On Functional Load and Its Relation to the Actuation Problem

Andrea Ceolin*

1 Introduction

The term 'functional load' can be traced back to Martinet (1955), where he refers to the number of minimal word pairs whose difference relies on a single phonemic contrast as the *rendement fonctionnel* of such a contrast. In Martinet's work, we find the first explicit suggestion that this measure can be a factor in sound change, though similar ideas had been around in the first half of the 20th century (Gilliéron 1918, Mathesius 1929, Jakobson 1931). His hypothesis is that the likelihood of a particular sound change does not entirely depend on the articulatory and acoustic properties of two sounds, but also on their contrastive function: if a language relies on a particular contrast to distinguish words in the lexicon, then we might expect a pressure towards preserving that contrast, while the same pressure would not be present in languages where the contrast is redundant.

King (1967) stressed the importance of investigating the hypothesis, since it is a fundamental step towards a greater understanding of the causality of sound change, or the actuation problem in Weinreich et al. 1968. A theory of sound change that only relies on the articulatory and acoustic properties of sounds fails to account for the fact that sound change is not uniform across languages: some phonemes might frequently merge in the varieties of one language, but remain distinct in other languages. On the contrary, if we integrate the notion of lexical contrast in sound change, we would expect that two phonemes which could merge as a result of misproduction or misperception are kept apart in languages in which they are contrastive.

King's attempts to test functional load found no relationship between phonemic contrast and sound change, and led to some skepticism towards its explanatory role. Moreover, functional load was not part of the variationist agenda proposed by Weinreich et al. (1968). In reference to functional load, Weinreich et al. argue that while Martinet's ideas about sound change being influenced by structural considerations were corroborated by empirical studies (Moulton 1962), it is also easy to identify cases where mergers that lead to homonymy are not blocked (Herzog 1965). These observations led Weinreich et al. to conclude that "the homonymy-prevention theory contributes little to the solution of the 'actuation riddle"' (Weinreich et al. 1968:137).

Recent works which looked at a wider sample of languages found a functional load effect in several different cases (Silverman 2010, Bouchard-Côté et al. 2013, Wedel et al. 2013, Eychenne and Jang 2018, Babinski and Bowern 2018), and in particular Wedel et al. (2013) present the most convincing argument for functional load. In their study, Wedel et al. collect a large sample of ongoing mergers in different contemporary languages, and use a mixed effect logistic model to test the correlation between factors associated to pairs of phonemes (like number of minimal pairs, and relative type and token frequency) and the presence of literature reports of mergers involving those pairs in some language varieties. Their model shows that the number of minimal pairs is the best predictor of a merger. This result is the most convincing evidence provided towards the functional load hypothesis, and it has had great influence in the literature (Sóskuthy 2013, Kiparsky 2016).

These results, however, have some odd implications. If minimal pairs have the effect of blocking sound change, then we would not expect homophony to be widespread in the lexicon of any language. Sampson (2013) shows that several sound changes that occurred between Middle Chinese and Mandarin had the effect of introducing a high amount of homophony in the lexicon, and this empirical observation should be sufficient to falsify a theory which predicts the neutralization of sound change in the presence of lexical contrast. Kaplan (2015) argues instead that if functional load is a statistical tendency rather than a universal law, one might expect to find exceptions in single cases, but still identify the pattern when a sufficient number of cases is considered.

*Thanks to Ryan Budnick, Spencer Caplan, Aletheia Cui, Jordan Kodner, Caitlin Richter, Ollie Sayeed, David Wilson, Robin Clark, Uriel Cohen Priva, Rolf Noyer, Don Ringe, Gareth Roberts, Geoff Sampson, Meredith Tamminga and Charles Yang for many comments and suggestions.

The tension between drawing universal conclusions from typological generalizations and providing a satisfactory account of the empirical data is not specific to the problem of functional load, but is an issue that affects linguistic theory in general. A proposal to reconcile the two pressures when studying sound change has been advanced by Honeybone (2016), and it can be summarized with the following points: i) start from a theory that makes universal predictions, ii) identify all the cases in which the theory is falsified, iii) list all the possible reasons why the predictions might have failed, and determine whether they are likely to have occurred in the case of interest or not. This is the strategy I adopt in this work to re-examine the functional load hypothesis.

2 Methods

In order to determine which are the cases in which functional load makes the right predictions and the cases in which the predictions are not met, I focus on the data from English and Dutch employed in Wedel et al. 2013, since those are among the best cases documented in the literature. Before proceeding, we must discuss some methodological issues that arise when testing functional load.

The first issue concerns the corpus to use to derive minimal pair counts. Databases like CELEX (Baayen et al. 1996) are popular choices to derive language statistics, but their use to derive minimal pairs is questionable. It seems unlikely that children have access to words like *abjuration*, *kilohertz* and *phlegmatic* in the construction of their phonemic inventory, and for this reason CELEX is an unrealistic sample of linguistic input. When counting minimal pairs, one might want to rely on corpora of child directed speech, like CHILDES (MacWhinney 2000), and calculate minimal pairs using a restricted list of frequent words. Therefore, I decided to create a list that includes the most frequent 5000 words in each set of corpora investigated. This is in the order of magnitude of the vocabulary of a 5-year-old child, an age in which the phonological grammar should be fully developed (Ingram 1989, Hoff 2009, McLeod and Crowe 2018).

The second issue concerns the phonemes to choose as *comparanda*. According to King (1967), one needs to make a choice between comparing all possible phonemes that are one or two features apart (the 'weak point' version of functional load) or only those which involve the phoneme that is lost after the merger (the 'least resistance' version). In Wedel et al.'s attempt to compare the mergers they identified with other mergers that could have occurred, the set of tested pairs contains pairs like /p/-/k/. Technically, /p/ and /k/ differ in only one feature (place of articulation), but mergers between them are rare, because they involve two places of articulation which are far apart.[1] It might well be that regardless of functional load considerations, it would be implausible to obtain this merger through misperception or misproduction, and thus it does not qualify as a 'potential merger' for our purposes.

In this situation, a more conservative approach would be: i) when studying consonants, limit the possible mergers to phonemes which are one feature apart, and ii) in the case of place of articulation, limit the possible mergers to those which are close enough that no other phoneme could serve as a 'bridge' between them. It appears reasonable to allow the comparison with phonemes which are two features apart in case of vowels, because we know that there are vowels which are close in the acoustic space, and mergers among them have been attested, despite being two features apart (e.g., [high] and [ATR], for /e/-/ɪ/ and /o/-/ʊ/ in Romance).

A final issue is which hypotheses should be tested. The 'weak point' version, according to King (1967), simply states that mergers are always associated with low contrast. The 'least resistance' version is more conservative, because it assumes that there are independent reasons that put phoneme *x* under pressure, and therefore limits its predictive power to determining the target of the merger (namely, given that *x* is under a pressure to merge, predict the phoneme *y* with which it will merge). The least resistance hypothesis is the easiest one to test, because the domain is restricted and the predictions are clear. The weak point hypothesis suffers from the problem that one would need to define how weak is weak, and the interpretation would be very different depending on whether one wants to test strong versions (e.g., no minimal pairs at all) or conservative versions (e.g., min-

[1] Such a merger is attested in Romanian, where we have Lat. OCTŌ, FRŪCTUS > Rm. *opt*, *frupt*. I thank Don Ringe for pointing this out to me.

imal pairs which are beyond one standard deviation from the overall mean). For this reason, the discussion in this paper will mostly refer to the least resistance version. In this case, 'functional' mergers are those mergers that carry a functional load which is the smallest among those associated with the pairs of phonemes investigated, while 'anti-functional' mergers are those mergers that carry the greatest functional load among those compared.

There are two specific instances where the weak point version makes more sense as the hypothesis to test. First, when testing cases of multiple sound changes attested in a single environment, it makes little sense to ask specifically why a particular sound change happened, but it seems more plausible to ask whether the sound changes attested carry a lower functional load comparable to many other phoneme pairs that could have merged, but did not. Second, when testing cases of several mergers which are not independent, because they involve the loss of the same feature (like for instance, the loss of a voicing contrast), the scenario is similar: we do not want to subject the single changes to tests, but we also want to check whether other contrasts involving the same feature (or the same phonemes) were associated to similar functional load values.

In this case, since we have several mergers occurring in the same environment, determining a single optimal outcome is not the best strategy, because it will rule out the possibility that all the mergers that occurred were optimal in that environment. Therefore, we need to relax the criteria as follows: i)'functional' mergers: mergers that carry a functional load which is the smallest among those associated with the pairs of phonemes investigated OR that is more than one standard deviation below the mean of the unmerged pairs, and ii) 'anti-functional' mergers: mergers that carry a functional load which is the greatest among those associated with the pairs of phonemes investigated OR that is more than one standard deviation above the mean of the unmerged pairs.

3 Results

3.1 English (UK Varieties)

For the study of English varieties in the United Kingdom, we can rely on many electronic resources. The language is well represented in CHILDES, with 18 corpora, and is one of the three languages for which phonological transcriptions are available in CELEX. I collected all the lemmas that appear in the speech of mothers and fathers in CHILDES and that have a frequency higher than one in CELEX, obtaining 7776 lemmas.[2] Then, I used the 5000 most frequent ones, according to CHILDES frequencies, to calculate minimal pairs. Proper nouns were excluded.

The consonant mergers reported for English are all taken from the classic Wells 1982, and they involve the two interdental fricatives /θ/ and /ð/. The merger of these two phonemes with /f/ and /v/, also known as th-fronting, is a characteristic of the Cockney dialect (Wells 1982:328). 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 (Wells 1982:129-430). Finally, a merger with /s/ and /ts/ is reported for Gaelic English (Wells 1982:413).[3]

Sampson (2019) correctly points out that Wells describes th-fronting as a variable rule, not as a merger. Adult speakers of Cockney clearly have the two fricatives in their inventory: they are able to distinguish between them, and they never hyper-correct. Honeybone (2016) makes a strong argument about th-stopping being both irregular and exogenous, namely attested only in bilinguals or language contact areas. The realization of the interdental fricatives as /s/ and /ts/ is also described as a substratum effect from Gaelic (Wells 1982:413). For these reasons, none of these cases qualifies as a spontaneous regular sound change.

For vowel mergers, Wells reports four cases.

The first merger is /aɪ/-/ɔɪ/ (lexical sets PRICE-CHOICE), and is attested in the West Country (Wells 1982:347) and in the North (Wells 1982:425), even though in this second case it might be a substratum effect from Irish, which lacks the second diphthong.

[2] The choice of lemma is motivated by the fact that we want to exclude redundancy due to inflectional morphology.

[3] Wedel et al. report a merger between /ð/ and /z/ instead, but I could not locate it in Wells (1982).

Phonemes	MinPairs
aɪ-ɔɪ	**7**
aɪ-eɪ	63
aɪ-aʊ	19
aɪ-i:	76

Phonemes	MinPairs
ʊə-ɔ:	2
ʊə-ɔɪ	0
ʊə-u:	0
ʊə-əʊ	0

Phonemes	MinPairs
ɪə-ɛə	**22**
ɜ:-ɛə	14
ɪə-i:	7
ɪə-ɜ:	11
ɪə-ʊə	1
ɜ:-i:	21
ɜ:-a:	13
ɜ:-ɔ:	26
ɜ:-u:	9
	μ=12.6 σ=8.48

Table 1: Functional Load for vowel mergers in English (UK varieties).

The second merger is /ʊə/-/ɔ:/ (lexical sets CURE-THOUGHT), and is well attested in different parts of England (Wells 1982:237,287,374). The merger is a recent development, and it occurred after the loss of historical /r/.

The third merger is /ɪə/-/ɛə/ (lexical sets NEAR-SQUARE). This merger is attested in East Anglia (Wells 1982:338) and across the English speaking world, with various outcomes included in the range of the two original diphthongs. In this case, we do not know if its origin is anterior, posterior or related to the loss of historical /r/, but Lass (1992) reports instances of this merger that predate the Great Vowel Shift, and therefore the loss of historical /r/.

The fourth merger is /ɜ:/-/ɛə/ (lexical sets NURSE-SQUARE), which is reported in Liverpool, Merseyside, adjacent parts of Lancashire and some parts of the Manchester area (Wells 1982:361). Since these are also areas where historical /r/ tends to be preserved, it is not clear if this merger is conditioned on the presence of /r/ or not (Sampson 2019).

The first two mergers can be studied in isolation, because they are both unconditioned. This means that we will be testing the least resistance hypothesis. The last two mergers, however, occur in the same environment, and involve the same vowel. Therefore, we have to test the weak point hypothesis.

The results in Table 1 show that minimal pairs are a good predictor for the PRICE-CHOICE merger. This is a sound change that is compatible with the functional load hypothesis. For the CURE-THOUGHT merger the evidence is unclear: the dictionary has two minimal pairs, *you're-your* and *cruel-crawl*, but this second one is a mistake, because the word *cruel* does not belong to the CURE set. The dictionary also has *poor* transcribed as /pɔ:/, with the merger already completed. For these reasons, CELEX is not reliable to address this case.

As for the other two cases, the NEAR-SQUARE and the NURSE-SQUARE mergers are both associated with a large amount of minimal pairs.[4]

Interestingly, the NEAR-SQUARE merger yields the second-highest amount of minimal pairs, and is beyond one standard deviation from the mean. For this reason, it is an anti-functional merger, according to our criteria. Apart from some obvious mistakes like *where'd* and *we're*, which are considered independent lemmas in CELEX even though it is not clear why they should be (but notice that CELEX excludes *were*, since it is an inflected form), it seems that English has a high amount of homophones in the historical rhotic environment to begin with. This sound change represents a problem for functional load: the hypothesis is at odds with the fact that the pre-rhotic environment exhibits a great amount of homonymy, because this should have been prevented to begin with.[5]

[4] ɪə-ɛə, (NEAR-SQUARE): ear-air, ear-heir, weir-where, we're-where, here-hair, hear-hair, here-hare, hear-hare, cheer-chair, beer-bear, beer-bare, peer-pair, sheer-share, fear-fair, peer-pear, dear-dare, deer-dare, weird-where'd, steer-stare, steer-stair, we're-where, really-rarely, deary-dairy ear-air, ear-heir, weir-where, we're-where, here-hair, hear-hair, here-hare, hear-hare, cheer-chair, beer-bear, beer-bare, peer-pair, sheer-share, fear-fair, peer-pear, dear-dare, deer-dare, weird-where'd, steer-stare, steer-stair, we're-where, really-rarely, deary-dairy. ɜ:-ɛə, (NURSE-SQUARE): her-air, her-heir, per-pair, per-pear, fur-fair, fur-fare, purr-pair, purr-pear, furry-fairy, her-hair, her-hare, word-where'd, steer-stair, steer-stare.

[5] One could argue that only minimal pairs for which the part of speech is the same are relevant, since the

Phonemes	MinPairs
w-ʍ	7
v-ʍ	1
f-ʍ	10
h-ʍ	15
b-ʍ	15
p-ʍ	12

Phonemes	MinPairs
t-d	**4**
p-b	2
f-v	3
s-z	2
ʃ-ʒ	0
tʃ-dʒ	0
k-g	0

Table 2: Functional Load for consonant mergers in North American English.

For the NURSE-SQUARE merger, the amount of functional load is within expectations.

3.2 North American English

For the study of North American English, we can rely on the North American section of CHILDES, which contains 63 corpora. While we still rely on CELEX to filter out inflected forms, mistakes, and words which appear only once, this time we use the CMU Dictionary (Weide 1998) to obtain a phonological transcription of the words, since it is an electronic resource developed for the study of North American English. CHILDES contains 9699 lemmas, and I used the 5000 most frequent ones to calculate minimal pairs.

3.2.1 Consonants

There are two consonant mergers reported in the Atlas of North American English (Labov et al. 2008): /w/-/ʍ/ and /t/-/d/ (flapping). In both cases, since we have a single sound change, we can check whether the functional load of these specific mergers is the lowest among the alternatives that the language had, according to the least resistance hypothesis. The counts are reported in Table 2.

The merger between /w/ and /ʍ/ (*witch-which*) is widespread across North America, and there are are seven minimal pairs associated with it.[6] In terms of minimal pairs, the merger between /w/ and /v/ would be preferred, as the merger with /w/ is only the second best option. In this case, we do not have clear evidence for functional load.

The neutralization of the voicing contrast between /t/ and /d/ after a stressed vowel is typical of North American English varieties. Lenition of voiceless consonants when they occur between vowels is a natural sound change, which is acoustically motivated. For this reason, a meaningful question to ask in this case is why the voicing contrast is neutralized for these two phonemes, but not for others. In terms of minimal pairs, as we see from Table 2, neutralizing the contrast is not optimal, because the voicing contrast between /t/ and /d/ is the one that yields the highest amount of minimal pairs (*writer-rider, writing-riding, petal-pedal, metal-medal)*, even though one of them is redundant because CELEX codes *writer-writing* and *rider-riding* as independent lemmas.

One could make the argument that while according to the least resistance hypothesis this merger is not optimal, the number of minimal pairs involved is small, but then we would be left with explaining why a sound change as general as neutralization of voicing does not apply across the board, but only between the two segments for which the contrast stands out from a distributional viewpoint. This merger, then, appears anti-functional.

others can all be disambiguated from the context. However, there are a couple of 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 would likely confirm the results; second, experimental studies have showed that homophones can influence lexical access even if the ambiguity is between two words of different syntactic categories (Boland and Blodgett 2001), and therefore the effect of homophones in perception goes beyond sytactic categories.

[6] w-ʍ (WITCH-WHICH): why-‘y’, where-wear, whether-weather, which-witch, whine-wine, wise-whys, win-when (only in varieties with the *pin/pen* merger).

Phonemes	MinPairs
ɑ-ɔ	**5**
ʊ-ɔ	6
ʌ-ɔ	11

Phonemes	MinPairs
ɔɪ-ɜr	**2**
aɪ-ɜr	23
eɪ-ɜr	25
aʊ-ɜr	9
oʊ-ɜr	21
i-ɜr	61
u-ɜr	18

Table 3: Functional Load for vowel mergers in North American English.

Phonemes	MinPairs
ɪ-ɛ	8
æ-ɛ	9
ʌ-ɛ	7

Table 4: Functional Load for some possible vowel mergers before nasals in North American English.

3.2.2 Vowels

There are two unconditioned mergers that are mentioned in Wedel et al. (2013): the first is the /ɑ/-/ɔ/ merger (LOT-THOUGHT, or *cot-caught*), which has been well documented in the literature (Labov et al. 2008); the second merger is /ɔɪ/-/ɜr/ (CHOICE-NURSE). This merger is discussed in Labov 1966, and was found in some speakers of New York City English that were born at the beginning of the 20th century. This merger causes homonymy among minimal pairs like *coil-curl*. Since these two mergers involve two different classes of vowels (short vowels for the first merger, and long vowels for the second merger), we can investigate them independently according to the least resistance hypothesis.[7]

Table 3 contains the results. Most of the minimal pairs for LOT-THOUGHT are resulting from noise. Apart from *stock-stalk*, we have the interjections *ha-ho*, *pa-po*, *don-dawn* (where the first word appears in CELEX and CMU, but is a misspelling for *don't* in CHILDES), and *la-law*, where the first word is an onomatopoeia. The contrast is the weakest one according to minimal pairs.

The second table in Table 3 shows counts for /ɜr/ and other diphthongs and long vowels, and it shows that this is a case where functional load seems to be fully compatible with the merger, since it only distinguishes two minimal pairs (*oil-earl, soy-sir*). Both mergers can be considered functional.

3.2.3 Vowels Before Nasals

Another well studied conditioned merger is the merger between /ɪ/ and /ɛ/ before nasal consonants, the *pin-pen* merger. Wedel et al. (2013) describe the merger as occurring before /n/, but actually the merger occurs before all nasals (/m/, /n/ and /ŋ/).

Table 4 contains the calculations for some other potential vowel mergers involving /ɛ/, and we can see that in this case functional load seems to play little role. In terms of minimal pairs, we find six pairs associated with this merger.[8] This is more or less the same amount of pairs that we find in other contrasts before nasals. In this case, it is unclear how functional load relates to this merger.

[7]Using the CMU dictionary for estimating functional load can be done with some caveats. In North American English, rhotic environments can preserve unconditioned mergers or trigger conditioned mergers, and this is a problem for functional load calculations. According to the CMU dictionary, *star-store* is a minimal pair which is associated to the LOT-THOUGHT merger, but this is a mistake, because the rhotic environment preserves the historical short-/o/, which instead merges with /ɔ/ in all the other environments in the majority of the North American English varieties. The CMU dictionary does not encode /o/ among its phonemes. For this reason, a proper study of the LOT-THOUGHT merger requires coding rhotic environments as independent symbols.

[8]ɪ-ɛ, (PIN-PEN): win-when, tin-ten, pin-pen, din-den, mini-many, in-'n', since-sense, gym-gem.

Phonemes	MinPairs
ɪ-i	**13**
ʊ-u	2
ʊ-oʊ	3
ʌ-ɔ	**0**
ʌ-ʊ	**0**
ɛ-eɪ	**12**
ʌ-oʊ	**0**
ɪ-ɛ	10
ɪ-ʌ	1
ʊ-ɔ	3
ʊ-aʊ	3
oʊ-aʊ	3
ʌ-ɛ	1
ʌ-ɑ	2
ʌ-eɪ	1
ʌ-aʊ	0
ɪ-eɪ	16
i-eɪ	17
ɑ-ɔ	0
æ-ɛ	1
ɑ-æ	2
ɔ-oʊ	10
ɔ-aʊ	0
u-oʊ	9
u-aʊ	0
	μ=4.39 σ=5.52

Table 5: Functional Load for some possible vowel mergers before /l/ in North American English.

3.2.4 Vowels Before /l/

In North American English we find many mergers before /l/. Wedel et al. (2013) cite, from Labov et al. 2008, the following mergers: /ɪ/-/i/ (*hill-heel*), /ʊ/-/u/ (*pool-pull*), /ʊ/-/oʊ/ (*bull-bowl*), /ʌ/-/ɔ/ (*hull-hall*) and /ʌ/-/ʊ/ (*hull-bull*). Additionally, Labov et al. (2008) report areas where the mergers involve /ɛ/-/eɪ/ (*fell-fail*) and /ʌ/-/oʊ/ (*hull-hole*), for a total of seven vowel mergers. This is another case that qualifies for the weak point hypothesis test.

Table 5 lists all phonemic pairs which could have potentially merged. From the viewpoint of minimal pairs, two of these mergers (/ɪ/-/i/ and /ɛ/-/eɪ/) appear suboptimal, because they lead to a great amount of homonymy.[9] These two mergers are beyond one standard deviation from the mean, and therefore they are both anti-functional. Some minimal pairs are also found for /ʊ/-/u/ (*pull-pool, full-fool*) and /ʊ/-/oʊ/ (*pull-pole, full-foal- bull-bowl*).

In terms of minimal pairs, we have two mergers which are clearly anti-functional, two mergers which are neutral, and three mergers which can be considered functional (namely, they yield no minimal pairs at all).

3.2.5 Vowels Before Intervocalic /r/

Many vowel mergers in North American English are limited to the pre-rhotic environment. Wedel et al. (2013) select only one of them, namely the merger between /ɑ/ and /ɔ/ (START-NORTH, or *far-for*). Some varieties of English merge instead /ɔ/ with /o/ in pre-rhotic environment (NORTH-FORCE, or *for-four*). The CMU phonological dictionary has a unique symbol in this case, because it neutralizes the distinction between /ɔ/ and /o/. Since these two mergers are mutually exclusive (namely, there aren't varieties where START-NORTH-FORCE share the same unique value), this makes it impossible to count minimal pairs for the START-NORTH merger, because the NORTH class and the FORCE class are overlapping.

However, there are several other mergers that are reported in Labov et al. 2008 before an intervocalic /r/: we have the merger between /ʌ/ and /ɜr/ (the *hurry-furry* merger, the merger among /eɪ/, /æ/ and /ɛ/ (the *Mary-marry-merry* merger), the merger between /ɛ/ and /ʌ/ (the *merry-Murray* merger) and the merger between /ɪ/ and /i/ (the *mirror-nearer* merger).

In this case we have a sparsity problem: phonemic contrasts in this environment are rare, and therefore almost all vowel contrasts yield zero minimal pairs in this context. This case is not ideal to test functional load.

[9] ɪ-i, (HILL-HEEL): will-wheel, fill-feel, hill-heel, sill-seal, pill-peal, mill-meal, filling-feeling, will-we'll, hill-heal, still-steal, still-steel, willing-wheeling. ɛ-eɪ, (FELL-FAIL): well-whale, tell-tale, tell-tail, bell-bale, sell-sale, sell-sail, belle-bale, hell-hail, cell-sale, cell-sail, 'l'-ale, cellar-sailor.

Phonemes	MinPairs
f-v	**1**
s-z	**3**
χ-ɣ	**0**
t-d	23
p-b	33
v-b	36
v-z	16
v-m	26
v-w	18
z-d	24
z-l	22
z-n	8
z-ɣ	4
h-ɣ	0
j-ɣ	4
ŋ-ɣ	0
	μ=16.5 σ=12.3

Table 6: Functional Load for voicing contrasts in Dutch.

3.3 Dutch

For Dutch, we have 14 corpora in CHILDES. Putting together the speech of mothers, fathers, siblings and investigators and filtering the words we have about 7700 lemmas that in CELEX have frequency of at least two, and we select the top 5000. Proper nouns are excluded.

Kissine et al. (2003) report that in the northern part of the Netherlands, voiced fricatives are disappearing, and they merge with their voiceless counterpart. This is another case where we need to test the weak point hypothesis, by comparing the mergers with other possible mergers involving the same feature or the same segments. In Table 6, we see that the contrast between /f/ and /v/ has only one minimal pair (*fee-'v'*) while the contrast between /s/ and /z/ has three (*zee-'c', set-'z', set-zet*). The merger between /χ/ and /ɣ/ has zero. All of them are below the mean of at least one standard deviation, and therefore can be considered functional.

4 Summary and Conclusion

The results of this investigation can be summarized in Table 7. First, most of the unconditioned mergers (6/7) are compatible with functional load. This is not true for the conditioned mergers (3/11). Second, all the mergers with an 'anti-functional' result are conditioned on a phonetic environment. These observations point towards the hypothesis that functional load is a factor that is only identifiable for unconditioned mergers, namely mergers that completely remove a contrast from the grammar. When the merger is limited to a specific environment, functional load considerations seem to be less relevant to decide whether to merge two sounds.[10,11]

The question at this point is: what is the source of the correlation between lexical contrast and sound change, and why would this factor influence one kind of sound change, but not the other? One potential place to look at is experimental findings in language acquisition. Experimental evidence shows that children lose sensitivity to phonetic differences very early in the course of the developmental period. While at birth they are able to distinguish among different sounds, within their first year of life their perceptual system adapts to the sounds which characterize the language to which they are exposed, to the point that they lose sensitivity to fine phonetic distinctions that are not associated with the phonemic contrasts of the language they are learning (Werker and Tees 1984, Kuhl et al. 2006). This implies that the development of the phonological inventory must begin when children are very young.

When children have settled on an inventory of phonemes, they can then pay attention to the alternations among them, for instance realizing that the sound /t/ has a different phonetic realization depending on whether it appears at the beginning of the word or intervocalically. The developmental window for learning these allophonic rules is estimated to be much longer, at least until children are 5 years old (Klein and Altman 2002, Richter 2018).

[10]The merger between /ɪə/ and /ɛə/ (the NEAR-SQUARE merger) represents a delicate data point, because its nature as a conditioned merger is unclear, and in other varieties of English (e.g. New Zealand English) it appears to be unconditioned.

[11]This is compatible with the finding that word-final lenition does not seem to be motivated in terms of functional load (Cohen Priva 2017).

Phonemes	Language	Type	Functional	Neutral	Anti-Functional
aɪ-ɔɪ	English (UK)	Unconditioned	X		
ɪə-ɛə	English (UK)	before /r/ (?)			X
ɜː-ɛə	English (UK)	before /r/ (?)		X	
t-d	North American Eng.	V́_V			X
w-ʍ	North American Eng.	Unconditioned		X	
ɑ-ɔ	North American Eng.	Unconditioned	X		
ɔɪ-ɛr	North American Eng.	Unconditioned	X		
ɪ-ɛ	North American Eng.	before [NASAL]		X	
ɪ-i	North American Eng.	before /l/			X
ʊ-u	North American Eng.	before /l/		X	
ʊ-oʊ	North American Eng.	before /l/		X	
ʌ-ɔ	North American Eng.	before /l/	X		
ʌ-oʊ	North American Eng.	before /l/	X		
ɛ-eɪ	North American Eng.	before /l/			X
ʌ-oʊ	North American Eng.	before /l/	X		
f-v	Dutch	Unconditioned	X		
s-z	Dutch	Unconditioned	X		
χ-ɣ	Dutch	Unconditioned	X		
TOTAL			9	6	4

Table 7: Functional Load measured with minimal pairs.

In this scenario, it is not implausible that unconditioned mergers occur in the first stage, while conditioned mergers occur in the second stage. Unconditioned mergers occur when a child fails to acquire a phonological boundary between two categories, which are then acquired as a unique entity. Conditioned mergers, on the other hand, do not affect the phonological inventory: the number of categories is the same before and after the merger. The only difference between the two stages is the presence of an allophonic rule, or the fact that an entire class of words (like the class of *pen* words, which contain the sequence /ɛ/ + [nasal]) shifts from one category to another. This would explain why functional load is strongly associated with unconditioned mergers, but does not explain conditioned mergers: by the time children reach the stage of allophonic rule-learning, their categories are already fixed. In this second stage, the learning strategy might be different, and other properties of the lexicon, like morpho-phonological alternations or finer phonetic distinctions, can play a greater role (cf. Richter 2018).

In conclusion, this work corroborates previous attempts to show a correlation between mergers and lexical contrast (Silverman 2010, Wedel et al. 2013, Eychenne and Jang 2018, Babinski and Bowern 2018), but restricts the functional load hypothesis to unconditioned mergers: if homonymy results from different processes, like the loss or creation of an allophonic rule, then homonymy seems to be tolerated. This finding has some implications for the actuation problem, because it can explain why certain sound changes occur in some languages but not in others. Unfortunately, it does not explain why the changes happen at the exact place and time in which they happen, but points to lexical contrast as a place to look at to try to address the 'actuation riddle'.

References

Baayen, R Harald, Richard Piepenbrock, and Leon Gulikers. 1996. *CELEX2*. Linguistic Data Consortium, Philadelphia.

Babinski, Sarah, and Claire Bowern. 2018. Mergers in Bardi: Contextual probability and predictors of sound change. *Linguistics Vanguard* 4, s2.

Boland, Julie E, and Allison Blodgett. 2001. Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements. *Journal of Memory and Language* 45:391–411.

Bouchard-Côté, Alexandre, David Hall, Thomas L Griffiths, and Dan Klein. 2013. Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences* 110:4224–4229.

Cohen Priva, Uriel. 2017. Informativity and the actuation of lenition. *Language* 93:569–597.

Eychenne, Julien, and Tae-Yeoub Jang. 2018. On the role of contrastivity in the development of the /e/-/ɛ/ merger in Korean. *Journal of East Asian Linguistics* 27:111–139.

Gilliéron, Jules. 1918. *Généalogie des mots qui désignent l'abeille d'après l'Atlas linguistique de la France*, volume 225. Champion.

Herzog, Marvin. 1965. *The Yiddish language in northern Poland*. Indiana University.

Hoff, Erika. 2009. Language development at an early age: Learning mechanisms and outcomes from birth to five years. In *Encyclopedia on early childhood development*, ed. Susan Rvachew.

Honeybone, Patrick. 2016. Are there impossible changes? θ>f but f≯θ. *Papers in Historical Phonology* 1:316–358.

Ingram, David. 1989. *First language acquisition: Method, description and explanation*. Cambridge University Press.

Jakobson, Roman. 1931. Prinzipien der historischen phonologie. *Travaux du Cercle Linguistique de Prague* 4:247–267.

Kaplan, Abby. 2015. Discussion 4: A highly improbable data point. *Journal of Chinese Linguistics* 43:710–713.

King, Robert D. 1967. Functional load and sound change. *Language* 43,4:831–852.

Kiparsky, Paul. 2016. Labov, sound change, and phonological theory. *Journal of Sociolinguistics* 20:464–488.

Kissine, Mikhail, Hans Van de Velde, and Roeland Van Hout. 2003. An acoustic study of standard dutch /v/ , /f/, /z/ and /s/. *Linguistics in the Netherlands* 20:93–104.

Klein, Harriet B, and Elaine K Altman. 2002. The acquisition of medial /t,d/ allophones in bisyllabic contexts. *Clinical linguistics & phonetics* 16:215–232.

Kuhl, Patricia K, Erica Stevens, Akiko Hayashi, Toshisada Deguchi, Shigeru Kiritani, and Paul Iverson. 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science* 9:F13–F21.

Labov, William. 1966. *The social stratification of English in New York City*. ERIC.

Labov, William, Sharon Ash, and Charles Boberg. 2008. *The Atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.

Lass, Roger. 1992. Phonology and morphology. *The Cambridge history of the English language* 2:1066–1476.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk. transcription format and programs*. Mahwah, NJ: Lawrence Erlbaum Associates.

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.

McLeod, Sharynne, and Kathryn Crowe. 2018. Children's consonant acquisition in 27 languages: A cross-linguistic review. *American journal of speech-language pathology* 27:1546–1571.

Moulton, William G. 1962. Dialect geography and the concept of phonological space. *Word* 18:23–32.

Richter, Caitlin. 2018. Learning allophones: What input is necessary? In *Proceedings of the 42nd annual Boston University Conference on Language Development*. Cascadilla Press Somerville, MA.

Sampson, Geoffrey. 2013. A counterexample to homophony avoidance. *Diachronica* 30:579–591.

Sampson, Geoffrey. 2019. An unaddressed phonological contradiction. *Journal of Chinese Linguistics* 6:221–237.

Silverman, Daniel. 2010. Neutralization and anti-homophony in Korean. *Journal of linguistics* 46:453–482.

Sóskuthy, Marton. 2013. Phonetic biases and systemic effects in the actuation of sound change. Doctoral dissertation, University of Edinburgh.

Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128:179–186.

Weide, Robert. 1998. *The CMU pronunciation dictionary, release 0.6*. Carnegie Mellon University Pittsburgh, PA.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. *Empirical foundations for a theory of language change*. University of Texas Press.

Wells, John C. 1982. *Accent in English: An Introduction*. Cambridge University Press.

Werker, Janet F, and Richard C Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development* 7:49–63.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104
ceolin@sas.upenn.edu