /t<sup>h</sup>/, /c<sup>h</sup>/ and /k<sup>h</sup>/ all developed from consonant clusters \*pk, \*tk, \*ck and \*kk. This would imply that also for Middle Korean, we have four new phonemes resulting from contraction.
- **Proto-Korean to Middle Korean, vowel inventory** The question of the phonetic nature of Proto-Korean seven-vowel system is debated, but there is an agreement that the seven-vowel system of Middle Korean has a direct reflex in Proto-Korean, either as a direct descent or as a result of a vowel shift.
- **Proto-Tungusic to Manchu, consonant inventory** The phonemic inventories of Tungusic languages are quite similar, and therefore reconstruction is not controversial. The reference here is Benzing (1955). If we follow the development into

Manchu, we would find a /p/ and /f/ contrast that is not reconstructed for Proto-Tungusic: since a change /p/ > /f/ is reported in the history of Manchu, modern /p/ can be explained through dialectal contact with other Tungusic languages in which the sound change did not occur. Another feature of modern Manchu is the contrast between /s/ and /ʂ/, which following Robbeet’s report resulted from a secondary split.

- **Proto-Tungusic to Manchu, vowel inventory** An eight-vowel system is reconstructed for Proto-Tungusic, but modern Manchu is described as a six-vowel system, as a result of the loss of a length contrast in /o/ and /u/.
- **Proto-Mongolic to Mongolian, consonant inventory** The reconstruction follows Poppe (1954), and the changes in the phonemic inventory until contemporary Mongolian can be tracked in detail through Svantesson et al. (2005). Even though the two sources have some minor disagreement, it is clear that a large number of contractions affected the consonant system. Proto-Mongolic lacked /w/ and /ʂ/, which entered the language as allophonic splits from \*p and \*s, and in the second case this might have been facilitated by the presence of loanwords, where the allophonic alternation was not respected. However, the majority of the phonemes are associated to the palatalized versions of the obstruents (eight), the nasals (two), the liquids (two) and the approximant /w/, which contrary to Proto-Mongolic exhibit contrast with their non palatalized version before pharyngeal vowels, and therefore add a total of 13 phonemes to the language.
- **Proto-Mongolic to Mongolian, vowel inventory** The only point of debate is whether the modern Mongolian seven-vowel system can be traced back to Proto-Mongolic, or if the proto-language had a length contrast in /i/ that was lost. For this reason, no sound change is reported in this case.
- **Proto-Turkic to Turkish, consonant inventory** The reconstruction of Proto-Turkic is more unclear than the others. In particular, there is no agreement about whether Turkish /z/ and /ʂ/ result from secondary splits of reconstructed liquid consonants or from some contractions. The only patterns which are uncontroversial are the emergence of /h/ as a result from a secondary split of /p/ in word initial position, while /f/ and /Z/ entered the language through loanwords.
- **Proto-Turkic to Turkish, vowel inventory** The eight-vowel Turkish system can be traced back to Proto-Turkic.

## C Contractions

Modeling contractions artificially is problematic because contractions affects the length of the word, a factor that is usually kept constant in models of sound changes: this true for all models that are cited in this paper (Nowak et al., 1999, Martin, 2007, Graff, 2012, Cui, 2019).

For this reason, I decided to implement ‘pseudo-contractions’: a sound change is applied to two adjacent segments and substitutes them with a new one, but instead of reducing the length of the word, an extra segment is added instead. The algorithm works as follows:

1. Pick one C-place in the syllable (onset, coda).
2. Select one segment in the inventory of the languages among those available for that syllable position (TARGET\_1\_C) and one available in the nucleus (TARGET\_1\_V).
3. Select a second C-segment NOT in the inventory in that syllable position (TARGET\_2\_C).
4. TARGET\_1\_C and TARGET\_1\_V contract to TARGET\_2\_C.
5. **TARGET\_1\_V is substituted with another nucleus present in the language in the current state.**

Point 5 is what makes the contraction a ‘pseudo’ one. Ideally we would like the vowel to be a ‘default’ value (like a /@/) to be closer to the real world, but this would introduce a strong bias. So we let the vowel to be picked at random. This is as close as we can get to the fact that contractions in the real world can both create and destroy minimal pairs: *shy-pie* is a minimal pair in English, while *shy-sky* is not because of a sound change involving contraction.

Consider the following example:

- pen, pin, p@t

In a classical sound change, the /n/ in the first word might be affected by a merger or a split, and therefore change its realization without affecting the preceding vowel. But in a contraction, for example if the /n/ becomes a syllabic consonant, the preceding vowel would be affected as well:

- pen, pin, p@t > p $\eta$ , pin, p@t

This outcome, however, would violate the equal length assumption. As a solution, we can simply pick the /@/ as a vowel replacement.

- pen, pin, p@t > p@ $\eta$ , pin, p@t

This would give us the desired outcome of having contractions both adding and removing minimal pairs. In the example above, one minimal pair is removed, but another is gained.

Running several simulations with and without contractions did not report any change in the main results or the paper, therefore reassuring us that the phenomenon did not introduce any bias in the model.

## D Proto-Romance and Italian

	Proto-Romance	Italian
<b>Morphological Change</b>		
Neuter -> Feminine	<b>RAPU</b> [rapu] ‘turnip’ <b>RAMU</b> [ramu] ‘branch’	rapa [rapa] ramo [ramo]
3rd declension -> 2nd declension	<b>LAVŌRE</b> [lavore] ‘job’ <b>LAVARE</b> [lavare] ‘to wash’	lavoro [lavoro] lavare [lavare]
4th declension -> 2nd declension	<b>CORPUS</b> [k0rpus] ‘body’ <b>CORNU</b> [k0rnu] ‘horn’	<b>corpo</b> [k0rpo] <b>corno</b> [k0rno]
<b>Split</b>		
/k/->/tʃ/	<b>CANTU</b> [kantu] ‘song’ <b>CĚNTU</b> [kĚntu] ‘hundred’	canto [kanto] cento [tʃĚnto]
/E/-/je/ in V̇	<b>VĚTARE</b> [vĚtare] ‘to prohibit’ <b>VŌTARE</b> [votare] ‘to vote’	vietare [vjetare] votare [votare]
/0/->/wo/ in V̇	<b>SANO</b> [sano] ‘healthy’ <b>SŌNO</b> [s0no] ‘sound’	sano [sano] suono [swono]
<b>Contraction</b>		
/lj/ -> /LL/	<b>PALĪA</b> [palja] ‘straw/hay’ <b>PALMA</b> [palma] ‘palm tree’	paglia [paLLa] palma [palma]
/gl/ -> /LL/	<b>TALIA</b> [talja] ‘size’ <b>TEGLA</b> [tĚgla] ‘pan’	taglia [taLLa] teglia [teLLa]
/dj/ and /j/ -> /ddZ/	<b>MAIUS</b> [majus] ‘May’ <b>RADIUM</b> [radjum] ‘ray’	<b>maggio</b> [maddZo] <b>raggio</b> [raddZo]
/aw/ -> /o/	<b>PAUSARE</b> [pawsare] ‘to pause’ <b>PESARE</b> [pesare] ‘to weight’	<b>posare</b> [pozare] <b>pesare</b> [pezare]
<b>Merger</b>		
/l/ -> /e/	<b>BELVA</b> [belva] ‘beast’ <b>SILVA</b> [sllva] ‘forest’	<b>belva</b> [belva] <b>selva</b> [selva]
/U/ -> /o/	<b>COLPU</b> [kolpu] ‘shot/blow’ <b>CULPA</b> [kUlpa] ‘fault’	<b>colpo</b> [kolpo] <b>colpa</b> [kolpa]
unstressed /0/ -> /o/	<b>VŌLARE</b> [volare] ‘to fly’ <b>VŌTARE</b> [votare] ‘to vote’	<b>volare</b> [volare] <b>votare</b> [votare]
unstressed /E/ -> /e/	<b>GĚLARE</b> [dZĚlare] ‘freeze’ <b>BĚLARE</b> [belare] ‘to baa’	<b>gelare</b> [dZelare] <b>belare</b> [belare]
/ct/ -> /tt/	<b>CATTU</b> [kattu] ‘cat’ <b>PACTU</b> [paktu] ‘pact’	<b>gatto</b> [gatto] <b>patto</b> [patto]
/ps/ -> /ss/	<b>CAPSA</b> [kapsa] ‘box’ <b>PASSA</b> [passa] ‘withered/pass-3SG’	<b>cassa</b> [kassa] <b>passa</b> [passa]
<b>h-deletion</b>		
/h/ -> /0/	<b>HORTU</b> [h0rtu] ‘garden’ <b>MORTU</b> [m0rtu] ‘dead’	orto [orto] morto [m0rto]

Table 14: Example of minimal pairs gains and losses between Proto-Romance and Italian.

## E Borrowing and Word formation in English and Italian

- **Borrowed words, English:** ability, able, accept, across, act, action, activity, add, administration, adult, affect, age, agency, agree, air, allow, american, amount, analysis, animal, appear, apply, approach, area, argue, arrive, art, article, artist, attack, attention, author, authority, avoid, bank, bar, base, benefit, big, billion, blue, call, camera, campaign, candidate, car, career, carry, case, catch, cause, cell, center, central, century, certain, chair, challenge, chance, change, character, charge, check, choice, city, claim, class, clear, close, college, color, common, community, company, compare, concern, condition, conference, congress, consider, contain, continue, control, cost, country, couple, course, court, cover, create, crime, culture, current, data, decade, decide, decision, defense, degree, democrat, democratic, describe, design, detail, determine, develop, die, difference, different, direction, director, discover, discuss, disease, doctor, drug, easy, economic, economy, education, effect, effort, election, energy, enjoy, enter, entire, establish, event, evidence, example, executive, expect, experience, expert, explain, face, fact, factor, fail, family, federal, figure, final, fine, finish, firm, focus, force, foreign, form, fund, future, general, get, give, government, group, guy, history, hit, hospital, hour, huge, human, husband, idea, identify, image, imagine, impact, important, improve, include, including, increase, indicate, individual, industry, information, institution, interest, international, interview, involve, issue, join, just, kid, language, large, law, leg, legal, letter, level, list, local, low, main, maintain, major, manage, market, matter, measure, media, medical, member, memory, mention, message, military, million, minute, moment, money, move, movement, music, nation, national, natural, nature, necessary, nice, note, number, occur, office, officer, official, oil, operation, opportunity, order, page, pain, paper, parent, part, particular, party, pass, patient, pay, peace, people, per, perform, period, person, personal, physical, pick, picture, piece, place, plan, point, police, policy, political, poor, popular, population, position, possible, power, practice, prepare, present, president, pressure, price, private, problem, process, produce, product, production, professor, program, project, property, protect, prove, provide, public, push, quality, question, quite, race, raise, range, rate, real, reason, receive, recent, recognize, record, reduce, region, religious, remain, remember, remove, report, represent, require, research, resource, respond, response, rest, result, return, reveal, risk, rock, role, rule, same, save, scene, school, science, season, seat, second, section, security, seem, sense, series, serious, serve, service, several, sex, sign, significant, similar, simple, single, sister, site, situation, size, skill, social, society, sort, sound, source, space, special, specific, stage, standard, state, station, stay, store, story, structure, student, study, stuff, style, subject, success, suggest, support, sure, system, table, take, tax, term, test, their, them, theory, they, though, treat, trial, trip, trouble, try, type, use, value, very, view, visit, voice, wait, want, war, window, wrong
- **Innovations, English:** according, actually, against, already, although, always, another, around, as, available, baby, bad, beautiful, because, behavior, bill, boy, card, certainly, clearly, computer, cultural, cut, despite, development, difficult, during, employee, environment, environmental, especially, evening, everybody, everything, exactly, finally, financial, girl, goal, growth, guess, gun, happy, herself, however, inside, its, itself, job, kill, later, lawyer, likely, management, manager, material, maybe, mister, morning, movie, myself, nearly, network, news, not, ok, or, organization, outside, particularly, past (n.), past (adj.), performance, perhaps, phone, politics, pretty, probably, radio, realize, really, recently, relationship, republican, run, share, she, simply, since, someone, sometimes, sport, statement, suddenly, talk, technology, television, themselves, these, those, throughout, traditional, training, treatment, tv, unit, until, usually, various, whose, yourself
- **Borrowed words, Italian:** affare, angelo, appartamento, appuntamento, aria, arriva, ar-

rivo, assassino, auto, autobus, bagno, balla, ballo, banca, bar, barca, base, basta, bastardo, bianco, bicchiere, biglietto, birra, bisogna, bisogno, blu, bordo, borsa, bottiglia, braccio, bravo, caffè, calma, camera, camion, carne, carriera, carta, cavallo, cavolo, cazzo, cinema, club, colazione, college, colpo, compleanno, computer, controlla, controllo, coraggio, detective, dettagli, diavolo, dipartimento, dna, dollari, drink, droga, energia, episodio, eroe, fantasma, fianco, film, foto, francese, funziona, furgone, gamba, gatto, gay, genio, giro, grandioso, gruppo, guai, guarda, guerra, guida, hotel, idiota, inglese, liceo, lista, macchina, magari, magia, mangia, metri, milioni, nord, okay, ore, orologio, palla, pantaloni, papa, papà, parco, passaggio, pausa, pensiero, periodo, pezzo, pista, pistola, politica, polizia, poliziotto, pratica, prigionia, problema, progetto, programma, puttana, ragazzo, razza, reale, regalo, responsabile, rischio, ristorante, roba, sabato, sala, sbagliato, sbaglio, scarpe, scena, sceriffo, scherza, scherzo, schiena, schifo, scopo, scuola, sembra, sergente, sexy, show, sindaco, sistema, stanco, storia, stronzo, sud, taxi, telefono, teoria, tesoro, test, tipo, treno, troppo, turno, viaggio, villaggio, zio, zona

- **Innovations, Italian:** abbastanza, accanto, accusa, addosso, aereo, affatto, affronta, alcun, almeno, altrimenti, alzati, arrabbiato, arrivederci, assistente, attacco, attraverso, bambino, bellezza, benvenuto, bomba, buco, buonanotte, buonasera, buongiorno, campione, cellulare, ciao, cioè, colonnello, combatte, comincia, compagnia, comunque, conta, contratto, corsa, davanti, davvero, destino, doccia, emergenza, entrambi, eppure, fanculo, ferita, fidanzato, finché, finta, fretta, gara, ghiaccio, giornale, giornata, glielo, guardia, importanza, impronte, incinta, indietro, indirizzo, infatti, infermiera, innamorato, insomma, intorno, invece, laggiù, livello, malato, malattia, manca, mantenere, mentre, meraviglioso, mossa, neanche, nemmeno, nessun, nessuna, nonostante, occhiata, offerta, ognuno, oppure, ormai, ovunque, partita, passato, pazzo, perchè, persino, piccolo, piuttosto, pomeriggio, posto, potere, presa, purtroppo, ricerca, ricordo, riesce, riguarda, riguardo, ringrazio, ritorno, riunione, riuscire, scelta, scommette, scomparsa, scoperto, scopre, semplicemente, sentimenti, serata, sguardo, sicurezza, sistema, situazione, smette, soldato, soltanto, soprattutto, sorpresa, spara, spiace, spiaggia, squadra, stamattina, stampa, stanotte, stasera, stavolta, sveglia, taglio, telefonata, tiro, tranne, troia, tv, video, visita, volo, zitto