Functional load: comparisons over measures and texts

John D. PHILLIPS

Abstract

本稿では、言語学の機能負担量という概念を取り調べる。ウェールズ語の文書を例示しながら言語の各音素の重要性を計る三つの方法を比べる、即ち頻度・最小対語による機能負担量・情報理論による機能負担量。文書を三つ、時代を二つ比べて、七百年間の間の変化を明らかにする。機能負担量によく使われる計り方である最小対語による方法と情報理論による方法を比べると相関が殆ど無いので違う特性を計っていると思わせられる。

1. Introduction

Functional load is a measure of the work that some feature of a language does, usually a phonological feature. It has its origins in the idea of the phonology of a language serving to differentiate the words of the language and avoid homonymy, as a measure of the degree of utilisation of a phonemic opposition or phonological unit. In English for instance *c* distinguishes *cat* from *bat, chat, fat, hat, mat, gnat, pat, rat, sat, tat, vat*, the pronunciations of which are similar except for the initial sound (phoneme). The functional load of the distinction between *t* and *d* is the number of word-pairs (minimal pairs) distinguished: *tin/din, ton/done, writer/rider, bit/bid*, etc. The concept can apply not just to phonemes but to phonological distinctions in general. The functional load of voice could be measured as the number of minimal pairs distinguished by voice: *tin/din, pin/bin, tie/dye, thigh/thy, pie/bye, fie/vye,* etc.

King (1967b) traces the idea back to the early twentieth century, noting similar thinking in the work of Gilléron (1918, p. 14), later developed by Mathesius (1929) and other linguists of the Prague Circle, and particularly by Martinet (1933; 1955).

Though the basic idea is clear enough, functional load is hard to define precisely in terms of minimal pairs (see below). Hockett in 1955 suggested conceiving of functional load in a different way, one based on the then-new

science of information theory. He defined the functional load of a contrast as the change in the entropy of the language which collapsing the contrast would cause. Though he published a formula for calculating functional load, Hockett never published any actual measurements.

A similar idea was developed and used independently in the field of automatic speech recognition: Carter (1987) used an information-theoretic measure similar to the one Hockett had suggested, to calculate the usefulness of recognising classes of phonemes in the initial stage of the automatic analysis of speech into phonemes. More recently, Surendran & Niyogi (2006) applied Carter's metric to functional load and used it to demonstrate how the functional load of distinctions and distinctive features can be calculated. The years since then have seen some other work in the same vein (Oh et al., 2015; Surendran & Levow , 2004; cf. also Coupé et al., 2009)

Apart from its intrinsic interest functional load has been applied to practical use in historical linguistics and in language teaching. In language learning an appeal is often made to functional load to say that for instance as the functional load of the distinction between *t* and *th* in English is small, it is not so important if a foreign learner pronounces *th* as *t*. It seems to have been Brown in 1988 who first advocated the appeal to functional load in teaching pronunciation, and the idea has taken root. Munro and Derwing (2006) presented experimental evidence that functional load, using the minimal-pair-based rankings of Catford (1987) and Brown (1991), is to some extent correlated with comprehensibility, and hence can be effectively employed in guiding some aspects of pronunciation instruction.

Application to historical linguistics has been less successful. Martinet (1933, 1955) suggested that functional load influenced language change in that contrasts with low functional load were liable to be lost over time. Testing this hypothesis, King (1967b) made some empirical comparisons of functional load, which he measured with a metric (King, 1967a) derived from textual frequency and possibility of contrast in minimal pairs, in relation to historical phoneme mergers in Germanic languages, concluding that 'functional load and relative frequency seem to be largely irrelevant in sound change' (p. 848). King's results were criticised at the time (Hockett, 1967), but more recent empirical work using information-theoretic measures has supported his conclusion: Surendran & Niyogi (2006) showed that before initial *l* and *n* merged recently in Cantonese, the contrast between *l* and *n* carried one of the highest functional loads.

Though recent work on functional load has mostly looked at phonological contrasts, the early work also applied to phonological entities: *Grad der Ausnützung von phonologischen Einheiten*, 'the degree of use of phonological units', as Mathesius put it (King, 1967b, p. 832). Oh et al., in their recent (2015) paper, propose an information-theoretic way of measuring the functional load of phonological units and use it to compute the correlation between functional load and frequency for the phonemes of Cantonese, English, Japanese, Korean and Mandarin. They also compare the functional loads of individual vowels between these five languages. They find the correlation between functional load and frequency to be rather strong in most cases, but it is not straightforward and may be limited: it is very low for Mandarin consonants for instance.

My own earlier work on change in Welsh (Phillips, 2009) has shown that the textual frequency of some Welsh phonemes has fallen over the last few centuries. There is an impression that these phonemes are less important to modern Welsh than they were to mediæval Welsh. In this paper I investigate this question by comparing three different ways of measuring how much use a language makes of a phoneme: textual frequency, functional load as measured by minimal pairs, and functional load measured in bits of information. I compare texts from different periods to look at how phoneme use has changed over time.

2. Texts

The analyses here are based on three texts in the Welsh language, chosen for their different style and period. Firstly the *Cronfa Electroneg o Gymraeg* (Ellis, 2001) is a sampled corpus of about a million words of late twentieth-century written Welsh, comparable to the Brown corpus of American English or the LOB corpus of British English. The second text is a recent (1988) translation of the Bible, slightly less than a million words in a fairly formal narrative style. The third text comprises a collection of mediæval stories, the four tales in *Pedair Cainc y Mabinogi* (Williams, 1930), and a fifth *Peredur* (P. W. Thomas, 2000), altogether about fifty thousand words. These tales are transcribed from fourteenth century manuscripts, though the Mabinogi tales are thought to have been copied from manuscripts a couple of centuries older. The language of the tales is Middle Welsh, an older form of the language of the first two texts. The narrative style is roughly comparable to that of the Bible.

All three texts have been converted automatically to a phonemic representation in the International Phonetic Alphabet using the software described by Phillips (2013). Because of the automatic nature of this conversion, there will of course be a proportion of mistakes. Specifically with the mediæval texts, the orthography used for Middle Welsh is highly ambiguous. The mediæval pronounciation of the dictionary form of a word can be fairly reliably deduced, as can that of word-endings, but morphological changes to the initial consonants of words are often not written and cannot be reliably restored. Word-initial /m̥ n̥ ŋ̊ r̥ ð/, in particular, are likely to appear much less common than they actually were in Middle Welsh. A check of 525 words of the Middle Welsh text used here reveals two errors of this type, and ten other mistranscribed phonemes out of a total of 1592, about ¾%.

A more important problem with a corpus of this type is that it transcribes the citation form of each word, without any of the elisions and assimilations which would happen in natural speech. The scarcity of phonetically transcribed speech corpora makes this problem hard to quantify, but Surendran and Niyogi (2006) find an 82% correlation between functional loads of consonant oppositions in a corpus of carefully transcribed American speech, and the functional loads calculated from word frequencies in written British English texts combined with British English citation form pronunciations.

The Welsh texts have had foreign words removed as far as possible. Those Biblical personal and place names not adapted to Welsh phonology were removed from the Welsh Bible text, just under 4% of word tokens. The CEG is tagged with part-of-speech and other information, and words tagged as *foreign* were removed as well as other words whose spelling contained any of the non-Welsh letters *qkvxz*: these are words and phrases in a language other than Welsh (mostly English) which the CEG's annotation has failed to tag. Again, just under 4% of word tokens were removed (361 word types out of 40,621).

In addition to the three texts, a Welsh Dictionary (Collins-Spurrell, 1996) is used for one comparison. The dictionary contains 13,491 entries (lexemes), midway between the Bible (approaching 8,000 lexemes) and the CEG (less than 20,000 lexemes).

3. Measures

I compare here three methods which have been used to measure the importance of a phoneme in a language: simple textual frequency, and two different ways of measuring functional load.

3.1. Frequency   The frequency of a phoneme is the simplest measure of how much use a language makes of it. It is measured over a corpus, the percentage of the total number of phoneme tokens comprising the corpus.

3.2. Minimal pairs   The original and still most widely used measure of functional load is the number of minimal pairs distinguished, a minimal pair being defined as two words differing in exactly one phoneme. Though the basic idea is clear enough, most scholars who have looked as the subject (e.g. Brown, 1988, 1991) have pointed out how hard it is to define precisely. The basic question is whether all minimal pairs are equal, or whether they should be weighted in some way. There are at least three possible types of weighting.

   1. The most obvious point is that common words seem more important than rare words. This can be measured objectively as textual frequency or subjectively by the psycholinguistic metric of word familiarity. The two measures correlate, but not particularly strongly (Tanaka-Ishii & Terada, 2011). Kitahara & Amano (2000), studying the functional load of accent in Japanese, weight minimal pairs by the inverse of the difference between the familiarity scores of the words of the pair. The logic "is that when the familiarity scores between the two items differ, the role of accentual contrast within that pair will be reduced. In other words, when the two items differ enough in familiarity, accent does not have to play a major role in identifying the target word" (p. 292). Others suggest that only when both members of a minimal pair score highly is the pair important. Frequency is generally preferred to familiarity, but there seems to be no evidence or argument for preferring one to the other. On the other hand Munro and Derwing (2006) found no conflict at all between the minimal-pair-based rankings of Brown (1991), who used frequency, and Catford (1987), who took no account of frequency or familiarity. Further, Kawashima (2012) conducted an experimental investigation with Japanese students learning English of the correlation between ability to distinguish minimal-pairs, the frequency of those minimal pairs, and general listening proficiency. He found no effect of frequency with vowels and only a small effect with

consonants. It appears then that word-frequency may not be so important.

2. Another contentious area is context: minimal pairs the members of which could be substituted sensibly for one another in sentences are perhaps more important. For instance in the last sentence, substituting *should* /ʃʊd/ or *would* /wʊd/ for *could* /kʊd/ maintains grammaticality but changes the meaning, so the distinction between *c/sh/w* is doing important work here. Conversely, replacing *could* with *mud* /mʊd/ or *good* /gʊd/ results in nonsense so the distinction between *c/g/m* is less important here: mishear *could* as *good* and the context will put you right, or so the argument goes. Some scholars partly address this point by giving greater weight to minimal pairs where both members have the same part-of-speech. A better way to address the problem would be by weighting each minimal pair with a measure of the interchangeability of the members, perhaps derived from a stochastic language model.

3. If 'minimal pair' is defined simply in terms of phoneme substitution, it follows that all phonemes are theoretically equal, so that for instance *flirt* and *alert* are a minimal pair, with *f* and *a* contrasting. However, some minimal pairs seem more minimal than others: *could* seems phonetically closer to *good* than to *mud*, though all differ by one phoneme. *Flirt* and *alert* differ prosodically in number of syllables, as well as in the initial phoneme. Some scholars partly address this point by considering only contrasts within groups of 'similar' phonemes, oftenest just the two groups consonants and vowels. From the standpoint both of language learning and of language change though, some consonants are more similar to some vowels than to other consonants. Some English pronunciations of /l/ are often misheard as /u/ and in some accents of English (e.g. Cockney, Scots) post-vocalic /l/ is vocalised. Post-vocalic /r/ of earlier English has been vocalised in most modern varieties, either lengthening the previous vowel or forming a diphthong with it, sometimes making the vowel retroflex (in 'rhotic' varieties). This can change the number of syllables, as in *fire* /fajr, fajə/. In the Celtic languages too, earlier monosyllabic /tarß/ 'bull' becomes bisyllabic /taru/ in modern Welsh and Gaelic, and earlier Welsh /eirɣ/ 'snow' becomes modern /eira/. In the opposite direction, late Brythonic vowel /i/ can give rise to Welsh consonant /ð/ (Morris Jones, 1913, p. 99). Examples are /r̥ɨð/ 'free', which is cognate with its English translation, and the ordinal suffix /−ɨð/ seen in /peduerɨð/ 'fourth', cognate with Gaulish /−i−/ in *petuarios*.

The specific distinction between vowels and consonants is then of doubtful relevance; what is needed is a general account of phonetic similarity. However, as Kessler (2005) notes in a review of measures of phonetic similarity, 'phonetic distance can mean quite a few different things'. He discusses (p. 247) measurement of phonetic similarity in terms of acoustic properties, articulatory gestures, perceptual features, and likelihood of diachronic change. Each of these can be measured in a multitude of ways differing in the size of the phonetic segment being compared (phoneme, phone, gesture), whether phonetic context is taken into account, the number of different features, whether features are binary-valued or multiple-valued, how the different features are combined to give a final measure, etc. Kessler concludes (p. 254) that 'it would be premature to call phonetic comparison a mature discipline, but at least it is fast becoming a discipline.' More recently, Mielke (2012) proposed a metric for phonetic similarity based on experimentally measured acoustic and articulatory data. Different types of measurements and analyses produced very different results, implying that a general account of phonetic similarity is still elusive.

Minimal pairs are used for other purposes besides the measurement of functional load. In psycholinguists, density of minimal pairs ('phonological neighbourhood density') has been shown to influence speech production and perception, and here minimal pairs are invariably defined simply as differing in one phoneme (e.g. Gahl, Yao & Johnson, 2012, and references therein). Likewise in a rather different field, the automatic measurement of phonetic similarity between lists of whole words for purposes of automatic language classification, the ASJP project have found (Brown et al., 2008; Holman et al., 2008) that weighting correspondences between phonemes by a measure of either similarity of phonemes or 'common sound changes' did not improve overall results compared to a simple same-or-different metric: 'better results (judged with respect to how ASJP automated language classification compares to expert classification) were achieved by requiring segments to be identical' (Holman et al., p. 333). It seems that more work needs to be done on measures of phonetic similarity before they can be useful in practice.

Apart from considerations of weighting, a separate point is that though a minimal pair is normally defined by a single phoneme substitution, in actual language use it is often the presence or absence of a phoneme which causes misunderstanding. The pairs *thirty/thirteen* and *can/can't* (in accents which

have the same vowel in both) are often confused for instance. It could be argued then that to the list of contrasts *bin, chin, din, fin, gin, kin, Lynne, pin, quin, sin, shin, tin, win* should be added *in*, with null as an additional contrast, as in some psycholinguistic work. Gahl, Yao & Johnson (2012), for instance, define two words as a minimal pair 'if they differ by deletion, insertion, or substitution of one segment'. In language change too, functional load would be a more useful concept if it could handle elision and epenthesis (deletion and addition of phonemes) as well as mergers.
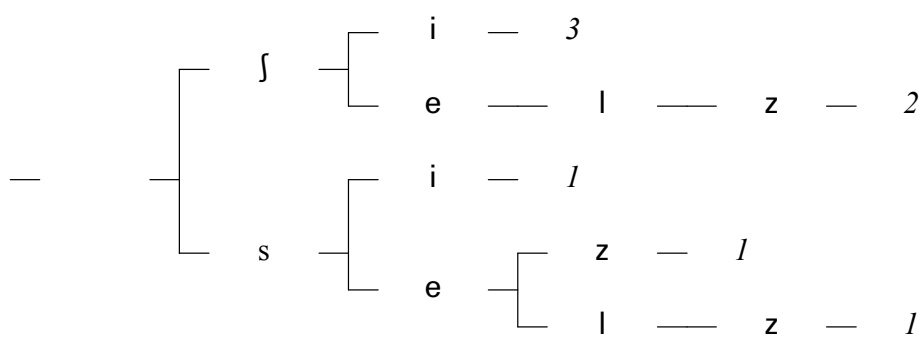
However, given the multitude of possibilities and the lack of any substantial argument or evidence preferring one definition of minimal pair over another, a minimal pair is defined for this paper in the simplest way, by a single phoneme substitution; word-frequency, phonetic similarity, prosody and context are ignored. Minimal pairs are extracted from a list of the word-types contained in the corpus being investigated.

3.3. Surprisal

Speech is a stream of consecutive sounds. As a listener hears each new sound, how much information does it contribute to the listener's understanding? The average information contributed by a particular phoneme is a measure of that phoneme's importance in the language, of its functional load.
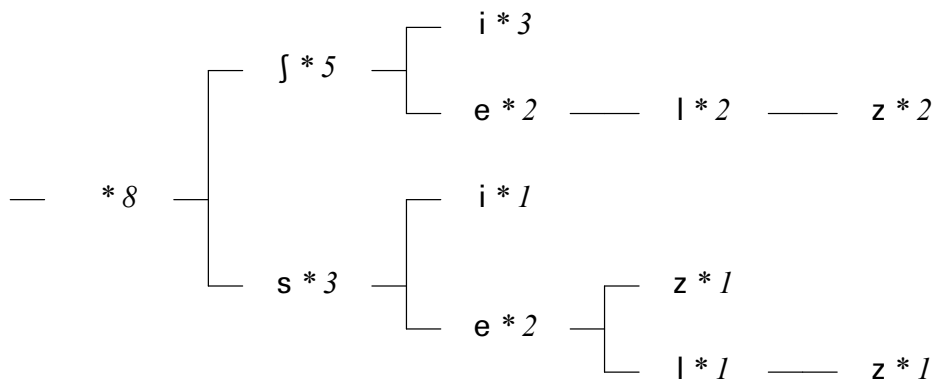
Surprisal is a measure of the information content associated with an event in a probability space. It is measured as the logarithm of the inverse of the event's probability, so if the event is hearing a particular phoneme, the surprisal is the logarithm of the inverse of the phoneme's probability in that context. If the logarithm is to base 2, the surprisal will be measured in binary bits of information. The average surprisal value of a phoneme will be a measure of that phoneme's contribution to the language.

Obviously we have no means of measuring the average surprisal of a phoneme in a language as a whole; we can only investigate concrete texts which we take to represent the whole. The surprisal value of a phoneme is then the sum of the information conveyed by the phoneme in the text as a proportion of the total information conveyed by all phonemes. A simple method of measuring surprisal in a text is as follows.
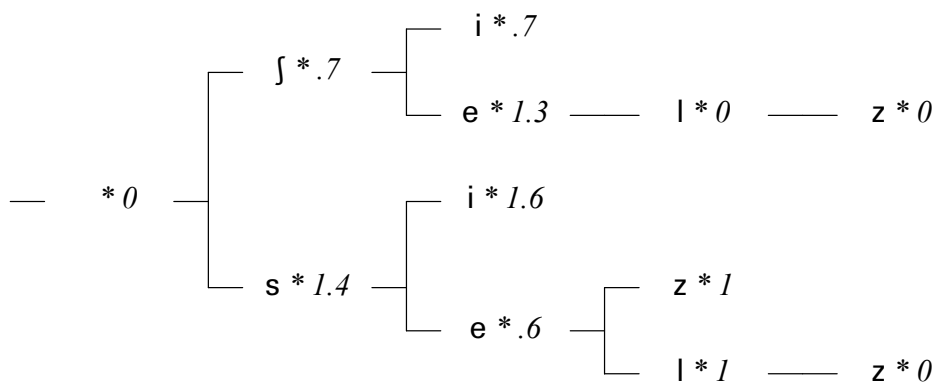
Take as a sample text *She says she sells sea shells she shells*, in IPA [ʃi sez ʃi selz si ʃelz ʃi ʃelz]. The frequencies of the vocabulary items are *she*/3, *says*/1, *sells*/1, *sea*/1, *shells*/2. Make a tree dictionary of the phonemic representation, marking the frequency of each word:

```
                      ┌─ i  ─  3
           ʃ ─┬─┤
              │   └─ e ── l ── z ─ 2
  ─┬─┤
   │  │        ┌─ i  ─  1
   │  └─ s ─┬─┤
              └─ e ─┬─ z ─ 1
                      └─ l ── z ─ 1
```

Sum the frequencies to the right of each node:

```
                      ┌─ i * 3
           ʃ * 5 ─┬─┤
                    └─ e * 2 ── l * 2 ── z * 2
  ── * 8 ─┤
           │         ┌─ i * 1
           └─ s * 3 ─┬─┤
                       └─ e * 2 ─┬─ z * 1
                                   └─ l * 1 ── z * 1
```

For each mother-daughter pair in the tree, the surprisal of the daughter is the logarithm of the mother's frequency divided by the daughter's frequency:

```
                      ┌─ i * .7
           ʃ * .7 ─┬─┤
                     └─ e * 1.3 ── l * 0 ── z * 0
  ── * 0 ─┤
           │          ┌─ i * 1.6
           └─ s * 1.4 ─┬─┤
                         └─ e * .6 ─┬─ z * 1
                                      └─ l * 1 ── z * 0
```

Here the logarithm is to base 2 so the unit is binary bits of information.

Note that for this text, the surprisal of [z] after [l] is zero, because [l] is always followed by [z] so there is no surprise: no information is conveyed.

The functional load of a phoneme is its percentage of the total information:

s	$1.4 \times 100 \div 8.3$	17%
ʃ	$.7 \times 100 \div 8.3$	8.4%
z	$(0+1+0) \times 100 \div 8.3$	12%
l	$(0+1) \times 100 \div 8.3$	12%
i	$(.7+1.6) \times 100 \div 8.3$	28%
e	$(1.3+.6) \times 100 \div 8.3$	23%

The metric could be improved by taking word-to-word transitions into account. Instead of the frequency of each word type calculated as the number of tokens, the sum of each token's probability in context could be a better base for the surprisal. I leave this for future implementation.

4. Phoneme inventory

Welsh has the following phonemes: /ptk bdg fθχ vð m̥n̥ŋ̊ mnŋ ɬr̥ lr sh iɨu eəo a/. Note that the voiced fricative series is lacking /ɣ/, which existed in Old Welsh but was lost to elision or vocalisation. Old and Middle Welsh had an additional vowel, usually analysed as /ʉ/, which later lost its rounding to merge into /ɨ/. The merger began in unstressed final syllables in the fifteenth century (Sims-Williams, 2013, note 21), though it was not complete until the eighteenth century. It is here projected anachronistically back into Middle Welsh to enable straightforward comparisons in the results below, except in the section where the merger's relation to functional load is analysed. In parts of south Wales the three vowels /i ɨ ʉ/ have all fallen together as /i/.

The only other relevant difference between the mediæval and modern language is in the addition of the marginal phoneme /ʄ/ in a few loan words from English. In the Bible text this phoneme is confined to personal and place names, the most frequent of which is *Job* which occurs 61 times. In the CEG text it occurs in a number of everyday words, the commonest of which are *project* (140 times) and *garage* /gareʄ/ (38 times).

5. Results

5.1. Checking the metrics   To check the internal consistency of each metric, the Welsh Bible text is divided into two, taking even and odd verses separately. It would be expected that each of the three metrics should give identical results for the two halves of the Bible. A scatter plot should and does line the phonemes up along the diagonal.

   In these and subsequent plots, the axes extend to 12½%, i.e. a phoneme at the margin has an eighth of the total frequency or information content.
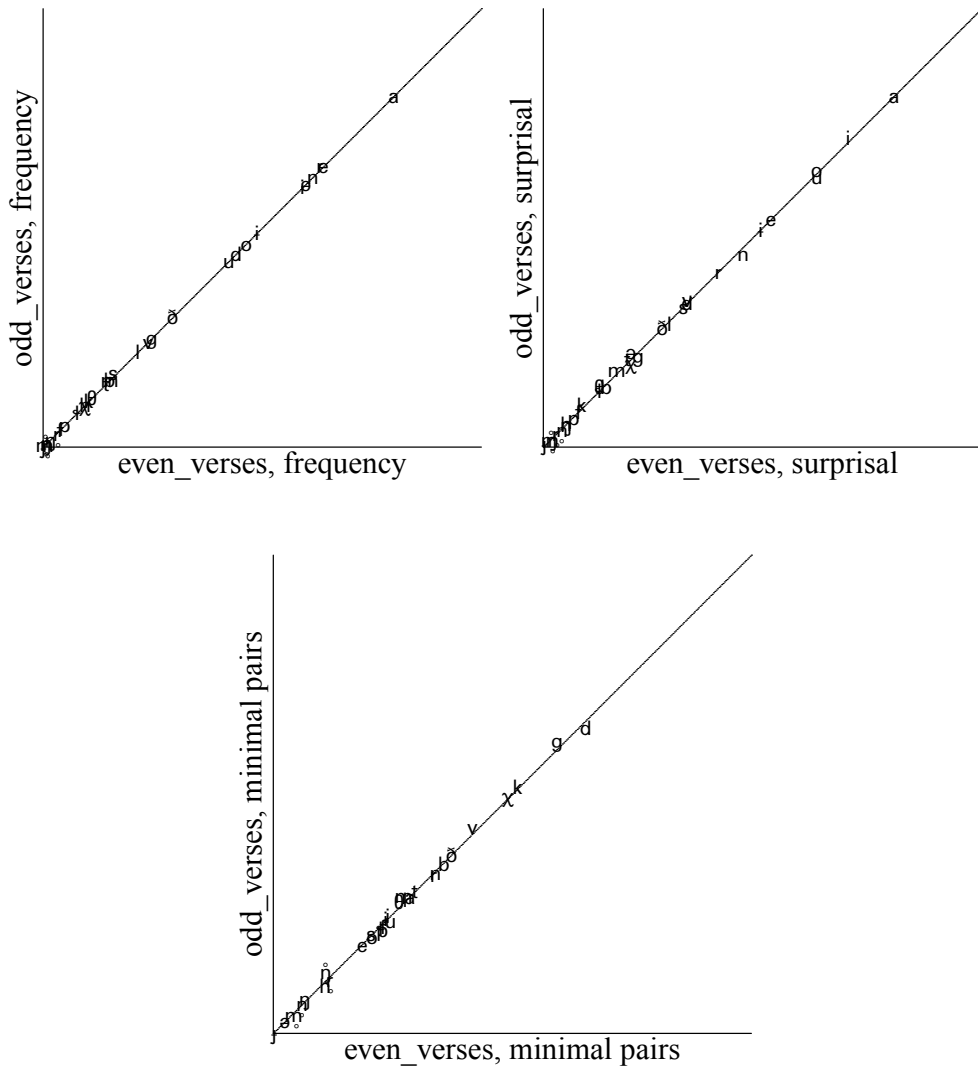
Following are the modern Welsh phonemes ranked using the Bible text from highest to lowest by:

frequency a e r n ə i ɨ o d u ð g v l s m b t θ k χ h ɬ p f r̥ ŋ n̥ ŋ̊ m̥ ɟ

surprisal a i u o e ɨ n r d s v l ð g t ə χ m b θ ɬ k f p h ŋ n̥ r̥ m̥ ŋ̊ ɟ

pairs d g χ k v ð b n r a t θ l m i ɨ u f p ɬ o s e ŋ̊ r̥ h ŋ n̥ m̥ ə ɟ

Measuring frequency or information content of phonemes in a dictionary seems uninformative, but minimal pairs have normally been taken to be dictionary words rather than inflected wordforms in a text, as is recommended by Brown (1988, p. 601). The following plot demonstrates how different the result is.
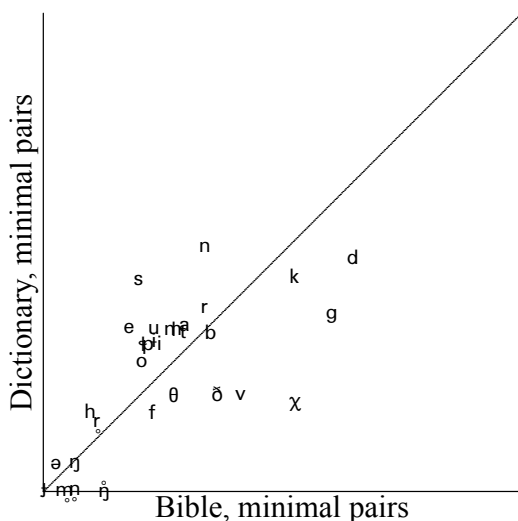

Different inflected forms of the same Welsh dictionary word often form minimal pairs, and the plot above shows that the consonants involved mostly have a higher functional load in the text-based measure than in the dictionary-based measure.

5.2. Functional load and frequency

The following scatter plots compare the frequency of each phoneme with its functional load in modern and medieval Welsh.
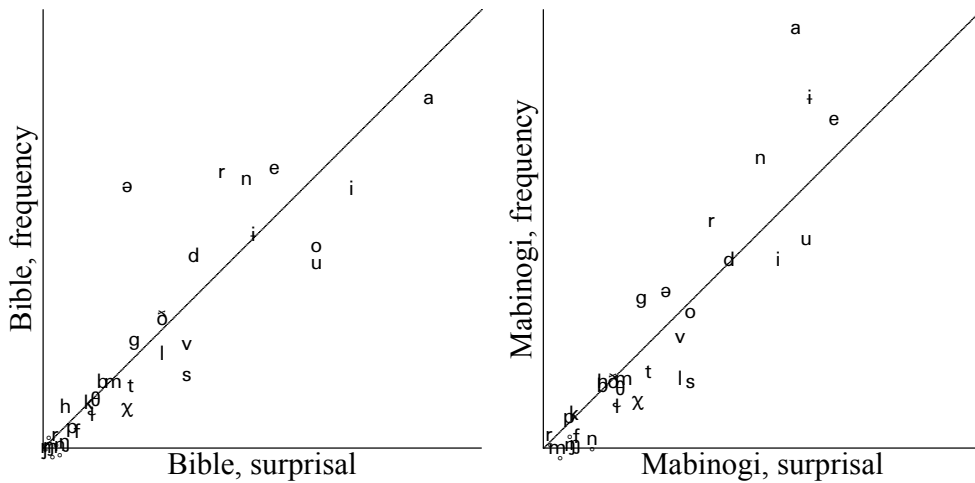
As Oh et al. (2015) note, the correlation between functional load and frequency is strong but not straightforward. As with the languages they looked at, in Welsh too functional load correlates roughly with frequency for most phonemes, but there are clear exceptions. Particularly noticeable in modern Welsh is /ə/ which is frequent but conveys little information. It occurs in a few frequent monosyllabic function words, but is otherwise restricted to non-final syllables of polysyllabic words. What has changed over the last seven centuries to reduce the information content of /ə/ slightly whilst making it much more frequent?

5.3. Minimal pairs and surprisal Here are compared the two ways of measuring functional load, by minimal pairs and by information content.
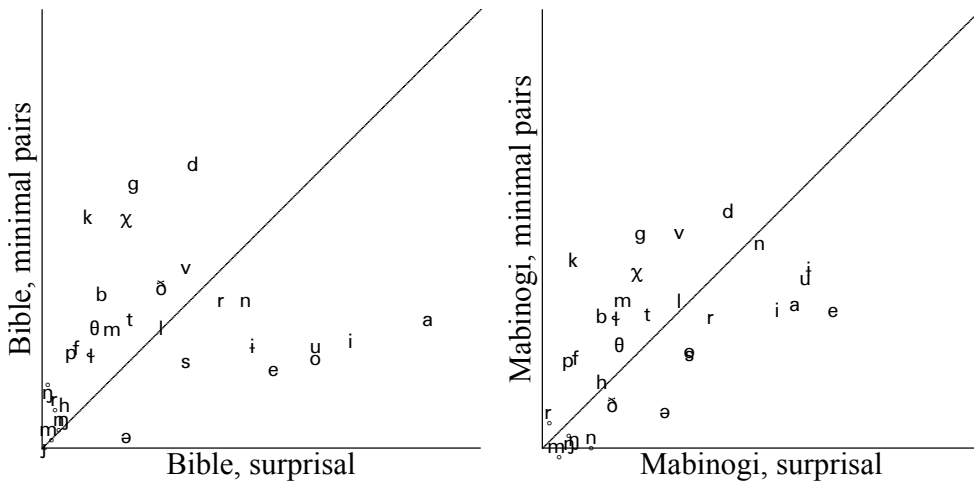
It can be seen that the correlation is weaker than between frequency and surprisal. Minimal pairs and surprisal are obviously not measuring the same linguistic property. The two concepts of functional load are quite different things.

5.4. Change over time First I compare two very different late twentieth century texts.
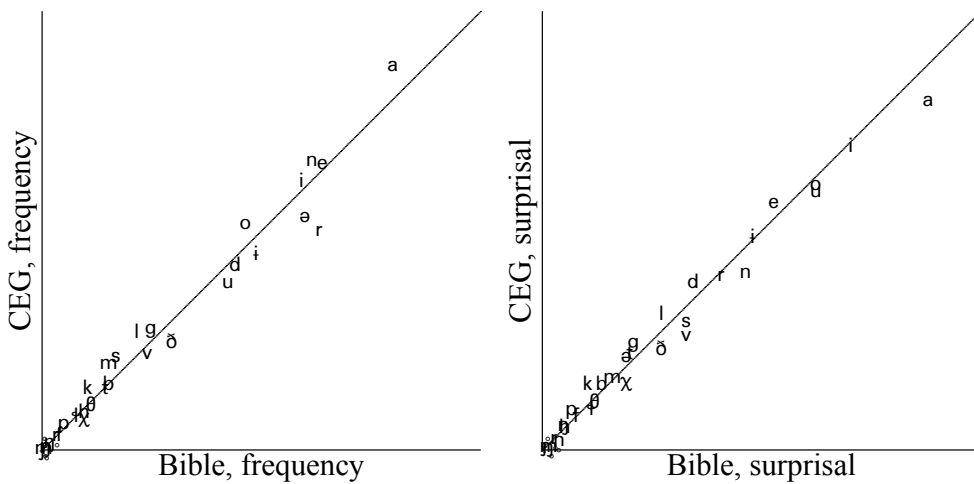
For both frequency and functional load, the correlation is not as strong as within the Bible text in §4.1 above, but is still very strong. The point here

is to show the range of variation that can exist between texts of the same language.

Next a comparison of two narrative texts separated in time by about seven centuries.
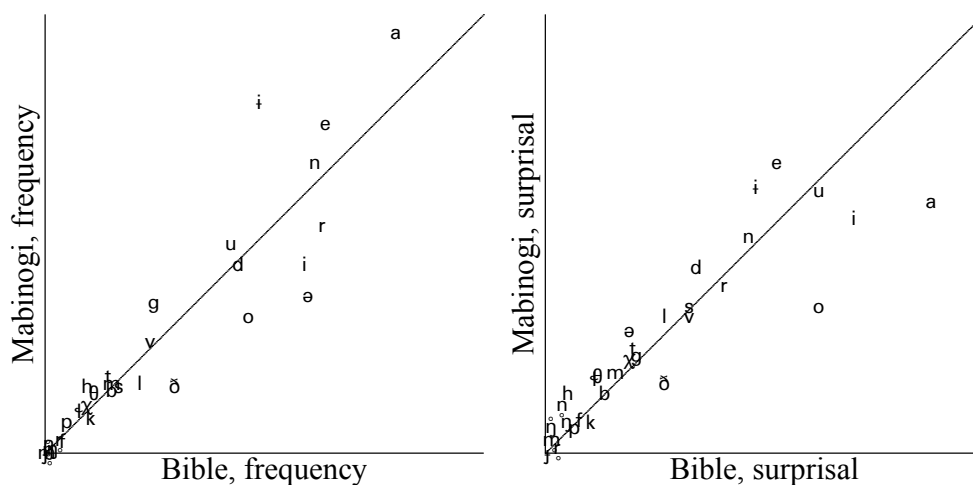
Taking the comparison of the two different modern texts above as an indication of the degree of variation that can be expected within different modern texts, it can be seen that the mediæval text is well outside this range. There has been a considerable change of both textual frequency and average information content for a number of phonemes over the seven centuries.

Future research will investigate individual changes, such as the seeming increase in frequency and information content of /ð/. It was noted above that /ð/ is underrepresented word-initially in the Middle Welsh text, appearing as /d/ because of errors in the restoration of pronunciation from spelling. However, even if the Bible text is used to artificially normalise the distribution of word-initial /ð/ and /d/ in the Middle Welsh text, the outlying status of /ð/ is only slightly reduced. The must be some other explanation for the increase over the centuries.

6. Phoneme merger and functional load

Middle Welsh had eight vowels: /iɨʉu eəo a/. There were two high central vowels, unrounded /ɨ/ and rounded /ʉ/. Later /ʉ/ lost its rounding to merge with /ɨ/, leaving the modern seven-vowel system /iɨu eəo a/.

A widespread idea, associated particularly with Martinet (1955), is that functional load influences language change, in that phonological contrasts with low functional load are more easily lost. As noted above, more recent research has tended to cast doubt on this hypothesis, so here I look at the Welsh merger. Using the surprisal metric for functional load and keeping the two vowels distinct, the ranking of phonemes from highest to lowest in the mediæval text is:

e u a i n d r ʉ ɨ s o l v ə t g χ m θ ɬ ð b h n̥ f ŋ k ŋ̊ p m̥ r̥

This text is from just before the merger began. For the modern Welsh of the CEG text, a couple of centuries after the merger completed, the ranking with /ɨ/ and /ʉ/ anachronistically distinguished is:

a i o u e r n d l ʉ s ɨ v g t ð ə m χ k b θ p ɬ f ŋ h r̥ n̥ m̥ ɟ ŋ̊

Both vowels have low functional load compared with the other vowels and phonologically they differ in only one feature (rounding). They would seem to have been good candidates for merger.

We can look at the functional load of the contrast between /ɨ/ and /ʉ/ using minimal pairs as Martinet might have. For both the Middle Welsh text and the CEG text with the contrast anachronistically restored, the /ɨ − ʉ/ contrast distinguishes by far the fewest minimal pairs of the twenty-one possible vowel contrasts. In the Middle Welsh text it distinguishes five minimal pairs out of a total of 554 distinguished by a vowel contrast: 0.9%. The raw figures for the modern text are 68 out of 4846, or 1.4%, but this is an overestimate as eleven of the pairs arise from misspellings (due to the identical pronunciation) which are not noted in the annotations attached to the CEG text. Typographical errors are corrected in the CEG with an annotation to note the correction, but these misspellings have been missed. A single example of misspelt *melus* (which would represent a historical /melʉs/) for correct *melys* /melɨs/ 'sweet', so spelt 27 times, makes a spurious minimal pair. Errors are many in the CEG text and annotations, but other types of error are likely to be distributed randomly, so 1.2% is probably a reasonable estimate.

Finally, in Middle Welsh /ɨ/ was more than twice as frequent as /ʉ/, so the result of the merger is the commoner of the pair, as Martinet hypothesised.

7. Conclusion

This paper has proposed and exemplified a more direct information-theoretic way of measuring the functional load of a phoneme or other phonological feature. It supports for a different language Oh et al.'s finding that the functional load of a phoneme measured information-theoretically correlates with textual frequency strongly but not consistently.

A comparison of the information-theoretic measure of functional load and the widely-used traditional measure based on minimal pairs, showed that these two measures do not correlate and seem to be measuring quite different linguistic properties.

A comparison of two narrative texts separated in time by seven centuries, showed considerable changes in frequency and functional load of phonemes. How these results relate to linguistic changes remains to be investigated.

Lastly a test was presented of Martinet's disputed hypothesis that mergers of phonemes are facilitated by low functional load. Data on a Welsh vowel merger showed that the vowels involved had comparatively low functional load compared with other vowels, when measured information-theoretically. The vowel pair involved was the least useful of all vowel contrasts in distinguishing minimal pairs. In this case then, the evidence is consistent with Martinet's hypothesis.

8. References

Adam Brown: 'Functional load and the teaching of pronunciation'. In the *TESOL Quarterly*, 22 (1988), pp. 593−606

——: 'Functional load'. In *Teaching English Pronunciation: A Book of Readings* ed. A. Brown, London : Routledge, 1991, pp. 221−224.

Cecil H. Brown, Eric W. Holman, Søren Wichmann, & Viveka Velupillai: 'Automated classification of the world's languages: A description of the method and preliminary results'. In *STUF — Language Typology and Universals*, 61.4 (2008), pp. 285−308.

David M. Carter: 'An information-theoretic analysis of phonetic dictionary access'. In *Computer Speech & Language*, 2 (1987), pp. 1−11.

J. C. Catford, 'Phonetics and the teaching of pronunciation: a systemic

description of English phonology'. In *Current Perspectives on Pronunciation: Practices Anchored in Theory*, ed. J. Morley, Washington : TESOL, 1987, pp. 87−100.

Collins-Spurrell Welsh Dictionary, Harper Collins, 1996.

Christophe Coupé, Egidio Marsico & François Pellegrino: 'Structural complexity of phonological systems'. In *Approaches to Phonological Complexity*, ed. F. Pellegrino, E. Marsico, I. Chitoran & C. Coupé, Berlin : Mouton de Gruyter, 2009.

Jules Gilliéron: *Genéalogie des mots qui designent l'abeille.* Paris : Champion, 1918.

Charles F. Hockett: *A Manual of Phonology*. Volume 21 of the *International Journal of American Linguistics*, Baltimore : Waverly Press, 1955.

——: 'The Quantification of Functional Load'. In *Word*, 23 (1967), 320−339.

Eric W. Holman, Søren Wichmann, Cecil H. Brown, Viveka Velupillai, André Müller & Dik Bakker: 'Explorations in automated language classification'. In *Folia Linguistica*, 42.2 (2008), pp. 331-354.

John Morris Jones: *A Welsh Grammar, Historical and Comparative.* Oxford : Clarendon, 1913.

Hirokatsu Kawashima: 'Frequency effects upon perceptual discrimination of vowel & consonant minimal pairs'. In 長崎外大論叢, 16 (2012), pp. 191−202.

Brett Kessler: 'Phonetic comparison algorithms'. In *Transactions of the Philological Society* 103 (2005), 243−260.

Robert D. King: 'A measure for functional load'. In *Studia Linguistica*, 21 (1967), pp. 1−14.

——: 'Functional load and sound change'. In *Language*, Vol. 43 (1967), pp. 831−852.

Mafuyu Kitahara & Shigeaki Amano: 'Functional load of pitch accent based on word familiarity.' In 日本音響学会講演論文集, 2-2-16 (2000), pp. 291−292.

André Martinet: 'Remarques sur le Système Phonologique du Français'. In the *Bulletin de la Société de Linguistique de Paris*, 34 (1933), 191−202.

——: *Économie des changements phonétiques*. Bern : Francke, 1955.

Vilém Mathesius: 'La structure phonologique du lexique du tchèque moderne'. In the *Travaux du Cercle Linguistique de Prague*, 1 (1929), pp. 67−84.

Jeff Mielke: 'A phonetically based metric of sound similarity'. In *Lingua*, 122 (2012), pp. 145−163.

Murray J. Munro & Tracey M. Derwing: 'The functional load principle in ESL pronunciation instruction: an exploratory study'. In *System*, 34 (2006), 520−531.

Yoon Mi Oh, François Pellegrino, Christophe Coupé, Egidio Marsico: 'Cross-language comparison of functional load for vowels, consonants, and tones'. In the *Journal of Phonetics*, 53 (2015), pp. 153−176

John D. Phillips: 'Cross-linguistically unusual features of Welsh are declining'. In *Current Issues in Unity and Diversity of Languages*, Linguistic Society of Korea, 2009, pp. 763−774

——: 'Automatic conversion between historical spelling, modern spelling, and IPA for Welsh'. In 山口大学文学会志, 63 (2013), 53−77.

Susanne Gahl, Yao Yao, Keith Johnson: 'Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech.' In the *Journal of Memory and Language*, 66 (2012), pp. 789−806

Patrick Sims-Williams: 'Variation in Middle Welsh conjugated prepositions: chronology, register and dialect'. In the *Transactions of the Philological Society*, 111:1 (2013), pp. 1−50.

Dinoj Surendran & Gina-Anne Levow: 'The functional load of tone in Mandarin is as high as that of vowels'. In the *Proceedings of the International Conference on Speech Prosody, 2004,* pages 99−102, Nara, Japan.

Dinoj Surendran & Partha Niyogi: 'Quantifying the Functional Load of Phonemic Oppositions, Distinctive Features, and Suprasegmentals'. In *Competing Models of Language Change: Evolution and Beyond*, ed. O. Nedergaard Thomsen, John Benjamins, 2006.

Kumiko Tanaka-Ishii & Hiroshi Terada: 'Word familiarity and frequency'. In

Studia Linguistica, 65 (2011), pp. 97−116.

Peter Wynn Thomas: *Peredur: Golygiad Lleiafol.* Online publication, 2000, *http://www.cardiff.ac.uk/welsh/contactsandpeople/academicstaff/thomas-peter.html*

Ifor Williams: *Pedeir Keinc y Mabinogi.* Caerdydd : Gwasg Prifysgol Cymru, 1930.