On the role of contrastivity in the development of the /e ∼ ɛ/ merger in Korean

Julien Eychenne
Department of Linguistics and Cognitive Science,
Hankuk University of Foreign Studies, South Korea

Tae Yeoub Jang
Department of English Linguistics and Language Technology,
Hankuk University of Foreign Studies, South Korea

May 2, 2018

This is a post-peer-review, pre-copyedit version of an article published in the *Journal of East Asian Linguistics*. The final authenticated version is available online at: http://dx.doi.org/10.1007/s10831-018-9171-7.

Abstract

This paper seeks to assess the role of functional factors in the development of the /e ∼ ɛ/ unconditioned merger that has taken place in Seoul Korean. We explore several alternative measures of functional load, one measure of phonetic distance evaluated on a perceptual scale, and a simple integrative measure, which combines functional load and phonetic distance into a single 'contrastivity index'. Our results show that, while both functional load and perceptual distance hint at the fact that /e ∼ ɛ/ was a weak contrast, this effect is clearest when these two factors are integrated together, especially when functional load is measured as phone-level entropy loss. The proposed metric thus offers a fruitful way of modeling the interaction between functional factors in the development of mergers.

Keywords – contrastivity, functional load, phonetic distance, merger, Korean

1 Introduction

Phonological inventories are dynamic systems which, over time, can acquire new sounds or lose existing ones. From a functionalist point of view, phonological contrasts that better serve the purpose of communication should be more likely to be preserved in a language because they fulfill an important signaling role in the identification of lexical distinctions; conversely, contrasts that are somehow less well suited for this purpose should be more prone to undergo a *merger*, i.e. to be lost in whole or in part. How the usefulness of a contrast should be measured is, however, an open question. Theories of sound change have often emphasized the importance of phonetic similarity in the development of mergers (e.g. Martinet 1955, 29-30 et passim; Labov 1994, 329; see also Garrett and Johnson 2013, 58-59). Indeed, most, if not all, mergers discussed in the literature involve segments which are phonetically similar. See among many others the mergers w ∼ ʍ and k ∼ x in Scottish English (Scobbie and Stuart-Smith, 2008); o ∼ ɔ before /r/ in American English (Labov, 1994, 315-316); ɛ̃ ∼ œ̃ in French (Martinet, 1955, 35); ɟ ∼ ʎ in Spanish (Navarro, 1967). This fact is consistent with the finding that phonological systems tend to favor segments that are maximally dispersed in acoustic space (see Beddor 1991; Flemming 2004; Lindblom 1986; Oudeyer 2005; Schwartz et al. 1997 among others), thus avoiding segments that are phonetically too similar.

Phonetic similarity is undeniably a crucial factor but, as it has been recognized early on, it is not sufficient; *functional load*, preliminarily understood as the relative importance of a phonological contrast, has been argued to be another significant factor driving sound change (Hockett, 1955, 211-220; Labov, 1994, 328; Martinet, 1952, 1955, 35-39). The underlying assumption, at least according to the simplest interpretation, is that a contrast is less likely to be lost if it distinguishes more words, especially frequent ones. The precise statement of this notion has been a matter of debate (Hockett 1955; King 1967; Wang 1967; see below), some scholars even questioning the relevance of this factor in sound change (e.g. King, 1967). Nevertheless, several recent studies have suggested that functional load, in one form or another, appears to play *some* role. Silverman (2010) investigated the effect of consonantal neutralization in Korean nouns and showed that existing neutralization processes, despite being numerous in this language, created unexpectedly little derived homophony when compared to other possible (but unattested) patterns of neutralization. This appears to suggest that there is an anti-homonymy effect, and that the (Korean) lexicon can tolerate a substantial amount of information loss as long as lexical distinctions are well preserved overall. Kaplan (2011) confirmed this general tendency in Korean using computer simulations. In the realm of vowels, Eychenne and Jang (2015) investigated the merger of /e/ and /ɛ/ in Korean and found that the phoneme-level functional load of this pair of vowels was among the lowest. They hinted that this low functional load, coupled with the small phonetic distance between the two vowels, may have triggered their merger, although they did not provide any indication as to how these factors might have interacted. Furthermore, typological investigations appear to provide further support to the role of functional load in mergers. In two related studies, Wedel et al. (2013a,b) investigated its role across nine languages (including Korean) and found that functional load (more specifically, as we shall see, minimal pair counts) appeared to be predictive of the likelihood of a pair of phonemes to merge; as expected, contrasts whose functional load is lower are more likely to merge. Although they differ in methodology and scope, these studies, taken together, strongly suggest that functional considerations play a role in the maintenance and loss of phonological contrasts, especially in Korean.

The goal of this paper is to test Eychenne and Jang's conjecture on the interplay between functional load and phonetic distance in the development of the /e ∼ ɛ/ merger in Korean. We set out to test the simple hypothesis that contrasts can be lost when they contribute little to the purpose of communication. We consider three alternative operationalizations of functional load, one measure of phonetic distance (measured on a perceptual scale), and a simple integrative measure, dubbed 'contrastivity index', which is the geometric mean of functional load and phonetic distance. Our results show that, while both functional load and perceptual distance hint at the fact that /e ∼ ɛ/ was a weak contrast, this effect is clearest when these two factors are integrated together than when they are considered in isolation.

The remainder of this paper is organized as follows: Section 2 provides an overview of the vowel system of contemporary Korean and its recent diachrony; Section 3 provides an overview of the factors that can contribute to contrastivity and lays out the formal details of our integrative metric; Section 4 reports on the results of an empirical study of the e ∼ ɛ merger in Korean based on the proposed method, whereas Section 5 discusses the significance of this proposal, as well as potential venues for future research.

2 The vowel system of Korean

This section provides an overview of the vocalic system of Korean for readers unfamiliar with this language. One will find general descriptions of Korean in Sohn (1999) and Lee and Ramsey (2000), and a thorough survey of its history in Lee and Ramsey (2011). Shin et al. (2013) give an overview of the sound system of Korean, while Ahn (1998) offers a fairly comprehensive treatment of its phonology.

2.1 Contemporary Seoul Korean

In its maximal extension, the ‘Standard Korean Pronunciation’, which roughly corresponds to the pronunciation of people living in Seoul, is assumed to possess 10 monophthongs, written ㅣ ㅔ ㅐ ㅟ ㅚ ㅡ ㅓ ㅏ ㅜ ㅗ in the Korean alphabet (*Hangeul*). Following Shin et al. (2013), these vowels will be transcribed /i e ɛ y ø ɯ ʌ ɑ u o/ respectively. Phonological descriptions traditionally recognize three degrees of height (high vs mid vs low), two degrees of backness (front vs back) and contrastive rounding (Ahn 1998, 34; Sohn 1999, 156).

This vowel system calls for several important remarks. First, the vowels /y ø/ are now pronounced by most speakers as diphthongs, often transcribed as /wi/ and /we/, respectively. The development of front rounded vowels in Korean is a relatively recent innovation since, according to Lee and Ramsey (2011, 294), they were absent in the 19th century (see also Hong 1988, 41). Given that they have now disappeared from the speech of young Seoul speakers, it seems that they were not stable members of the inventory for a long period of time, if ever at all.

Secondly, Korean is often assumed to have contrastive length. Even though there are a number of minimal pairs distinguished by length (e.g. /pam/ ‘night’ vs /paːm/ ‘chestnut’), this feature appears to play a marginal role and is restricted to the word-initial position (Shin et al., 2013, 152-153; Lee and Ramsey, 2011, 296-297). According to Sohn (1999, 14), the length distinction is now ‘obsolete in the younger generation’.

Last but not least, the /e ∼ ɛ/ pair of vowels, which is the focus of this paper, is widely reported to be the target of an ongoing unconditional merger (see Eychenne and Jang, 2015; Hong, 1988, 36-89; Moon, 2007, 11; Seong 2004, 461; Shin et al., 2013, 99-101), to such an extent that ‘almost no one in Seoul under the age of about fifty can tell the two sounds apart’ (Lee and Ramsey, 2011, 295). The quality of the resulting vowel is reported to be intermediate between /e/ and /ɛ/, that is to say /e̞/. In Labov’s (1994,

321) terminology, this seems to be a *merger by approximation*, i.e. a merger whereby two phonemes gradually get closer to each other and typically merge into an intermediate sound.

2.2 Origins of the e ∼ ɛ merger

There is a consensus, based on much evidence from various diachronic studies, that the vowels /e/ and /ɛ/ had in the recent past an independent phonemic status with a distinct pronunciation, which is reflected in their different orthographic notation (ㅔ and ㅐ respectively). It is also true, according to most recent phonetic studies (see Eychenne and Jang, 2015 and references therein), that no distinction between the two vowels exists when they are produced and/or perceived by speakers of the younger generation, leading scholars to infer that there is an ongoing merger. These facts raise an interesting question as to when the merger began to take place.

Understanding the genesis of Korean non-high front vowels starts from recognizing the dynamicity of the inventory of Korean vowels. Evidence from Old Korean seems difficult to interpret reliably, but diachronic studies have shown that from the 15th century, there was no mid front or low front monophthong in Korean (Sohn 1999, 46; Sohn 2012, 81; Lee and Ramsey 2011, 95). According to Lee and Ramsey (2011, 94), Mongolian words containing an /e/ which were borrowed in Early Middle Korean (around the 13th century) were consistently realized with a central vowel noted /ə/ in Late Middle Korean (15th century). The lack of mid front vowels is confirmed by more explicit evidence made available thanks to the establishment and promulgation, in the middle of the 15th century, of Hangeul, which adopted a phonetic writing system to replace logographic Chinese characters. Interestingly, there was no symbol in the list of the basic vowels of Hangeul to represent either /e/ or /ɛ/, demonstrating without ambiguity that no such vowels existed in Late Middle Korean.

Furthermore, morphological studies have revealed that the vowels /e/ and /ɛ/ have not been abruptly added to the inventory, but emerged as a consequence of monophthongization of off-glide diphthongs that mostly disappeared in contemporary Korean. The vowel change that started to take place in the late 18th century caused the diphthongs /əj/ and /aj/ to change into /e/ and /ɛ/, respectively. It has been suggested that the merger first occurred in non-initial syllables and was later extended to initial syllables[1]. For instance, Martin (2006, 25) and Lee and Ramsey (2011, 295) both hint at the fact that initial syllables were more resilient to this merger[2]. It is not clear whether it is the initial position itself, or length, which initially prevented vowels from merging in that position since, as we mentioned in 2.1, long vowels are only found in initial syllable. In any event, this merger, which was completed by the beginning of the 19th century, is considered to be a major yardstick that distinguishes between Middle Korean and Modern Korean (Lee and Ramsey, 2011, 161). Given the relatively recent incorporation of the monophthongs /e/ and /ɛ/ into the Korean vowel inventory, the subsequent merger of these two vowels into an intermediate vowel is striking.

Several scholars have suggested that the ongoing merger of the two vowels is likely to

[1]We thank the editors of the journal for drawing our attention to this fact.

[2]Interestingly, Yong Heo (p.c.) pointed out to us that some professional news anchors are able to distinguish /e/ and /ɛ/, but only in word-initial syllable.

have originated from the southeastern dialect spoken in the Gyeongsang Province (Ahn 1998, 57, Lee and Ramsey 2000, 326), and that it may be due to a massive influx of south-easterners into Seoul (Lee and Ramsey 2011, 295; Hong 1988, 43-46): Speakers of the southeastern dialect who had already lost the distinction would have affected the pronunciation of Seoul speakers, eventually causing the entire loss of this distinction in the dialect of Seoul Korean. For this hypothesis to be tenable, appropriate answers to the following two questions should be provided. First, why have other major dialects of Korean (e.g., the southwestern dialect spoken in the Jeolla Province) also undergone the same merger? Second, what makes the e ∼ ɛ merger of Gyeongsang Korean powerful enough to affect the system of Seoul Korean, whereas another widely known merger of ɯ ∼ ʌ in Gyeongsang Korean (see Hong 1988, 45; Sohn 1999, 156; Yeon 2012, 12) is not influential at all. Conclusive answers to these questions require a close empirical investigation in both synchrony and diachrony, which is beyond the scope of the current study. With respect to the second question, however, we hope that the computational approach developed in the following pages will provide some elements of answer.

3 Contrastivity

Traditional structuralist phonology was chiefly concerned with the identification of contrastive units and their allophonic realizations, and this approach still underlies much theorizing in phonology to date[3]. However, many scholars have pointed out the inadequacy of this approach, because it fails to take into account the relative importance of a contrast (e.g. Hall, 2013; Hockett, 1955, 213; Scobbie and Stuart-Smith, 2008). In this work, we follow a number of previous proposals which argue that the continuum between allophony and full contrastivity is best modeled quantitatively (Hall, 2009; Hume et al., 2013; Peperkamp et al., 2006).

3.1 Gradient contrastivity

Hall (2013) provides a comprehensive overview of the range of intermediate relationships between full contrastivity and allophony, and we shall take her discussion as a starting point. Which situations should be recognized as cases of gradient contrastivity depends on the criteria that are used to define contrastivity.

The first, most obvious criterion is *predictability of distribution* (Hall, 2013, 223). For any given pair of segments x and y, there are two extreme cases: (i) x and y always appear in the same contexts, in which case their distribution is strictly *equivalent*; (ii) x and y never appear in the same context, in which case their distribution is *non-intersecting*[4] (we reserve the term *complementary* for a non-intersecting distribution that is also allophonic, as explained below). Many sounds, however, have partially overlapping distributions: this is the case in particular when two segments contrast in a set of environments but the contrast is lost in some contexts, a phenomenon known as *contextual neutralization* – a typical example is word-final devoicing in German. Another criterion, which we view

[3]A full discussion of contrastivity is beyond the scope of this paper. See, among others, Hall (2009, 2013); Hume et al. (2013); Flemming (2004); Steriade (2007) and references therein.

[4]We borrow this term from Hockett (1966).

as a special case of predictability of distribution, is the existence of *lexical distinctions.* Two sounds *x* and *y* contrast in a language if substituting *x* with *y* creates a *minimal pair*, *i.e.* a pair of existing words that differ only by one sound, such as [tɑl] 'moon' vs [tʰɑl][5] 'mask' in Korean. The existence of minimal pairs is usually *sufficient* to assess the existence of a phonemic contrast, but it is not *necessary*; indeed, phonologists often use near-minimal pairs to establish that two sounds contrast. More generally, there exist many degrees of predictability (Hall, 2013, 230-234), and the existence of minimal pairs simply contributes to making the distribution more unpredictable.

Although non-intersecting and complementary distributions are often treated as one and the same thing, they must be distinguished because not all non-intersecting distributions are complementary. The criterion that is usually used to identify a complementary distribution is *phonetic distance.* Thus, the high phonetic dissimilarity between [h] and [ŋ] explains why those sounds are usually treated as two separate phonemes despite the fact that they have a (quasi-)non-intersecting distribution (Hockett, 1966, 9). These two criteria (predictability of distribution and phonetic distance) are fairly uncontroversial, although Hall (2013, 228) notes that some authors reject phonetic distance as a useful criterion to identify allophonic relationships.

A third criterion which has been appealed to is word *frequency*: everything else being equal, a contrast will be more salient if it appears in frequent pairs of words, such as *pack-back*, than in an infrequent ones, for example *prig-brig* (Martinet, 1978, 130; see also Carter, 1987; Hockett, 1955, 213). In addition, it is now well known that frequent forms tend to be realized with greater gestural overlap and less precise articulatory control, frequently leading to assimilation and phonetic reduction (see Hall, 2009, 82–89; Bybee, 2015; Ernestus, 2014 and references therein). Such frequency effects have been argued to differentially affect homophone words, which suggests that word forms are associated with frequency information in the mental lexicon (Gahl, 2008).

A fourth criterion that has been argued to play a role is the existence of *morpho-phonemic alternations* (Hall 2013, 224; see also Blevins 2004, §8.3 on the role of morphological paradigms). In Korean, laryngeal contrasts are neutralized in coda position, as in /piʨ/ 'debt' and /piʨʰ/ 'light' which are both realized as [pit̚]. Nevertheless, the contrast is still apparent in derived forms, such as [piʨɯl] 'debt.ACC'[6] vs [piʨʰɯl] 'light.ACC', and such alternations could protect the segments from merging completely.

Other potential criteria mentioned by Hall (2013) include native speaker judgments, orthography and possibly other factors that may rely on more abstract analyses of the phonology of a language. Native speaker judgments certainly provide an important source of evidence, but we believe they are a consequence, rather than a cause of the existence of a contrast. Furthermore, we view orthography as an external factor which may interfere with phonological factors, but does not enter into the definition of contrastivity per se.

To summarize the above, four criteria seem potentially relevant to a satisfactory definition of contrastivity: predictability of distribution, phonetic distance, frequency and the existence of morpho-phonological alternations. Morpho-phonological alternations seem to us rather difficult to integrate into a quantitative metric (but see Calamaro and Jarosz, 2015 for a proposal which relies on a combination of statistical learning and linguistic

[5]Korean has three series of (voiceless) stops, namely lax, aspirated and tense, as well as two series of fricatives (lax vs tense). In this paper, the diacritic [͈] is used to represent a tense articulation.

[6]Note that /ʨ/ further gets voiced between two sonorants.

filters). Since our goal is to test the role of contrastivity in the emergence of mergers, we chose to focus on a case of *unconditioned* merger[7]. As its name implies, an unconditioned merger is a merger that applies in all contexts, and which is not, by definition, sensitive to phonological or morphological structure. Focusing on this type of merger therefore allows us to disregard the potential effect of morpho-phonological alternations, and lets us focus more specifically on the other three factors. As we shall now see, functional load allows us to quantify the predictability of distribution, possibly taking into account the effect frequency, depending on how this notion is formalized.

3.2 Functional load

The observation that contrasts are not all equally important in a language led to the development of the notion of functional load (or *functional yield*, see Hockett 1955, 211–220, Martinet 1955, 35–39), which attempts to quantify the work done by a contrast in a system. Since functional load can be conceptualized in several ways (see Oh et al. 2015 and Wedel et al. 2013b for recent overviews), we will examine three alternative operationalizations of this notion.

The first definition of functional load that we will consider was proposed by Hockett (1955) and formulated within the framework of information theory, which was developed in the late forties (Shannon, 1948). Surendran and Niyogi (2003, 2006) refined Hockett's approach and applied it to the analysis of several languages, further consolidating his seminal work. Hockett (1955, 217) defined the functional load of a pair of phonemes, i and j, as follows:

$$\lambda(i,j) = \frac{H - H^*}{H} \tag{1}$$

where H is the *entropy* of the linguistic system L, *i.e.* the average amount of information (or equivalently, uncertainty) in the system, and H^* is the entropy of the corresponding pseudo-system L' where i and j have been replaced by a common segment in all contexts, everything else remaining constant. Hockett's metric can therefore be interpreted as the proportion of information that is lost when i and j are merged.

In order to make use of equation (1), it is necessary to discuss how entropy is calculated. Let $\mathcal{L}$ denote a sample representative of the language of interest. Its entropy $H(\mathcal{L})$ can be approximated by assuming that it was generated by a k^{th}-order Markov process, i.e. that the probability of occurrence of a unit only depends on the k previous units, the estimate of H becoming more precise as k grows larger. H is usually estimated from unigrams ($k = 0$), bigrams ($k = 1$) or trigrams ($k = 2$), using the following equation (Hockett, 1955, 217):

$$H(\mathcal{L}) = -\frac{1}{k+1} \sum_{x \in X} p(x) \log_2 p(x) \tag{2}$$

[7]As pointed out in Section 2.2, there is anecdotal evidence that this merger was not a true unconditioned merger since the word-initial position may have been more resilient to the change. But that seems to be the case for most, if not all, Korean mergers (see Lee and Ramsey, 2011), so this is the closest thing to an unconditioned merger in this language.

where k represents the order of the Markov process assumed to have generated the data, X represents the set of all n-grams in the corpus, and $p(x)$ is the probability of n-gram x in the data, which can be straightforwardly estimated from raw counts.

There are many ways in which Hockett's metric can be put to use in practice. First, different types of units (e.g. phonemes, syllables, words) can be used to compute the entropy of the system (Surendran and Niyogi, 2003, 2006). A merger between two sounds lowers the entropy of the system only if it decreases the number of n-grams at the chosen level of analysis, by merging contexts that would otherwise be distinct (Wedel et al., 2013b, 181). Second, one can use different levels of abstraction to represent a sample of the language (phonetic vs phonological transcriptions, types vs tokens). Here, we follow Hockett (1955) who assumed entropy to be a measure of uncertainty in the speech signal, and not at some more abstract level of representation. As a result, we will use phonetic transcriptions of word forms, taking into account their token frequency to estimate probabilities. We will consider two operationalizations of Hockett's metric: *phone-level entropy loss*, calculated using phone trigrams (*i.e.* with $k = 2$ in equation (2)), and *word-level entropy loss*, calculated using word unigrams.

Hockett's measure of functional load, in one form or another, has now been applied to the analysis of a number of languages (Oh et al., 2015; Surendran and Levow, 2004; Surendran and Niyogi, 2003, 2006). Nevertheless, as mentioned above, recent work has suggested that it is not the most useful measure of functional load when it comes to *predicting* mergers. In a cross-linguistic analysis of mergers, Wedel et al. (2013b) found that minimal pair counts were a significantly better predictor of the likelihood of a pair of segments to merge than were word-level or phone-level entropy loss. In a related study, Wedel et al. (2013a) further found that minimal pair counts appeared to be more predictive when they were calculated on lemmas within the same syntactic category, rather than on word forms or lemmas of different categories. These results call for a few remarks. First, it must be pointed out that most cases of mergers taken into account in these two studies are *conditioned mergers*, i.e. cases of contextual neutralization, rather than unconditioned mergers. Indeed, all cases of mergers which were considered for Korean are processes of consonantal neutralization, which are synchronically active – many of them have been stable for several centuries (Martin, 2006, 49). We remain agnostic as to whether synchronic neutralizations and diachronic mergers should receive a unified treatment, but it seems to us that their nature should nevertheless be controlled for. Secondly, while the distinction between lemma and word form is relatively straightforward in languages such as English, it is not clear to us that this distinction is equally helpful in an agglutinative language such as Korean, where the base often corresponds to a single syllable and can be accompanied by many affixes, as in /t͡ɕu+si+ʌs͈+tɑ+go/ 'give.HONORIFIC.PAST.STATEMENT.QUOTATIVE'. Unfortunately, Wedel et al. (2013a) had to set Korean (as well as Cantonese) aside since they did not have both lemmas and surface forms for these two languages, so that out of the seven remaining languages, six are Indo-European (including two varieties of English, and two other Germanic languages). Their results, while suggestive, would need to be confirmed on a genetically and typologically more balanced sample of languages. These remarks notwithstanding, the findings they report *do* deserve to be investigated further, and we included minimal pair counts on lemmas as an alternative measure of functional load.

To summarize the above, three measures of functional load will be considered: phone-level entropy loss (on tokens, with trigrams), word-level entropy loss (on tokens, with unigrams), and minimal pair counts (on lemmas, within the same syntactic category). These measures account for the distributional equivalence of segments in different but related ways. The token-based information-theoretic metrics measure functional load in the speech signal, whereas lemma-based minimal pair counts assume a more abstract level of representation in the mental lexicon. Furthermore, phone-level entropy loss uses more local information than the other two word-based measures since, according to the former, a merger between two phonemes a and b will affect the entropy of the language whenever they are found in the same local environment, for instance XaY and XbY, whereas according to the latter, a change only occurs when the two segments distinguish word forms (inflected forms or lemmas, depending on which measure is used).

3.3 Phonetic distance

Although phonetic distance can be estimated in a number of ways, in this paper we shall follow the line of inquiry which considers that perception plays a crucial role in driving sound change and shaping phonological systems (e.g. Ohala 1981; Lindblom 1986; Hume and Johnson 2001; Hume et al. 2013). The model that we adopt solely relies on information from the first three resonances (formants) of the vocal tract. In order to build an auditory representation of a vowel, raw frequencies (in Hertz) are first converted to normalized values on the bark scale, which accounts for the perceptual warping of frequency. Furthermore, to account for the way formants are integrated in perception, our perceptual model makes crucial use of the notion of effective second formant (F'_2) (see Bladon and Fant, 1978; Fant, 1973), which relies on the theory of spectral integration (see Beddor 1991 for an overview, and references therein). This theory predicts that when two formants are within a distance of about 3 to 3.5 bark of each other, they are perceptually integrated and perceived as a single spectral peak whose 'center of gravity' depends on the formants' center frequency and magnitude. F'_2 is often used in computational modeling of vowel systems to represent the contribution of higher formants (F_2, F_3, and possibly F_4) to vowel perception (e.g. de Boer, 2001; Oudeyer, 2005; Schwartz et al., 1997) and it has been argued to correlate with the size and shape of the vocal tract's front cavity (Hermansky and Broad, 1989). F'_2 can be estimated in a number of ways; here we rely on a formulation which only requires information from the first three formants (F_1, F_2, F_3), as proposed by Fant (1973, 52):

$$F'_2 = f_2 + \frac{\frac{1}{2} \times (f_3 - f_2) \times (f_2 - f_1)}{f_3 - f_1} \tag{3}$$

where f is the frequency, either in Hertz or mel. We depart slightly from Fant's original formulation and use the bark scale instead of the mel scale, since the former is now more widely used to model perceptual warping, especially for modeling F'_2 (de Boer, 2001).

To derive perceptual representations of vowels, frequencies for the first three formants (in Hertz) were converted to bark using Traunmüller's (1990, 99) classical formula. Next, using equation (3), we computed the (bark transformed) effective second formant (z'_2) for all vowels. We thus obtained for each vowel a two-point representation (z_1, z'_2), corresponding to the first formant and effective second formant, both expressed in bark.

In order to estimate the distance between vowels i and j, we used a weighted Euclidean distance, using the following equation (after de Boer, 2001, 49):

$$\delta(i,j) = \sqrt{(z_1^i - z_1^j)^2 + \lambda(z_2'^i - z_2'^j)^2} \qquad \text{with } \lambda = 0.3 \tag{4}$$

The parameter λ has the effect of attenuating the influence of the effective second formant. Its value is based on typological observations about vowel systems and perceptual experiments (de Boer, 2001, 49-50; Schwartz et al., 1997), which together suggest that F_2' contributes less than F_1 to vowel identification.

For the sake of clarity, we will illustrate the whole procedure with a concrete example. Consider the pair of vowels e ∼ ɛ, which have the following formant values (in Hz): $F_1 = 490$, $F_2 = 1968$, $F_3 = 2644$ for /e/, and $F_1 = 591$, $F_2 = 1849$, $F_3 = 2597$ for /ɛ/. (Formant measurements for all vowels are reported in Table 1 below.) We first convert these values to bark, which yields $z_1 = 4.83$, $z_2 = 12.90$, $z_3 = 14.87$ for /e/, and $z_1 = 5.68$, $z_2 = 12.48$, $z_3 = 14.75$. Next, we derive the effective second formant for both vowels from equation (3), which yields the values $z_2' = 13.69$ for /e/, and $z_2' = 13.33$ for /ɛ/. Finally, we obtain the weighted Euclidean distance between /e/ and /ɛ/ by plugging the values for z_1 and z_2' into equation (4), as follows:

$$\begin{aligned} \delta(e,\varepsilon) &= \sqrt{(z_1^e - z_1^\varepsilon)^2 + 0.3 \times (z_2'^e - z_2'^\varepsilon)^2} \\ &= \sqrt{(4.83 - 5.68)^2 + 0.3 \times (13.69 - 13.33)^2} \\ &\approx 0.87 \end{aligned} \tag{5}$$

3.4 Contrastivity index

We have argued that at least two fundamental factors may contribute to contrastivity: distributional equivalence (measured more or less locally, and more or less abstractly), and phonetic distance (measured perceptually). A number of recent proposals point in the same direction. Peperkamp et al. (2006) used a combination of statistical learning and linguistic filters to infer allophonic rules in a corpus of French. This approach is well suited to identify allophonic rules, but it cannot be used as a measure of contrast in the sense that it does not integrate distributional and phonetic factors. More recently, Hume et al. (2013) developed an information-theoretic analysis of vowel epenthesis, arguing that ideal epenthetic vowels have *weak contrastivity*: they do little work to keep lexical items apart and, at the same time, are minimally distinct from other vowels from a perceptual point of view. To determine the strength of a contrast, the authors considered both an information-theoretic formulation of functional load, and perceptual distance, measured as a function of miscategorization probability. Our approach builds upon this notion of weak contrastivity, and takes it one step further by proposing an explicit integrative metric, instead of considering each dimension separately[8].

We define the contrastivity index (K) of a pair of sound units as the *geometric mean* of its functional load and phonetic distance. Like the more common arithmetic mean, the geometric mean is a measure of central tendency, but it is the only one which is

[8]In a related study, Hume and Mailhot (2013) did propose an explicit integrative measure, but they did not put it to a test.

appropriate when averaging data points that are normalized or measured on different scales. Formally, the geometric mean corresponds to the n^{th} root of the product of n numbers. In our case, this corresponds to:

$$K(i,j) = \sqrt{\lambda(i,j) \times \delta(i,j)} \tag{6}$$

where λ and δ represent functional load and phonetic distance, respectively. In other words, contrastivity corresponds to the square root of the product of functional load and phonetic distance. (It would of course be possible to integrate other factors into this index, an issue to which we will return in Section 5.) As a matter of example, consider the case where functional load is measured in terms of minimal pair counts. As we report in Section 4.2.1 below, there are 406 minimal pairs for e $\sim$ ɛ in our corpus, and the phonetic distance between these vowels is 0.87 bark, as we have seen in the previous section. The corresponding contrastivity index is therefore $\sqrt{406 \times 0.87} \approx 18.79$.

In order to better appreciate the interaction between functional load and phonetic distance, Figure 1 provides a three-dimension plot of contrastivity as a function of phonetic distance and functional load. As can be seen, contrastivity is highest when both phonetic distance and functional load are high and, conversely, it is lowest when both are low. As should by now be clear, the notion of contrastivity that emerges is not categorical, but gradient: a pair of segments is more or less contrastive depending on the relative numerical value of its contrastivity index[9]. As the index gets lower, the distribution of the segments gets closer to an allophonic distribution. In the limiting case, K becomes zero if the segments are phonetically identical or if they are in a strictly non-intersecting distribution. Note that when $\lambda = 0$, phonetic distance becomes irrelevant, and the metric no longer distinguishes between pairs of segments which are phonetically close and segments which are phonetically very distinct. Whether such a situation actually occurs in real corpora, large enough to be representative, is an empirical question that deserves to be investigated. Should this prove to be a problem, one could simply add a small positive constant to equation (1), in order to ensure that λ is never null.

Before going further, it is important to clarify the status of the proposed metric: Is it simply a practical tool which conveniently summarizes the contribution of functional load and phonetic distance, or is it a measure which has an independent ontological status? Although this metric could indeed be used as a descriptive device, it should be regarded as an attempt to model contrastivity as the result of the *interaction*, rather than the mere *contribution*, of distributional and phonetic factors. We view contrastivity as an emerging property of the interaction of these more fundamental factors, which can weaken or strengthen each other, rather than an autonomous entity. Whether this approach, or some variant thereof, is ultimately correct is of course an empirical question, but we believe that it is one which is worth asking. We shall return to this issue in the discussion, but we will first put this approach to a test.

[9]Contrastivity indices should not be interpreted in absolute terms. They are relative values that are only meaningful within a given system, since they depend on how each dimension is measured.

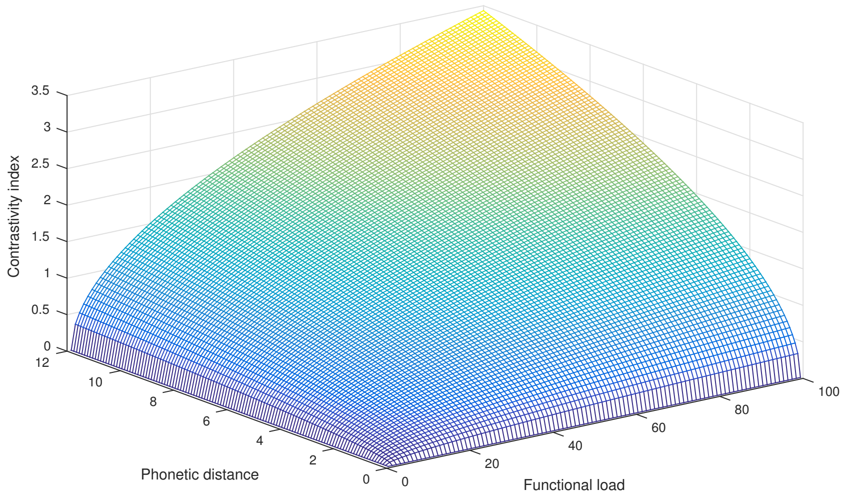

Figure 1. Contrastivity index as a function of phonetic distance and functional load

4 Analysis

4.1 Data

4.1.1 Textual data

In order to estimate the functional load of all pairs of monophthongs in Korean, it is necessary to use a corpus that is a reasonable approximation of the language. To this end, we chose the Korean National Corpus (better known as the Sejong corpus, see Kim 2006) which is a reference corpus for Korean available in digital form. The corpus contains a large amount of annotated textual data written in Hangeul. It contains two core subparts: a spoken corpus, which contains orthographic transcriptions of conversations and interviews, and a written corpus, which includes material such as press articles, textbooks, novels and poems from the 20th century. We used the written part of the corpus because it is much larger and the spoken corpus contains many free variants for the vowel pair o ∼ u (e.g. 사둔 [sadun] ∼ 사돈 [sadon] 'in-law relationship'; 하구 [hagu] ∼ 하고 [hago] 'after doing'), which would unduly have biased our results for functional load.

In order to obtain both tokens and lemmas, the following procedure was implemented. All word forms were first extracted, along with their corresponding frequency. Next, each word form was phonetized by first converting the orthographic form into a phonological form using a simple letter-to-phoneme algorithm, and then applying the phonological rules described in Shin et al. (2013, 178-198). This corpus yielded 2,734,643 types, representing 33,912,209 tokens all together. This dataset constitutes our *token corpus*. In order to ensure that trigrams could be computed for words shorter than three segments, we also added a word boundary symbol at the beginning and end of each word form. As a result, a word such as 달 /tɑl/ 'moon' is phonetized as #tɑl#, and contains the following trigrams: <#tɑ>, <tɑl>, <ɑl#>. This word shares the trigrams <#tɑ> and <tɑl> with the form #tɑli# 'moon.NOM', and shares the trigram <ɑl#> with #tʰɑl# 'mask' and /#t͈ɑl#/

‘daughter’.

To build the *lemma corpus*, we used the same part of the Sejong database, but processed it differently. First, we annotated the whole corpus with parts of speech (POS) information using KoNLPy (Park and Cho, 2014), set up with the KKMA tagging engine (Lee et al., 2010). Next, we extracted all unique { lemma, POS } pairs in the tagged corpus, discarding borrowings and grammatical morphemes. Therefore, inflected forms derived from the same base (e.g. /ton+i/ ‘money.NOM’, /ton+ɯl/ ‘money.ACC’) were collapsed into a single lemma (in this case, /ton/). In order to distinguish homonyms within a given syntactic category, we used Naver’s online dictionary[10], which lists all homonyms for a given word, ranked in order of decreasing frequency. The most frequent words are indicated with a number of stars (up to three); less frequent words are simply listed without any stars. For instance, the form 가격 /kakjʌk/ corresponds to three nouns, meaning ‘price’, ‘strike’ and ‘lineage’. The first one (price) is assigned two stars, whereas the other two have no stars. We adopted the following strategy: we extracted all the forms that were starred, together with the first unstarred form. This ensured that we extracted at least one form for rare words[11]. Finally, since Naver also provides phonetic forms, we used these forms directly as the phonetic representation of the lemma. The final lemma corpus contains 63,835 items. This corpus was used to calculate minimal pair counts: counts were obtained independently for each grammatical category, and then summed over all categories. To be clear, two words were treated as a minimal pair if and only if they met the following conditions: (a) they belonged to the same syntactic category, (b) their *surface* forms had the same length and (c) their surface forms differed by exactly one phoneme. For instance, [tɑl] forms a minimal pair with [tʰɑl] ‘mask’, but also with [tɑk] ‘chicken’, even though the latter has the phonological form /tɑlk/. However, [tɑl] and [tɑ] ‘everything’, even though they belong to the same grammatical category in Korean, were not treated as a minimal pair since their lengths differ. Finally, note that the dissimilarity between the mismatched phonemes, be it expressed in terms of acoustic distance or feature differences, was not taken into account for the purpose of counting minimal pairs. For example, pairs such as [tɑl] *vs* [mɑl] ‘horse’ on the one hand, and [tɑl] *vs* [tʰɑl] on the other, were treated identically.

4.1.2 Acoustic data

In order to get an accurate representation of the contrastivity of a pair of vowels in a vocalic system, one would ideally need acoustic data from the period of time before the merger took place. This is unfortunately impossible in our case since the onset of this merger most likely predates the development of reliable acoustic measurement techniques (see Section 2.2). As a consequence, we have to resort to use measurements from talkers that still display the contrast to some extent.

Umeda (1995) provides acoustic measurements from 18 Seoul speakers born between 1916 and 1964. Ten of these speakers appear to display a contrast (2 females and 8 males). The average F_1 values for the 8 male speakers are 426 (45) Hz and 578 (52)

[10]See `http://krdic.naver.com`.

[11]We submitted a random sample of 20 words to a native Korean speaker not involved in this study, and she confirmed that the distribution of stars (and absence thereof) in Naver matched her intuitions regarding the relative frequency of the homonyms.

Hz for /e/ and /ɛ/, respectively. Unfortunately, we could not use these data since we identified issues with the reported values of F_3 for the back rounded vowels /o/ and /u/, which are implausibly low (< 1500 Hz). Instead, we used data from Yang (1992, 1996), who provides acoustic measurements of F_1, F_2 and F_3 of 10 male and 10 female speakers of the Seoul Korean dialect. Recordings were done in the early nineties in the United states, but subjects were all Korean students (from 18 to 27 years old) who spent most of their life in Korea. We only used data from Yang's male subjects because /e/ and /ɛ/ appeared to be much closer for females, which suggested that the merger was more advanced in the latter group. Yang reports the following average values[12] for F_1 (the standard deviation is given in parentheses): for males, /e/: 490 (105) Hz and /ɛ/: 591 (75) Hz; for females, /e/: 650 (113) Hz and /ɛ/: 677 (108) Hz. The mean difference in F_1 is only 27 Hz for females (much less than the standard deviation of either vowel) whereas it is 101 Hz for males. Although Yang's subjects are younger than Umeda's and represent a state of the language where the merger seems to be more advanced, Yang's data are the best approximation of a pre-merger state that we could use, since our acoustic model relies on F_3 to estimate the effective second formant.

The formant values from Yang's male speakers are given in Table 1. (Phonetic symbols have been adapted to match those used in this paper.) The front rounded vowels will be ignored since, as mentioned in Section 2.1, they did not remain stable phonemes for very long in the history of Korean and, additionally, they have followed a trajectory of successive monophtongization and diphtongization, which is not the focus of this paper. By the same token, length will also be disregarded since it now plays a marginal role, and no data was available to reliably estimate its role.

Table 1. Mean and standard deviation of F_1, F_2, F_3 for Seoul male speakers (after Yang 1996, 251)

	F_1 (SD)	F_2 (SD)	F_3 (SD)
ɑ	738 (87)	1372 (124)	2573 (127)
ɛ	591 (75)	1849 (106)	2597 (110)
e	490 (105)	1968 (150)	2644 (94)
i	341 (29)	2219 (176)	3047 (146)
o	453 (47)	945 (134)	2674 (156)
(ø)	459 (69)	1817 (163)	2468 (134)
u	369 (43)	981 (141)	2565 (173)
(y)	338 (30)	2114 (140)	2729 (213)
ʌ	608 (76)	1121 (110)	2683 (145)
ɯ	405 (37)	1488 (176)	2497 (80)

4.2 Results

We now turn to the analysis of these data. We will first consider the three measures of functional load discussed in Section 3.2 and phonetic distance independently, and then see how they integrate to quantify the contrastivity of Korean monophthongs.

[12] Vowels were recored in hVd context and measured at $\frac{1}{3}$ of the vowel's duration (Yang, 1992, 2281).

4.2.1 Minimal pair counts

As we have seen, Wedel et al. (2013b) have suggested that minimal pair counts appear to be the best measure of functional load in their cross-linguistic study. Table 2 reports the results on the lemma corpus, summed over morphological categories[13]. (In this table and all the following tables, e ∼ ɛ is highlighted in bold.)

Table 2. Minimal pair counts for Korean monophthongs (lemmas) (in %)

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		521	3331	1035	2448	2895	1499	3608
ɯ			387	58	604	558	243	1138
u				1174	2396	3132	1255	4360
e					476	964	**406**	1060
ʌ						1917	687	3988
o							1973	3945
ɛ								1665
ɑ								

Visual inspection of the data by means of quantile-quantile and probability density plots reveals that the data deviate from normality, and the distribution is in fact bimodal. This is consistent with the cross-linguistic observation made by Oh et al. (2015) that the distribution of functional load (here, as measured by minimal pair counts) tends to focus on a few contrasts. Indeed, out of the 28 pairs of monophthongs, the first 8 pairs in our data represent 58.1% of the minimal pairs. The pair with the highest functional load according to this measure, namely u ∼ a, accounts for 9.1% of the minimal pairs alone.

Since the data are not normally distributed, we report the median ($\tilde{x} = 1214.5$) instead of the mean. The minimal pair counts range from 58 for ɯ ∼ e to 4360 for u ∼ ɑ. The pair e ∼ ɛ is the fourth lowest pair ($N = 406$), above ɯ ∼ u, ɯ ∼ ɛ and ɯ ∼ e[14]. Minimal pair counts do suggest that the functional load of e ∼ ɛ is relatively low, although it is still noticeably higher than that of ɯ ∼ e ($N = 58$).

4.2.2 Word-level entropy loss

Next, we investigate how word-level entropy loss fares compared to minimal pair counts. Both measures are similar, but the former also takes into account token frequency and was calculated using word forms across all syntactic categories instead of lemmas within a syntactic category. The results are provided in Table 3 and were converted to percentages to improve readability.

The results are fairly similar to those obtained for minimal pair counts. The median across monophthongs is $\tilde{x} = 0.174\%$. The pair with the lowest score is still ɯ ∼ e (0.025%)

[13]Examples of minimal pairs that were thus obtained include, for /e/ *vs* /ɛ/, [peda] 'to mow' *vs* [pɛda] 'to abdicate', [seu] 'stew' *vs* [sɛu] 'shrimp', [te] 'place' *vs* [tɛ] 'stand'...

[14]It is interesting to note that /ɯ/ has a somewhat special status in Korean since it is the default epenthetic vowel in loanword adaptation (*truck* → [tʰɯɾʌk]), it is prone to deletion in hiatus (/kʰɯ+ʌ/ → [kʰʌ] 'big.PLAIN REGISTER'), it tends to be phonetically short and is often devoiced between voiceless obstruents or in word-final position.

Table 3. Word-level entropy loss (%) for Korean monophthongs (tokens)

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		0.165	0.237	0.391	0.220	0.330	0.123	0.435
ɯ			0.058	0.025	0.095	0.073	0.048	0.179
u				0.100	0.193	0.190	0.107	0.333
e					0.101	0.203	**0.073**	0.170
ʌ						0.185	0.087	0.348
o							0.146	0.395
ɛ								0.198
ɑ								

and e ∼ ɛ is now the fifth lowest pair (0.073%), just above ɯ ∼ o (0.073%). The pair u ∼ ɑ, which has the largest minimal pair count, is now the fourth highest (0.333%), the pair with the highest score being i ∼ a (0.435%). Word-level entropy loss also suggests that e ∼ ɛ is one of the pairs with the lowest functional loads.

4.2.3 Phoneme-level entropy loss

The third measure of functional load considered in this paper is phone-level entropy loss, computed on trigrams using the token corpus. Table 4 shows the results for all vowel pairs.

Table 4. Phoneme-level entropy loss (in %) for Korean monophthongs (tokens)

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		0.392	0.351	0.371	0.366	0.354	0.170	0.597
ɯ			0.108	0.127	0.172	0.167	0.071	0.302
u				0.116	0.298	0.279	0.128	0.442
e					0.123	0.192	**0.089**	0.224
ʌ						0.314	0.153	0.708
o							0.176	0.515
ɛ								0.355
ɑ								

The median value is $\tilde{x} = 0.247\%$. The pair whose functional load is highest according to this measure is ʌ ∼ ɑ. Like the word-based metrics discussed above, this measure assigns a relatively low functional load to e ∼ ɛ. It seems to do a slightly better job than the previous measures since it ranks e ∼ ɛ as the second lowest pair (0.089%), above ɯ ∼ ɛ (0.071%), which is now the lowest ranked pair.

To get a better sense of the similarity between these three measures, we assessed the amount of correlation between minimal pair counts and word-level entropy loss on the one hand (Figure 2), and minimal pair counts and phone-level entropy on the other (Figure 3)[15]. The correlation was evaluated in both cases by means of Spearman's rank correlation coefficient, and was found to be quite high in both cases, $r = 0.83$ for minimal

[15]The filled circle in both plots represents e ∼ ɛ.

pair counts and word-level entropy, and $r = 0.75$ for minimal pair counts and phone-level entropy.

It is not clear whether the differences observed among the lowest ranked vowels for a given metric, and across metrics, are linguistically meaningful. Nonetheless, all three metrics do hint at the fact that functional load may have been a contributing factor in the development of the e ∼ ɛ merger.

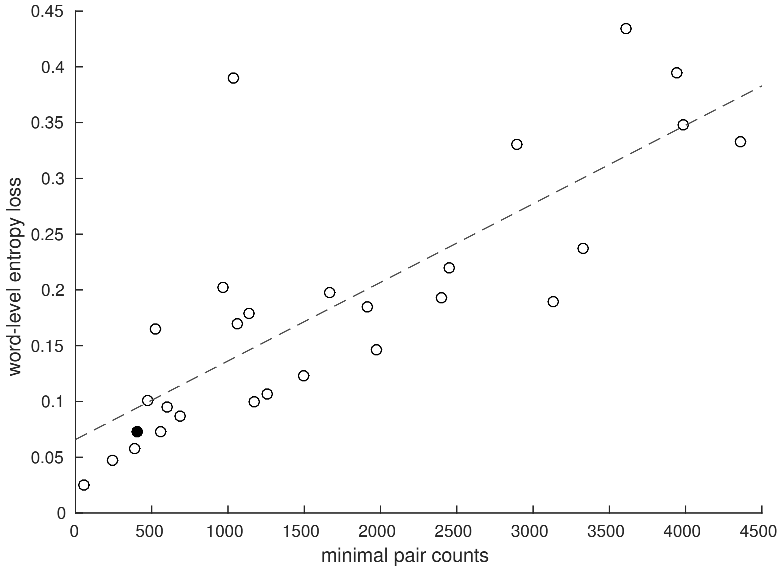

Figure 2. Correlation between minimal pair counts and word-level entropy loss

4.2.4 Perceptual distance

We have just seen that all the functional load metrics rank e ∼ ɛ among the lowest pairs of vowels. We now turn to the discussion of the perceptual distance based on the model presented in Section 3.3 and using the data from Yang (1996) reported in Table 1. The distance between all pairs of vowels (in bark)[16] is shown in Table 5.

The median distance was $\tilde{x} = 2.12$ bark, with values ranging from 0.81 to 3.76 bark. According to the model, the vowels perceptually closest are o ∼ u whereas the most distant vowels are i ∼ ɑ. As expected, the pairs u ∼ ɑ and i ∼ u are also among the most distant pairs, being ranked third and seventh respectively. Figure 4 provides a perceptual map obtained via multidimensional scaling of the distance matrix in Table 5. Axes have been rearranged in order to mirror the orientation familiar from articulatory maps and $F_1 \times F_2$ formant charts.

The e ∼ ɛ pair is the second lowest (0.87 bark) and is only slightly higher than the distance between o ∼ u. The fact that /o/ and /u/ are the most similar vowels is particularly interesting since this is a fact that has been previously noted in the literature. Hong (1988, 61-67) identifies o ∼ u as a 'false merger'. The results from her sociolinguistic investigation of Seoul Korean suggest that although the two vowels are acoustically very

[16]Although perceptual distances are expressed in bark, bear in mind that they are measured using a weighted Euclidean distance which dampens the effect of the effective second formant.

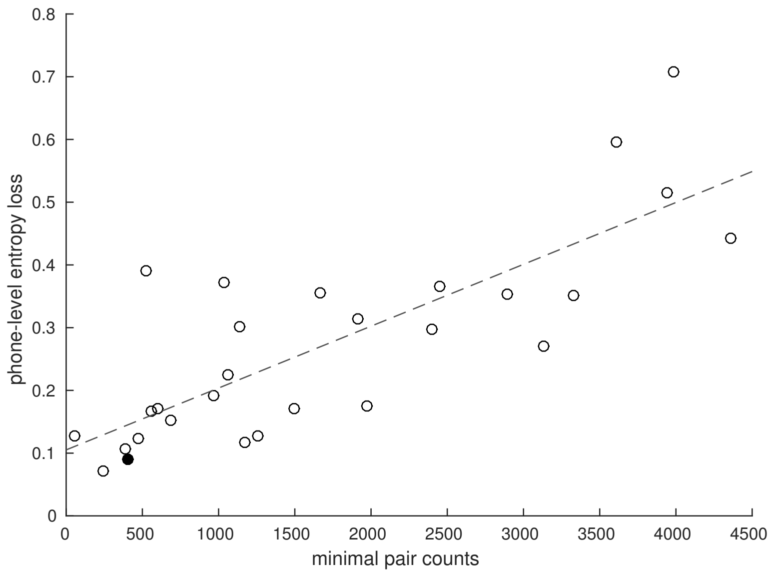

Figure 3. Correlation between minimal pair counts and phone-level entropy loss

Table 5. Perceptual distance for Korean monophthongs (in bark)

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		1.44	2.65	1.47	3.34	3.03	2.34	3.76
ɯ			1.38	1.13	2.04	1.6	1.74	2.77
u				2.43	2.12	0.81	2.77	3.22
e					2.11	2.38	**0.87**	2.31
ʌ						1.41	1.67	1.18
o							2.46	2.57
ɛ								1.51
ɑ								

close, listeners could nevertheless identify them correctly in a minimal pair test (e.g. [oɾi] 'duck' vs [uɾi] 'we'). More recent work has also shown that the distance between /o/ and /u/ has decreased among young Seoul speakers, especially for female speakers (Han and Kang, 2013). In addition, it is worth remembering that a number of words that originally had an /o/ now also have a free variant with an /u/, as noted in Section 4.1.1. This situation looks very much like an ongoing *merger by transfer* in Labov's terminology, i.e. a 'unidirectional process in which words are transferred gradually from one phonemic category to another' (Labov, 1994, 321).

Regarding the small score difference between the pairs e ∼ ɛ and o ∼ u, it must be emphasized that the acoustic data that we used represent contemporary Korean, where the merger is in an advanced stage, and in fact completed for many speakers. Since the merger seems to be in favor of an intermediate [e̞], it is plausible that the perceptual distance between /e/ and /ɛ/ was greater when the merger was initiated than it is nowadays, and this is indeed what Umeda's (1995) measurements suggest. Still, o ∼ u and e ∼ ɛ are the two closest pairs of vowels, and although their phonological confusion appears

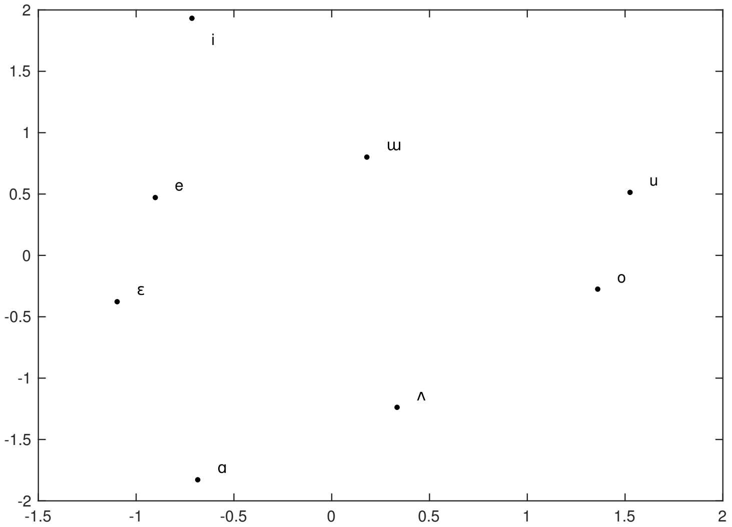

Figure 4. Perceptual map of Korean monophthongs

to follow a different path (merger by transfer vs merger by approximation, respectively) and e ∼ ɛ is in a far more advanced stage, perceptual distance most likely played a role in both mergers.

4.2.5 Contrastivity index

We now report on the contrastivity index of each pair of vowels, integrating functional load and phonetic distance into a single metric, as explained in Section 3.4. Since phone-level entropy loss ranked e ∼ ɛ lowest, we will use it as a reference measure. The results are shown in Table 6 and the corresponding contrastivity map, obtained by multidimensional scaling of the distance matrix, is displayed in Figure 5.

Table 6. Contrastivity (using phone-level entropy loss) for Korean monophthongs (×10)

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		0.751	0.964	0.739	1.106	1.036	0.631	1.498
ɯ			0.385	0.379	0.592	0.517	0.352	0.914
u				0.532	0.795	0.467	0.596	1.192
e					0.510	0.676	**0.279**	0.719
ʌ						0.665	0.506	0.913
o							0.657	1.151
ɛ								0.732
ɑ								

Let us first focus on the strongest contrasts in the vowel system as predicted by our metric. The three pairs i ∼ ɑ ($K = 1.498$), u ∼ ɑ ($K = 1.192$) and o ∼ ɑ ($K = 1.151$) are the most salient contrasts: these pairs of vowels combine a high perceptual distinctness and a high functional load. More interesting for us is the bottom end of the contrastivity

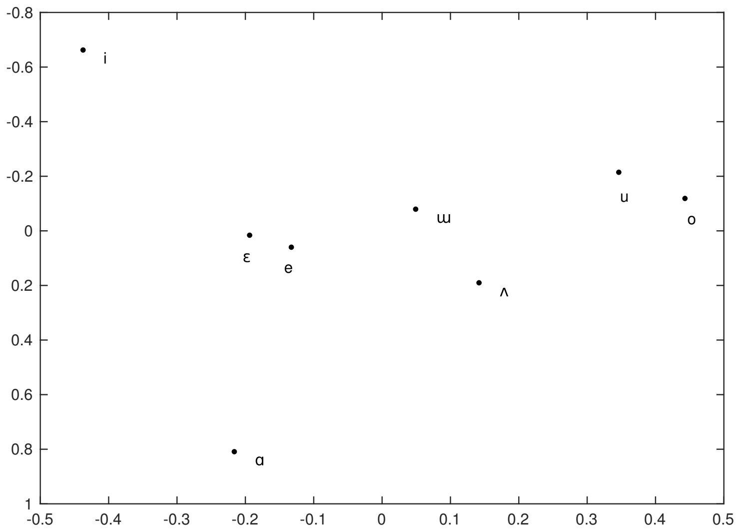

Figure 5. Contrastivity map of Korean monophthongs

scale. As we can see, the metric assigns e ∼ ɛ the lowest contrastivity index ($K = 0.279$); when combined, the relatively low phonetic distance and functional load of these vowels weaken each other and make this pair the least contrastive among all vowel pairs. The two pairs above it are ɯ ∼ ɛ ($K = 0.352$) and ɯ ∼ e ($K = 0.379$). The case of o ∼ u is particularly interesting; although its functional load is about average (just above the median), its contrastivity index is below the second decile (< 0.506), which is due to the perceptual similarity of /o/ and /u/ being the closest among all monophthongs.

Since we considered three alternative instantiations of functional load, we also attempted to measure contrastivity using minimal pair counts and word-level entropy loss instead of phone-level entropy loss. Due to lack of space, we only report on the most salient results. Using minimal pair counts, the three lowest pairs are ɯ ∼ e ($K = 8.10$), e ∼ ɛ ($K = 18.79$) and ɯ ∼ ɛ ($K = 20.56$), and the most contrastive pair is u ∼ a ($k = 118.49$). Using word-level entropy loss, the three lowest pairs are ɯ ∼ e ($K = 0.17$), e ∼ ɛ ($K = 0.25$) and w ∼ u ($K = 0.28$), whereas the most contrastive pair is i ∼ a ($K = 1.28$). It thus appears that all three measures rank e ∼ ɛ lower than does the corresponding measure of functional load, when considered in isolation. Unsurprisingly, the measure which uses phone-level entropy loss ranks e ∼ ɛ lowest since phone-level entropy loss itself ranked this pair lower than did the other measures of functional load.

4.3 Simulations

The acoustic data that we used do not represent the merger at its inception point. In particular, it is quite likely that /ɛ/ was closer to an [æ]-like vowel than it is in Yang's data. In order to put our metric to a test, we carried two simulations, using phone-level entropy loss as a measure of functional load. For both of these simulations, we first created an artificial [æ] vowel. We used F_1 and F_2 values for a synthetic [æ] reported in the literature (Hockett, 1955, 195), with $F_1 = 750$ Hz and $F_2 = 1650$ Hz. Given our data, and assuming that /ɑ/ did not change much, this is probably an unrealistically low vowel, much closer to /ɑ/ (F_1 = 738) than would be expected. Indeed, Schwartz et al. (1997, 266) give the values F_1 = 648 Hz and F_2 = 1712 Hz for their prototypical [æ], which is probably a more realistic value. Nonetheless, we decided to use this extreme point for [æ] to test how far the metric could be stretched.

Since Hockett (1955) did not provide any value for F_3, we fit a simple linear regression model that tried to predict F_3 from F_1, using Yang's values for [i e ɛ a] (measured in bark), the idea being that for unrounded vowels, F_3 is inversely correlated with the degree of lowering of the mandible. Based on this model, we predicted F_3 for [æ] at F_1 = 750 Hz. We thus obtained a reference [æ] with the following values, with F_3 converted back to Hertz: F_1 = 750 Hz, F_2 = 1650 Hz, F_3 = 2565 Hz.

4.3.1 Simulation 1: stepwise lowering of /ɛ/

The first simulation assumed that all vowels but /ɛ/ remained constant. We created nine intermediate points (linearly spaced on the bark scale in the $F_1 \times F_2 \times F_3$ space) between the vowel [ɛ], as reported in Yang (1996), and our reference [æ]. For each of the nine values, we recomputed the contrastivity index of all pairs of vowels, only updating the value for [ɛ]. Figure 6 shows the evolution for the five contrasts which have the lowest scores on our metric: o ∼ u, ɯ ∼ u, ɯ ∼ e, ɯ ∼ ɛ and e ∼ ɛ. The leftmost value is the value from Yang (1996), whereas the rightmost value is our reference [æ]. Note that only pairs that involve the vowel [ɛ] change. We see from that figure that the metric still ranks e ∼ ɛ as the least contrastive pair up to step 5, with values for [ɛ] of F1 = 668 Hz and F_2 = 1744 Hz. Assuming that the vowel system remained constant, we see that even if /ɛ/ was much closer to [æ], as it is often assumed, e ∼ ɛ would still be ranked as the least contrastive pair of monophthongs.

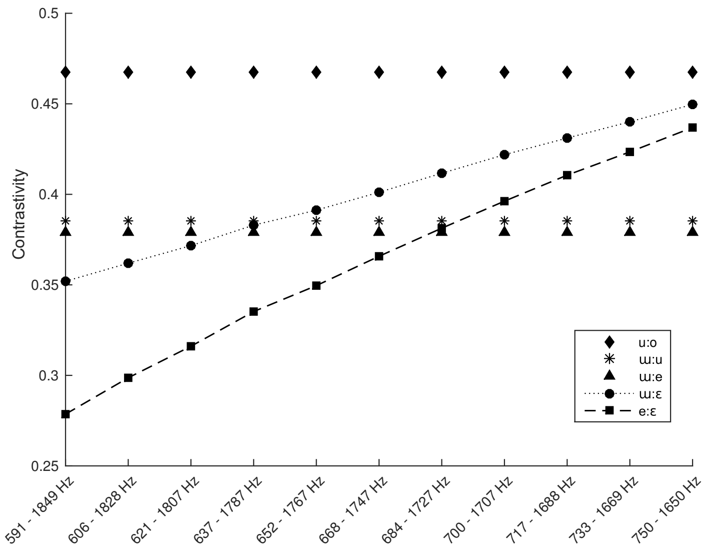

Figure 6. Contrastivity for the 5 weakest contrasts for different values of [ɛ]

4.3.2 Simulation 2: merger by approximation

Since some authors have argued that the e ∼ ɛ merger is a merger by approximation, we attempted to model this scenario as well. Assuming that /e/ and /ɛ/ have been converging towards each other, we first calculated the centroid value that would result from this merger (once again, with all values in bark). This gave a vowel with the following formants: F_1 = 539 Hz, F_2 = 1908 Hz, F_3 = 2620 Hz. This is not very far from the measurements reported by Shin et al. (2013, 103) for Seoul male speakers (F_1 = 490 Hz, F_2 = 1828 Hz), and thus seems like a plausible value. Next, we calculated a 10-step shift from this centroid vowel toward /i/ for /e/, and toward our reference /æ/ for /ɛ/. As in simulation 1, we re-run the calculations one step at a time, this time updating formant values for both /e/ and /ɛ/. The ranking of e ∼ ɛ remained unchanged up to step 4. Reading the formant values for vowels /e/ and /ɛ/ at that point, we obtain the following measurements: F_1 = 456 Hz, F_2 = 2027 Hz, F_3 = 2782 Hz for /e/; F_1 = 619 Hz, F_2 = 1800 Hz, F_3 = 2598 Hz for /ɛ/. This is a little higher/lower than the average values we observe for /e/ (F_1 = 490 Hz) and /ɛ/ (F_1 = 591 Hz) respectively.

5 Summary and discussion

Let us first summarize the findings from this study. Following a number of recent works, we have proposed a quantitative metric of contrastivity, modeled as the geometric mean of distributional equivalence, measured using three different operationalizations of functional load (token-based phone-level and word-level entropy loss, and lemma-based minimal pair counts), and phonetic distance, represented as the weighted Euclidean distance between the perceptual representations of two sounds. The resulting contrastivity index is interpreted as a measure of the strength of a contrast in a given phonological system. The main results are as follows:

- the three measures of functional load ranked e ∼ ɛ as one of the lowest pairs of vowels, but never the lowest
- phone-level entropy loss assigns a lower functional load to e ∼ ɛ than do word-level entropy loss and minimal pair counts
- e ∼ ɛ is one of the closest pairs of vowels in terms of perceptual distance, although o ∼ u is slightly closer
- the contrastivity index systematically ranks e ∼ ɛ lower than does the corresponding measure of functional load, when considered alone.
- e ∼ ɛ has the lowest contrastivity index among all vowel pairs when using phone-level entropy loss.

At the most basic level, these results appear to be consistent with the hypothesis that functional load (in one form or another) and phonetic distance both played a role in the development of the e ∼ ɛ merger. The proposed approach goes one step further, however, since it suggests that these two factors interacted and actually weakened each other, resulting in a weak contrast. From the point of view of the talker-listener interaction, pairs of phonemes such as e ∼ ɛ, which contrast very weakly, are suboptimal in the sense that they carry little information and are relatively difficult to distinguish because of their perceptual similarity. The merger of weakly contrasting segments must in some sense optimize the phonological system, at least locally.

While the three measures of functional load display a similar pattern (they all rank e ∼ ɛ low, and they rank it even lower when combined with phonetic distance), the only measure which ranks e ∼ ɛ as the lowest pair is the contrastivity index based on phone-level entropy loss. Although it might be tempting to consider it as the best metric since it appears to predict the merger most clearly, two remarks are in order. First, it would be a fallacy to infer that only the pair with the lowest index should be expected to merge or that pairs with a higher index should never merge; mergers are complex phenomena which involve both internal and external factors, and contrastivity alone most likely cannot account for the diversity of mergers that occur in languages (Martinet, 1955, 34-35). For instance, another internal factor that appears to play a major role in the preservation and loss of contrasts is feature economy, which is the strong tendency for phonological systems to make maximum use of the distinctive features already present in the system (Martinet, 1955; Clements, 2003). A consequence of this cross-linguistic tendency is that phonemes which are poorly integrated in correlation series are more likely to get lost. There is little doubt that contrastivity and feature economy will sometimes interact, and sometimes compete. Second, we must remain cautious in our interpretation of the results of this study since we only considered a single case of merger. As we have noted several times, Wedel et al. (2013a,b) suggest that minimal pair counts are actually a better predictor than phone-level entropy loss from a cross-linguistic perspective. The discrepancy between our results and theirs could simply mean that Korean e ∼ ɛ is an atypical data point, but it might also stem from different methodologies: we measured phone-level entropy loss using trigrams and word tokens, whereas they used unigrams and word types but considered the effect of frequency separately (see below).

Despite these words of caution and the limited scope of our study, these results are encouraging and open up a number of avenues for further research. A good place to start would be to investigate other mergers found in several Korean dialects. As noted in Section 2.2, in addition to e ∼ ɛ, the Gyeongsang dialect (South East) also merged ɯ ∼ ʌ, a change which was completed before the beginning of the 20th century (Jang and Shin, 2006). Likewise, the northern dialects spoken in the Pyeongan and Hamgyeong regions also underwent a merger of ɯ ∼ u and ʌ ∼ o, respectively. It would be interesting to test whether our approach works equally well in these varieties, and this method should of course be extended to other languages before any strong claim about its validity can be made.

Another area of further research is the measure of phonetic distance. Hall-Lew (2012) points out that using the Euclidean distance of averaged values (as we have done here) hides an important part of the variability within each vowel class. Although it is not always possible to use measurements from groups of vowel tokens to quantify phonetic distance as she recommends, it must be acknowledged that accounting for the dispersion of vowels would certainly be an improvement; at the very least, the standard deviation of the distribution could be included in the calculation of phonetic distance. Taking into account each vowel's variance would certainly yield a more fine-grained understanding of phonetic distance. This could also shed some light on the directionality or type of the merger, which the current metric is unable to do. As we have seen in Section 2.1, the Korean merger of e ∼ ɛ appears to be a merger by approximation, which we transcribe as /e̞/ (Lee and Ramsey 2000, 326, Shin et al. 2013, 99-101), but the current metric is uninformative on this issue; understanding the dispersion of each vowel before the merger takes place may help better predict its outcome. Besides the problem of variance, it must also be noted that the measure of formants, while useful for vowels (at least for oral, modal vowels), does not generalize well to all types of segments, especially obstruents. In addition, dynamics are not accounted for at all, yet these play a critical role in many sound classes, such as long vowels and diphthongs or plosives and affricates (with respect to voice-onset time and burst release).

Finally, the most crucial issue is perhaps the precise nature and number of dimensions that should be included in the metric. For example, although we have only considered the interaction between functional load and phonetic distance, token-based measures of functional load also take into account usage frequency, whereas minimal pair counts do not. The role of frequency in the Korean e ∼ ɛ merger is unclear[17]. As we discussed in Section 3.1, several authors have suggested that low token frequency should favor a merger, everything else being equal elsewhere. The vowels /e/ and /ɛ/ are indeed the least frequent monophthongs in our corpus, and this is true for both types and tokens. We also found very little difference between token-based word-level entropy loss and lemma-based minimal pair counts, which could mean that (token) frequency did not play an important role in this merger. However, in their original study, Wedel et al. (2013b) considered several measures of phoneme frequency and found that the probability of the more frequent member of a pair actually correlated *positively* with the likelihood of a merger, and was a better predictor than were the probability of the less frequent member, the sum of their probabilities or their ratio. This effect was rather subtle, since it was only significant for pairs of phonemes which did have minimal pairs, and it was the same regardless of whether phoneme probability was measured on types or tokens. The role of frequency certainly warrants further investigation but it might well be the case that it should be considered as an independent dimension instead of being subsumed under functional load, as we have done here. Further work might also reveal the need to include additional dimensions, such as morpho-phonological conditioning or paradigmatic effects, or to assign a different weight to each dimension in a way similar to the weighted Euclidean distance of the perceptual model that we used. This could be easily accommodated by using an n-dimension weighted geometric mean instead of the simpler 2-dimension geometric mean used here. This is an issue that we leave open, pending further inquiry.

Clearly, there are a number of ways in which our contrastivity index could be extended and refined, but how precise it can be is ultimately largely conditioned by the type and amount of data available. As annotated and aligned speech corpora become larger and more common, and as automatic phone alignment becomes more reliable, we can expect

[17] In this respect, it was also pointed out to us that /e/ is under-represented compared to /ɛ/ in the Sinitic stratum of the Korean vocabulary. We consulted the dictionary published by the National Institute of the Korean Language (Kim, 2005) and looked up the frequency of /e/, /ɛ/, /je/ and /jɛ/ in the 10,000 most frequent Sino-Korean nouns (i.e. those which were provided in both Hangeul and Hanja). Out of 23,100 vowels, we found 1,451 instances of /ɛ/, 437 instances of /e/ and 274 instances of /je/. (The diphthongs /jɛ/, /we/ and /wɛ/ were absent.) It thus appears that /e/ is about twice as rare as /ɛ/ in this subset of the Sinitic vocabulary. It is unclear to us whether this imbalance could have played a role in the merger.

that studies such as the one we conducted will be possible on a whole new scale across a wide range of languages. This should offer a firmer base for refining the approach we have developed and evaluating it against alternative models of mergers. As modest as it is, we nonetheless hope that this contribution will stimulate others to investigate the interaction between functional factors in the development of mergers more closely.

Acknowledgements

We would like to thank three anonymous reviewers as well as the editors of the journal for their insightful comments and suggestions. This research was supported by the Hankuk University of Foreign Studies Research Fund 2018.

References

Ahn, Sang-cheol. 1998. *An introduction to Korean phonology.* Seoul, Korea: Hansin Munhwasa.

Beddor, Patrice Speeter. 1991. Predicting the Structure of Phonological Systems. *Phonetica* 48 (2-4): 83–107.

Bladon, Anthony, and Gunnar Fant. 1978. A two-formant model and the cardinal vowels. *Speech Transmission Laboratory, Quarterly Progress and Status Report* 19 (1): 1–8.

Blevins, Juliette. 2004. *Evolutionary Phonology. The Emergence of Sound Patterns.* Cambridge: Cambridge University Press.

Bybee, Joan. 2015. Articulatory Processing and Frequency of Use in Sound Change. In *The Oxford Handbook of Historical Phonology*, eds. Patrick Honeybone and Joseph Salmons, 467–484. Oxford: Oxford University Press.

Calamaro, Shira, and Gaja Jarosz. 2015. Learning General Phonological Rules From Distributional Information: A Computational Model. *Cognitive Science* 39 (3): 647–666.

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

Clements, George Nick. 2003. Feature economy in sound systems. *Phonology* 20 (3): 287–333.

de Boer, Bart. 2001. *The Origins of Vowel Systems.* Oxford and New York: Oxford University Press.

Ernestus, Mirjam. 2014. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua* 142: 27–41.

Eychenne, Julien, and Tae-Yeoub Jang. 2015. On the Merger of Korean Mid Front Vowels. *Journal of the Korean Society of Speech Sciences* 7 (2): 119–129.

Fant, C. Gunnar M. 1973. *Speech Sounds and Features*. Cambridge: The MIT Press.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In *Phonetically Based Phonology*, eds. Bruce Hayes, Robert Kirchner, and Donca Steriade, 232–276. Cambridge: Cambridge University Press.

Gahl, Susanne. 2008. "Time" and "Thyme" Are not Homophones: The Effect of Lemma Frequency on Word Durations in Spontaneous Speech. *Language* 84 (3): 474–496.

Garrett, Andrew, and Keith Johnson. 2013. Phonetic bias in sound change. In *Origins of Sound Change: Approaches to Phonologization*, ed. Alan C. L. Yu, 51–97. Oxford: Oxford University Press.

Hall, Kathleen Currie. 2009. A Probabilistic Model of Phonological Relationships from Contrast to Allophony. PhD diss, The Ohio State University.

Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30 (2): 215–275.

Hall-Lew, Lauren. 2012. Improved representation of variance in measures of vowel merger. *Proceedings of Meetings on Acoustics* 9 (1): 060002.

Han, Jeong-Im, and Hyunsook Kang. 2013. Cross-generational Change of /o/ and /u/ in Seoul Korean II: Spectral interactions in normalized vowel space. *Journal of the Korean Society of Speech Sciences* 5 (2): 25–31.

Hermansky, Hynek, and D. J. Broad. 1989. The effective second formant F2' and the vocal tract front-cavity, 480–483. IEEE.

Hockett, Charles Francis. 1955. *A manual of phonology*. Baltimore: Waverly Press.

Hockett, Charles Francis. 1966. The Quantification of Functional Load: A Linguistic Problem., Rand Corporation memorandum RM-5168-PR, Santa Monica, California.

Hong, Yunsook. 1988. *A sociolinguistic study of Seoul Korean*. Seoul, Korea: Hanshin Publishing Co..

Hume, Elizabeth, and Keith Johnson. 2001. A model for the interplay of speech perception and phonology. In *The Role of Speech Perception in Phonology*, eds. Elizabeth Hume and Keith Johnson, 3–26. New York: Academic Press.

Hume, Elizabeth, and Frédéric Mailhot. 2013. The role of entropy and surprisal in phonologization and language change. In *Origins of Sound Change: Approaches to Phonologization*, ed. Alan C. L. Yu, 29–47. Oxford: Oxford University Press.

Hume, Elizabeth, Kathleen Currie Hall, Andrew Wedel, Andrew Ussishkin, Martine Adda-Decker, and Cédric Gendrot. 2013. Anti-markedness patterns in French epenthesis: An information-theoretic approach. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 37 (1): 104–123.

Jang, Hye Jin, and Jiyoung Shin. 2006. An acoustic study on the generational difference of the monophthongs in the Daegu Dialect. *Malsori* 57 (1): 15–30.

Kaplan, Abby. 2011. How much homophony is normal? *Journal of Linguistics* 47 (03): 631–671.

Kim, Hansaem. 2005. *Hyeondae Gugeo Sayong Bindo Josa 2*. Seoul: National Institute of the Korean Language.

Kim, Hansaem. 2006. Korean national corpus in the 21st century Sejong project. In *Proceedings of the 13th National Institute of Japanese Literature (NIJL) International Symposium*, 49–54.

King, Robert D. 1967. Functional Load and Sound Change. *Language* 43 (4): 831–852.

Labov, William. 1994. *Principles of linguistic change. Volume 1: Internal factors*. Oxford and Cambridge: Blackwell.

Lee, Dong-Joo, Jong-Heum Yeon, In-Beom Hwang, and Sang-Goo Lee. 2010. Kkokkoma: gwangyehyeong deiteobeiseureul hwaryonghan sejong malmungchi hwaryong dogu [KKMA: A Tool for Utilizing the Sejong Corpus based on Relational Database]. *Journal of KIISE: Computing Practices and Letters* 16 (11): 1046–1050.

Lee, Iksop, and S. Robert Ramsey. 2000. *The Korean Language*. Albany: State University of New York Press.

Lee, Ki-Moon, and S. Robert Ramsey. 2011. *A History of the Korean Language*. Cambridge and New York: Cambridge University Press.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In *Experimental phonology*, eds. John J. Ohala and Jeri J. Jaeger, 13–44. Orlando: Academic Press.

Martin, Samuel E. 2006. *Reference Grammar of Korean: A Complete Guide to the Grammar and History of the Korean Language*. Tokyo: Tuttle Publishing.

Martinet, André. 1952. Function, structure, and sound change. *Word* 8: 1–32.

Martinet, André. 1955. *Economie des changements phonétiques. Traité de phonologie diachronique*. Paris: Maisonneuve & Larose. New edition 2005.

Martinet, André. 1978. Function, Structure, and Sound Change. In *Readings in Historical Phonology: Chapters in the Theory of Sound Change*, eds. Philip Baldi and Ronald N. Werth, 121–159. University Park and London: The Pennsylvania State University Press.

Moon, Seung-Jae. 2007. A fundamental phonetic investigation of Korean monophthongs. *Journal of the Korean Society of Phonetic Sciences and Speech Technology* 62: 1–17. [In Korean].

Navarro, Tomás. 1967. Nuevos datos sobre el yeísmo en España. *Thesavrvs: Boletín del Instituto Caro y Cuervo* 19 (1): 1–17.

Oh, Yoon Mi, Christophe Coupé, Egidio Marsico, and François Pellegrino. 2015. Bridging phonological system and lexicon: Insights from a corpus study of functional load. *Journal of Phonetics* 53: 153–176.

Ohala, John J. 1981. The listener as a source of sound change. In *Papers from the parasession on language and behavior*, 178–203. Chicago: Chicago Linguistics Society.

Oudeyer, Pierre-Yves. 2005. The self-organization of speech sounds. *Journal of Theoretical Biology* 233 (3): 435–449.

Park, Eunjeong L., and Sungzoon Cho. 2014. KoNLPy: Korean natural language processing in Python. In *Proceedings of the 26th Annual Conference on Human and Cognitive Language Technology*. Chuncheon, Korea.

Peperkamp, Sharon, Rozenn Le Calvez, Jean-Pierre Nadal, and Emmanuel Dupoux. 2006. The acquisition of allophonic rules: statistical learning with linguistic constraints. *Cognition* 101 (3): 31–41.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée, and Christian Abry. 1997. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* 25 (3): 255–286.

Scobbie, James M, and Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In *Contrast in phonology: theory, perception, acquisition*, eds. Peter Avery, B. Elan Dresher, and Keren Rice, 87–114. Berlin and New York: Mouton de Gruyter.

Seong, Cheol-Jae. 2004. An acoustic analysis on the Korean 8 monophthongs with respect to the acoustic variables on the F1/F2 vowel space. *Journal of the Acoustical Society of Korea* 23 (6): 454–461. [In Korean].

Shannon, Claude Elwood. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423.

Shin, Jiyoung, Jieun Kiaer, and Jaeeun Cha. 2013. *The sounds of Korean*. New York: Cambridge University Press.

Silverman, Daniel. 2010. Neutralization and anti-homophony in Korean. *Journal of Linguistics* 46 (02): 453–482.

Sohn, Ho-Min. 1999. *The Korean Language*. Cambridge: Cambridge University Press.

Sohn, Ho-min. 2012. Middle Korean. In *The Languages of Japan and Korea*, ed. Nicolas Tranter, 73–122. Abingdon and New York: Routledge.

Steriade, Donca. 2007. Contrast. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, 139–157. Cambridge: Cambridge University Press.

Surendran, Dinoj, and Gina-Anne Levow. 2004. The functional load of tone in Mandarin is as high as that of vowels. In *Proceedings of Speech Prosody*, 99–102. Nara, Japan.

Surendran, Dinoj, and Partha Niyogi. 2003. Measuring the usefulness (functional load) of phonological contrasts, Technical report, Technical Report TR-2003.

Surendran, Dinoj, and Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Amsterdam studies in the theory and history of linguistic science. Series 4* 279: 43.

Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* 88 (1): 97–100.

Umeda, Hiroyuki. 1995. Age differentiation of the vowel system in the Seoul Korean – Acoustic measurements. *Journal of Asian and African Studies* 48-49: 443–453.

Wang, William S-Y. 1967. The measurement of functional load. *Phonetica* 16 (1): 36–54.

Wedel, Andrew, Scott Jackson, and Abby Kaplan. 2013a. Functional Load and the Lexicon: Evidence that Syntactic Category and Frequency Relationships in Minimal Lemma Pairs Predict the Loss of Phoneme contrasts in Language Change. *Language and Speech* 56 (3): 395–417.

Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013b. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128 (2): 179–186.

Yang, Byunggon. 1992. An acoustical study of Korean monophthongs produced by male and female speakers. *The Journal of the Acoustical Society of America* 91 (4): 2280–2283.

Yang, Byunggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24 (2): 245–261.

Yeon, Jaehoon. 2012. Korean dialects: a general survey. In *The languages of Japan and Korea*, ed. Nicolas Tranter, 168–185. Abingdon and New York: Routledge.