Language and Linguistics Compass 2/1 (2008): 103–118, 10.1111/j.1749-818x.2007.00040.x

Interlanguage Segmental Mapping as Evidence for the Nature of Lexical Representation

Yoonjung Kang*
University of Toronto Scarborough

Abstract

The traditional view of phonological representation assumes that lexical representation is economical and free of redundancy and that a phoneme is represented as a combination of contrastive phonetic features. The traditional view, however, is challenged by the recent developments in phonological theories. In Optimality Theory, the distinction between contrastive vs. non-contrastive aspects of pronunciation is made not by the lexical representation but by the constraint ranking and the commonly adopted principle of Lexicon Optimization requires lexical representation to be as close to the surface representation as possible. An exemplar model of phonology also assumes that predictable and redundant phonetic properties are also specified in the lexical representation and that each contextual variant of a phoneme forms a separate phonetic category. The current paper reviews recent studies on segmental mapping in interlanguage phonology, examines the nature of the L1 phonological representation that the mapping crucially refers to, and discusses how the interlanguage data bear on the debate on the nature of the lexical representation.

1. Views on Lexical Representation and the Status of the Phoneme

The traditional view of lexical representation in generative phonology assumes that human memory capacity is limited and that lexical representation should be economical (Halle and Clements 1983). The lexical representation should only contain unpredictable information and all predictable details should be derived by rules or constraints. Storing any predictable details of surface pronunciation is equated with incorrectly declaring them unpredictable. As phonetic properties of a segment that are not contrastive are excluded from lexical representation, a given phoneme has a consistent representation across contexts regardless of their differing allophonic realizations at the surface. For example, the coronal stops in [t^{h}]*op* and *s*[t]*op* are represented uniformly in the lexical representation, as /t/, because aspiration is not contrastive in English stops. Different models of feature underspecification limit the amount of lexical specification to a different degree (Steriade 1987; Archangeli 1988; Avery and Rice 1989; Dresher 2003)

but they all agree in assuming that phonetic features that do not serve a phonemic contrast of the language are left out of lexical representation. Economy can also be sought for the language as a whole by limiting the number of features used to distinguish a given set of phonemes (Clements 2001).

Such abstract and minimal lexical representation of speech sound has been proposed as a solution to the problem of variation in speech perception, namely, a given phonological category is realized differently depending on the speaker and the phonological context, but speech perception, nevertheless, is carried out seamlessly. Lahiri and Marslen-Wilson (1991) propose that recognition of phonological categories in the face of such variation is possible, because lexical representation is sufficiently abstract and devoid of phonetic details to be compatible with all surface variants of a given category.

Recently, however, there have been serious challenges to this traditional view of lexical representation. First of all, human memory capacity seems surprisingly vast and there is no a priori basis to assume that lexical storage should be maximally economical (see references cited in Johnson 1997). Also questionable is the related implicit assumption that deriving predictable aspects of surface pronunciation by online computation is more efficient and less costly than storing these details as such and accessing them directly (Nooteboom et al. 2002). Also, in the traditional model, a distinction is made between contrastive and redundant aspects of phonetic details of a phonemic contrast, and only the contrastive phonetic feature is specified lexically and plays a role in speech perception (Lahiri and Marslen-Wilson 1991). However, there is evidence that supposedly redundant features play a crucial role in speech perception and, moreover, supposedly contrastive features may not be consistently available in all phonological contexts – a problem referred to as THE LACK OF INVARIANCE PROBLEM (Miller 1994; de Jong 1995; Ohala and Ohala 1995).

For example, in English stressed monosyllables, a vowel is longer before a voiced consonant than before a voiceless consonant, and the vowel length difference constitutes an important perceptual cue in distinguishing word-final voicing contrasts in obstruents (Port and Dalby 1982). At the same time, this is precisely the context where the voiced obstruents tend to devoice, creating a surface contrast like [fʌs] 'fuss' vs. [fʌːz̥] 'fuzz'. As a result, the so-called voicing contrast is signaled, not by voicing of the consonants per se, but by the length of the preceding vowel. In situations like this, where more than one phonetic parameter is yoked together to signal a single contrast, choosing one phonetic parameter as contrastive and others as redundant is quite arbitrary and does not reflect the reality of speech perception (see more examples mentioned in de Jong 1995).

Alternatives to the traditional view of lexical representation assume a rich lexical representation that is full of redundant details. Bybee (2001) proposes a usage-based model of phonology that denies the very premise of the traditional view that predictable and productive properties of

pronunciation should not be specified in the lexical representation. Under this model, allophonic details of surface pronunciation are stored as such and their distributional regularities emerge as generalizations on the stored representations, not as rules or constraints that exist independent of stored items. An exemplar model of speech perception and production similarly assumes that every experienced token of speech sound leaves a trace in the lexicon (Goldinger 1997; Johnson 1997; Pierrehumbert 2001, 2003). Under this view, phonetic categories are formed based on distributional patterns of experienced exemplars over phonetic space (Maye and Gerken 2000; Maye et al. 2002) and speech perception circumvents the problem of variability by storing all surface variants of a lexical item and directly accessing them as such. It is crucial to note that the LACK OF INVARIANCE PROBLEM, that is, the phonetic properties that define a given contrast may be different from context to context, is handled by positing phonetic categories that are context dependent (Miller 1994; Maye 2000; Bybee 2001; Pierrehumbert 2003).

For example, word-final instances of /z/ will overlap significantly with word-initial instances of /s/ in acoustic properties due to final devoicing of /z/ and combined distribution of the two fricatives over all contexts will fail to show discrimination of /z/ and /s/. The model of speech perception that assumes context-dependent phonetic categories succeeds in perceiving context-dependent variants of a phoneme by giving up (and in fact denying) an inherent connection between surface variants of a phoneme in speech perception. In other words, in this model, there are only phonetic categories (equivalent to allophones) but there is no formal representation for PHONEMES as we know them. Pierrehumbert (2003) states that 'the position-dependent distinctions appear to be tractable if [variants in one context] are inactive when [variants in other contexts] are being perceived.' In this model, various contextual allophones may be associated 'through an analysis of homologies' due to their phonetic similarity, but importantly, speech perception accesses the context-specific phonetic categories, not phoneme-like categories that are defined across all different contexts.

Similarly, in the standard model of Optimality Theory (OT), lexical representation is not limited to unpredictable or contrastive information. In OT, the distinction between predictable and unpredictable aspects of pronunciation is made, not by whether they are specified in the lexical representation or not, but by constraint ranking – predictable aspects of pronunciation are those for which the relevant markedness constraints dominate faithfulness constraints (Kirchner 1997; Kager 1999; Pater 2003). The commonly adopted principle of Lexicon Optimization in OT dictates that lexical representation should be as close to surface realization as possible (Prince and Smolensky 1993). Under this view, lexical representation of *top* and *stop* are identical to their surface representation – [t^hɑp] and [stɑp].

To summarize, the traditional model posits one consistent lexical representation for a given phoneme that is defined by a set of contrastive

phonetic features. On the other hand, the exemplar model of lexical representation posits separate phonetic categories for context-specific instances of phonemes. Such phonetic categories are distinguished by multiple phonetic parameters that can differ from context to context, even for a given phonemic contrast.

In the rest of this article, I will examine how researches on interlanguage segmental perception and production can shed light on this issue. One's perception of foreign language sounds is filtered by the system of one's native language and therefore, by examining the way a speaker perceives or fails to perceive foreign contrasts, we can find evidence for the nature of native phonological representation that acts as a sieve for the foreign contrasts. Also, by examining how segmental equivalences between the foreign and the native language are established in second language production and loanword adaptation, we can gather evidence about the nature of native phonological representation by which the foreign inputs are categorized.

2. Interlanguage Mapping and the Measure of Similarity

One's perception of foreign sounds is affected by the sound structure of one's native language. Listeners have a tendency to perceive acoustic signals as sound sequences that conform to the sound structure of one's own language. For example, Japanese does not allow a non-prevocalic obstruent and indeed Japanese speakers tend to hear an extra vowel in words like *ebzo*, which violates the language's phonotactic constraints. When presented with a series of stimuli ranging from *ebzo* to *eb<u>u</u>zo* with a vocalic portion of varying length, the majority of Japanese speakers perceive *ebzo* as *eb<u>u</u>zo* EVEN WHEN THERE IS NO VOCALIC SIGNAL IN THE STIMULI while the majority response of French speakers, for whom *ebzo* is a licit sequence, shifted from *ebzo* (no vowel) to *ebuzo* (vowel) only when the vocalic portion is lengthened to around 50 ms (Dupoux et al. 1999).

The influence of native phonology on the perception of a foreign language is also ubiquitous at the level of single segment mapping (Hancin-Bhatt 1994; Best 1995; Flege 1995; Brown 2000). When foreign sounds are close enough to certain sounds of one's native language, equivalence is established between them, affecting the perception and production of the foreign (and sometimes native) sounds. For example, Japanese speakers have difficulty distinguishing /l/ and /r/ in English, as these two sounds are equated with the single liquid phoneme in Japanese. Also, English speakers equate the voiceless stops of Spanish with that of English, and produce the Spanish equivalent with voice onset time (VOT) values close to that of English. Such cases of interlanguage equivalence are abundantly found in second language phonology and also in loanword adaptation.

One of the interesting research questions that arise here is specifically how such an equivalence relationship is established. In other words, what is the nature of the phonological representation through which segmental mapping

is carried out? Much has been said about the nature of the L2 (or input language) representation that guides the interlanguage mapping, especially in the area of loanword phonology. Some claim that the foreign input to the loanword adaptation is phonetic in nature with all non-contrastive details specified (Silverman 1992; Peperkamp and Dupoux 2003). Others claim that the foreign input is phonological and only includes contrastive aspects of input segments (Oh 1996; LaCharite and Paradis 2005). Yet, others take a more measured view that acknowledges the role of both phonetic and phonological aspects of the input language (Kang 2003; Heffernan 2005).

This article will concentrate on the other side of the issue, that is, the nature of the L1 (or borrowing language) representation that acts as a filter of foreign input in interlanguage segmental mapping. Specifically, are only contrastive features of the L1 relevant in cross-language mapping (Section 2.1)? Does a phonemic category of L1 receive a consistent treatment throughout the context, despite their surface variations, or does the mapping tailor to phonetic details of allophonic variants of the L1 phoneme (Section 2.2)? I now consider these two questions in turn.

2.1. ROLE OF CONTRASTIVE AND NON-CONTRASTIVE FEATURES IN INTERLANGUAGE PHONOLOGY

As mentioned above, in L1 speech perception, evidence indicates that phonetic details that are not considered to be underlyingly contrastive, but are nevertheless robustly present in the surface contrast, can play a significant role in speech perception. But do such non-contrastive phonetic details of L1 play a role in filtering the foreign sounds and mapping them onto native categories?[1] Quite a few studies of interlanguage segmental mapping start with the assumption that contrastive features of the L1 play a primary role in determining the pattern of interlanguage mapping, although the specific model of feature specification may vary. Two such explicit models of interlanguage segmental mapping are those of Hancin-Bhatt (1994) and Brown (2000); however, neither of these models is fully empirically supported.

Hancin-Bhatt (1994) proposes that interlanguage segmental mapping is driven by contrastive features of the L1. When different contrastive features of the L1 make conflicting demands, the feature that has a higher functional load – as defined by the proportion of phoneme inventory that relies on the feature for contrast – determines the mapping. The model is applied to differential substitution of English interdental fricatives /θ ð/ in German, Hindi, Japanese, and Turkish, but the predictions of the model are only partially borne out.

Brown (2000) assumes that a new L2 phoneme contrast that involves a feature that is already available in the L1 is easily acquired while a new L2 contrast that requires the introduction of a new feature is more difficult to learn. For example, Mandarin Chinese and Japanese both lack phonemic contrast of /l/ and /r/, but Chinese speakers are better at discriminating /l/

and /r/ in English than Japanese speakers. Brown (2000) assumes that English /l/ is distinguished from /r/ by the presence of [Coronal] node for /l/. She goes on to propose that Chinese speakers are good at /l/ vs. /r/ contrast, because [Coronal] is already active in Chinese to house a subcoronal place distinction between alveolar and retroflex fricatives (e.g., /s/ vs. /ʂ/). On the other hand, Japanese lacks a subcoronal place contrast and therefore the [Coronal] node is unavailable. Brown's (2000) analysis is, however, problematic in that it relies on a non-standard assumption that /l/ and /r/ are distinguished by the feature [Coronal] rather than [lateral], which would have been a novel contrastive feature to both Japanese and Mandarin Chinese. Also, Flege and Port (1981) found native Arabic speakers fail to readily extend the voicing feature available from their L1 /t d/ contrast to a novel /p b/ contrast in L2, casting doubt on the abstract feature-based account along the line of Brown (2000).[2]

Note, however, that even if there are cases where interlanguage segmental mapping can be shown to be carried out based on contrastive features, it does not necessarily rule out the possibility that other phonetic details are also accessed; it is quite likely that what are considered contrastive features are also the most robust phonetic cues at the surface and, therefore, are relied upon in perception more than other phonetic cues. In other words, it could be the case that all phonetic details are accessed, but when different cues are in conflict, the most robust cues take precedence over others. Therefore, the evidence for the primacy of contrastive features should come from cases where a given contrast is signaled by more than one phonetic cue of comparable robustness but in interlanguage mapping, the contrastive one takes precedence over others. I will discuss two such potential cases, namely, Ito et al. (2006) and Herd (2005).

Ito et al.'s (2006) study of Japanese loanwords in Korean shows that the Japanese high back vowel /u/ is adapted to the Korean rounded high back vowel /u/, maintaining the labial gesture of the Japanese vowel, despite the fact that the second formant (F2) value of this Japanese vowel is more closely matched by Korean /ɨ/ than /u/. One of the possible explanations suggested by Ito et al. (2006) is that the phonologically contrastive feature that distinguishes /ɨ/ and /u/ in Korean is rounding, not the tongue backness, which is manifested in the F2 value acoustically. In other words, the multiple phonetic features in the foreign input make conflicting choices in interlanguage mapping and a priority is given to the contrastive feature (i.e., rounding) over the redundant phonetic detail (i.e., tongue backness).

Herd (2005) provides another example where the contrastive feature has a final say in segmental mapping when the phonetic similarity is indeterminate. Herd (2005) examines the differential adaptation of English consonants to various Polynesian languages, which are well-known for their simple consonantal inventory, and proposes that the mapping should be sensitive to the minimal contrastive feature specification (cf. Clements 2001; Dresher 2003). Hawaiian and New Zealand Maori both lack sibilants /s z ʃ ʒ/ and

these English sibilants are adapted to /h/ in New Zealand Maori, but the same English sibilants map to /k/ in Hawaiian, despite the fact that both /k/ and /h/ are available in both languages.[3] Herd (2005) proposes that the crucial difference between Hawaiian and New Zealand Maori is that, in Hawaiian, the non-labial obstruent inventory is /h ʔ k/ while in Maori it is /h t k/. In other words, in Hawaiian, /h/ contrasts with another laryngeal consonant /ʔ/ and this contrast requires /h/ to be specified for the laryngeal feature – Glottal Width and [spread] (cf. Avery and Idsardi 2001). This laryngeal specification creates a mismatch with English sibilants and the English sibilants are mapped to the other non-labial obstruent, /k/. In New Zealand Maori, on the other hand, /h/ does not contrast with another laryngeal consonant and, therefore, /h/ can be underspecified for the laryngeal feature causing no mismatch with English sibilants. /k/ in New Zealand Maori, on the other hand, is specified as [Dorsal], as it contrasts with /t/ minimally in its place of articulation, and /k/ cannot be matched with English sibilants, which are not dorsal. It does not seem likely that the acoustic properties of /h/ and /k/ of the two languages are significantly different enough such that English sibilants are more closely matched by [h] in New Zealand Maori but by [k] in Hawaiian acoustically.

To summarize, Ito et al. (2006) and Herd's (2005) results provide strong evidence for an abstract featural representation that cannot be deduced from surface phonetics alone by showing that contrastive features determine the choice in interlanguage mapping when phonetic similarity fails to provide a clear choice.

On the other hand, there are a few studies that emphasize the role of non-contrastive phonetic features of the L1 in interlanguage mapping. The English interdental fricative /θ/ is adapted differentially to European French and Quebec French, mapping to /s/ in European French and to /t/ in Quebec French. According to Brannen (2002), the differential adaptation cannot be due to a difference in the contrastive feature specification of the two dialects, as they have an identical set of phonemes. Rather, the crucial difference between the two dialects is that in European French, the coronal fricative /s/ has a dental place of articulation while in Quebec French, /s/ is alveolar, a phonetic detail that is not contrastive in each of the dialects. Based on this differential substitution, Brannen (2002) proposes that interlanguage mapping should crucially refer to phonetic details that are not contrastive in native language.

Curtin et al. (1998) found that English speakers were better at distinguishing the voicing contrast of Thai stops ([b] vs. [p]) than the aspiration contrast ([p] vs. [p^h]) in vocabulary learning tasks, despite the fact that surface VOT contrast of English stops is more closely matched by the aspiration contrast of Thai. Curtin et al. (1998) interpret this result as evidence that the English learners are utilizing the underlying contrast available in English, namely, [voice], rather than the surface contrast in aspiration. However, Pater (2003) failed to replicate Curtin et al.'s (1998) result and rather found a better

discrimination of the Thai aspiration contrast than the Thai voicing contrast by English speakers in various experimental tasks. Such results indicate that the supposedly non-contrastive feature, that is, aspiration of English stops, can play a salient role in interlanguage perception.[4]

The emerging picture is that even the phonetic details that are considered non-contrastive in the traditional view may be accessible in cross-language segmental mapping, especially when they are salient phonetically. At the same time, evidence from Herd (2005) and Ito et al. (2006) suggests that the contrastive status of features that cannot necessarily be deduced from phonetic saliency at the surface plays a role in interlanguage mapping, especially when the phonetic similarity does not provide a clear choice.

2.2. CONTEXTUAL DIFFERENCES: PHONEMES OR ALLOPHONES

Another difference between the traditional view of redundancy-free lexical representation and the surface-true view of lexical representation is that, in the former, a phoneme is given a consistent lexical representation. In the latter view, on the other hand, each contextual variant of a phoneme will form a separate phonetic category and be accessed as such and no formal status is given for the phoneme as we know it.

There is evidence that interlanguage segmental mapping is sensitive to contextual variations of a phoneme in the L1. Korean has a laryngeal contrast between the so-called lax and tense stops (/p t k/ vs. /p* t* k*/). Phonetically, the contrast is signaled by a combination of features: quality and pitch of the following vowel, duration of the stop closure, and VOT (see Cho et al. 2002 for a comprehensive review). But, these cues are not equally significant in all contexts. Specifically, in word-medial position, the tense stops have a closure duration that is over two times longer than that for lax stops, but the difference in closure duration is significantly reduced in word-initial position (Han 1996). Accordingly, duration cue plays a significant role in perception of the contrast only in word-medial position but not in word-initial position, where the closure duration cue may not be reliably available, if at all (Han 1996; Kang and Kang 2006). The evidence that such contextual variation of the L1 category may affect intersegmental mapping comes from loanword adaptation of Japanese voiceless stops to Korean (Ito et al. 2006). Word-initial /k/ of Japanese is adapted as lax /k/ of Korean (e.g., J. /katuo/ > K. /kas*ɨo/ 'bonito'), but Japanese /k/ maps to tense /k*/ of Korean word medially (e.g., J. /mikaɴ/ > K. /mik*aŋ/ 'tangerine'). In word-initial position, the closure duration cue does not play a role in the distinction between lax and tense stops and the mapping relies on other cues, namely, quality and pitch of the following vowel, according to which Japanese voiceless /k/ is more closely matched by Korean lax /k/ than tense /k*/. In word-medial position, on the other hand, the long closure duration of Japanese voiceless stops is perceptually relevant and Japanese /k/ is perceived as closer to Korean tense /k*/ than lax /k/.

Such evidence suggests that interlanguage mapping should be able to access context-dependent subphonemic details of L1 sounds.[5] However, a theory that actively denies a unity of allophones of a single phoneme seems too strict to be viable. Flege (1995) puts forward a very specific hypothesis that '[s]ounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level.' As Flege (1995) discusses, a strict interpretation of this hypothesis predicts that 'speakers of L1 without word-final stops [such as Italian] will not relate English word-final stops perceptually to word-medial or word-initial stops in their L1.' Under this view, the word-final stops of English (as in *tag* and *tack*) are not equated with any existing category of L1 and as a result, it is predicted that there is no interference from native categories and that the word-final stops should be learned perfectly given enough exposure to the English input. However, Flege et al. (1995) found that many highly experienced Italian learners of English who started learning English after the age of 12 failed to produce thoroughly English-like /k/ and /g/ contrast word finally. Acoustic measurements show that these speakers produced a smaller vowel length difference in *tag* vs. *tack* than native English speakers and also produced a slightly larger stop closure duration difference than native English speakers, possibly transferring the pattern found in word-medial /k/ vs. /g/ contrast of their L1. In other words, these Italian speakers are equating the word-final /k/ and /g/ of their L2 with the word-medial /k/ and /g/ of their L1, contrary to the hypothesis that position-sensitive allophones are the primary units of segmental equivalence. A theory that denies cross-context consistency of a phoneme-like unit runs into a problem with such cross-context transfer effect.

Also potentially problematic for a theory that does not recognize a phonemic level of representation is the observation that one of the hardest contrasts to acquire in L2 phonology is the L2 phonemic contrast that splits the allophones of a single L1 phoneme (Hammerly 1982; Eckman et al. 2003). For example, in Spanish, [d] and [ð] are allophones of the same phoneme with [ð] occurring after a continuant sound and [d] occurring elsewhere, and Spanish speakers have difficulty producing and perceiving the /d/ vs. /ð/ contrast of English (Hammerly 1982; Sánchez 2003). If the L2 input is filtered by the phonemic representation of L1, we expect both /d/ and /ð/ in English to map to the single phoneme category in Spanish, and the Spanish speakers' difficulty with this contrast is expected. Under an exemplar model of lexical representation or Flege's (1995) allophone-based mapping hypothesis, [d] and [ð] in Spanish will form separate position-sensitive phonetic categories that do not share a common abstract representation as a phoneme. Therefore, a /d/ vs. /ð/ contrast is not expected to be any more difficult than other novel contrast of comparable phonetic similarity. As /d/ vs. /ð/ contrast is indeed difficult for Spanish speakers to learn, the phonemic category model appears to be more successful than the phonetic category model. However, the assertion that the L2 contrast that

is allophonic in L1 is harder to acquire than other novel L2 contrasts needs to be examined more carefully. Some recent perception studies suggest that the discriminability of novel foreign contrasts depends not so much on the phonemic status of the contrast in L1 per se, but whether they contrast in the surface in L1 and how phonetically distinctive they are.

Hall et al. (2004) examined the discrimination of [d], [ð], and [ɾ] in intervocalic contexts by Spanish speakers. In Spanish, [d] and [ð] are allophones of a single phoneme category. [ɾ] is phonetically quite similar to [d] and [ð], but phonologically [ɾ] belongs to a phoneme category separate from [d] and [ð]. If the phonemic status of these segments determines the relative difficulty in discrimination, we expect Spanish speakers to have difficulty with [d]–[ð] distinction but not with [d]–[ɾ] or [ð]–[ɾ] distinction, as the latter pairs involve distinction across different phonemic categories.

The result, however, shows that while Spanish speakers are good at distinguishing [ɾ] from [ð], as expected under the phonemic analysis, Spanish speakers have almost as much difficulty with the [d]–[ɾ] distinction as with the [d]–[ð] distinction, contrary to the predictions of phonemic analysis. Hall et al. (2004) propose that the crucial factor is not whether the pair of segments contrasts phonemically, but whether it contrasts in the particular context at the surface. Of the three segments, [ð] and [ɾ] occur intervocalically, but [d] does not, and therefore, only the [ð]–[ɾ] pair contrast in intervocalic position in Spanish. [d]–[ɾ] and [d]–[ð], on the other hand, are discriminated equally poorly, because they do not contrast in intervocalic context. This is the case regardless of whether the contrasting segments belong to different phonemic categories, as in [d]–[ɾ], or not, as in [d]–[ð].

Sung (2003) studied American English speakers' perception of [d]–[ɾ] and [t^h]–[ɾ] contrasts in intervocalic position. Despite the fact that both pairs are allophones of a single phoneme, English speakers were good at discriminating [t^h] from [ɾ], performing at a level comparable to a native phonemic contrast ([l] ~ [ɾ]), but not good at discriminating [d] from [ɾ]. One plausible explanation for this difference is that [t^h] and [ɾ] are phonetically sufficiently distinctive, while [d] and [ɾ] are not. In other words, even allophonic contrasts may be easily discriminated depending on the degree of phonetic distinctiveness.

Similarly, Peperkamp et al.'s (2003) study of French speakers' perception of [χ] and [ʁ] shows that the discriminability of novel contrasts can differ depending on the phonetic distinctiveness, even for the same allophonic contrasts. In French, [χ] and [ʁ] are allophones of a single phoneme with [χ] occurring next to a voiceless obstruent and [ʁ] occurring elsewhere. Peperkamp et al. (2003) find that French speakers were good at discriminating between [χ] and [ʁ] when the sounds were presented in V_# context, as in [aχ] vs. [aʁ]. In this context, their discrimination for the allophonic distinction of [χ] and [ʁ] was comparable to their discrimination of phonemic distinction of [m] and [n], contrary to the prediction of phoneme-based mapping. But, when the stimuli were embedded in a

preconsonantal context as in [aχCV] vs. [aʁCV], the discrimination of the allophonic distinction fell dramatically, regardless of whether the following consonant matched the voicing of the segment under focus or not. In other words, the phonemic status of contrasting sounds alone does not determine the discriminability of pairs of sounds; allophonic contrasts may still be perceived accurately if presented in contexts where the acoustic cues are salient enough to allow the formation of separate phonetic categories. Therefore, the perception of novel contrasts appears to access the level of representation that is more concrete than the phonemic level.

Also, there is evidence that even two segments that belong to different L1 phoneme categories are not properly distinguished if they do not have surface contrasts in L1 and they are not phonetically distinctive enough. Pegg and Werker (1997) examined the discrimination of word-initial [d̥a] and [t̄a] by adult native speakers of English and 10–12-month-old infants. In English, the devoiced [d̥] occurs in word-initial position and unaspirated [t̄] occurs following [s], as in *stop*. The two phones belong to two different phonemic categories but they do not contrast in word-initial position. The result shows that while the English speakers could discriminate between [d̥a] and [t̄a] better than chance, the discrimination was worse than between [d̥a] and [tʰa], which do contrast word initially in English. This is also in line with Hall et al.'s (2004) observation that it is the context-specific surface contrast that enables speakers' discrimination, not the phonemic contrast per se.[6]

What such results suggest is that the supposed phonemic effect in L2 perception (Hammerly 1982) may be epiphenomenal; allophones of single phonemes are hard to distinguish because they are usually phonetically similar and lack surface contrast, not because they are variants of a single phoneme and share an identical lexical representation. In other words, the cases reviewed in this section support or are at least compatible with the view that interlanguage discrimination and mapping refers to context-specific phonetic categories that may be confusable to a varying degree depending on their inherent acoustic distinctiveness, not necessarily to an abstract phonemic category per se.

3. Conclusion

In this article, I reviewed recent studies on interlanguage segmental mapping and examined the nature of phonological representation that the mapping crucially refers to. Two specific questions are addressed: (i) Do non-contrastive phonetic details of the L1 contrast play a role? and (ii) Are the basic segmental units of interlanguage mapping phonemic or allophonic?

Concerning the first question, there is some evidence that interlanguage mapping DOES refer to non-contrastive phonetic details of L1 contrast (Brannen 2002; Pater 2003), and at the same time there is also evidence that contrastive features play a crucial role when phonetic similarity alone

does not make a clear choice (Herd 2005; Ito et al. 2006). In other words, the studies reviewed here suggest that the phonological representation that the interlanguage mapping refers to should contain phonetic details that are considered non-contrastive in L1 and at the same time should somehow acknowledge the privileged status of contrastive features.

An important question brought up by an anonymous reviewer is what happens when the two types of information actually conflict? In other words, can a salient non-contrastive phonetic feature override a contrastive feature that is less saliently represented in its phonetic realization? This seems to be exactly the kind of research question that needs to be pursued to refine our model of lexical representation. My hunch, however, is that such cases will be hard to come by for two reasons. The first reason is the learnability consideration. If a given contrast is signaled by two phonetic features, one more salient than the other, other things being equal, the learners will likely latch onto the more salient phonetic feature as the primary contrast opting for a transparent phonetics–phonology mapping. As a result, the type of cases we are seeking, where the contrastive feature is phonetically less well-represented, are unlikely to persist. The second related reason is that from an analytical point of view, if we find a case where supposedly non-contrastive but salient feature overrides a contrastive but less salient feature, chances are that there is an alternative phonological analysis that treats the salient feature as contrastive. This is exactly the situation we face in 'aspiration' vs. 'voicing' analysis of laryngeal contrast in English stops (see note 4).

This being said, a more careful and diligent search may yield cases where the phonology of the language points to one feature as contrastive while the phonetic salience points to another feature as more prominent and we await further research to find out how such conflicts are resolved.

In regard to the second question on the role of phonemic categories in interlanguage mapping, there is evidence that contextual cue variation of L1 phonemic contrast does play a role (cf. adaptation of Japanese stops in Korean: Ito et al. 2006). Also, the claim that novel L2 contrasts that split the allophones of a single L1 phoneme is harder to acquire than other novel L2 contrasts needs to be critically examined. It is suggested that the reason that splitting allophones of a single L1 phoneme into new contrasts tends to be difficult is that the allophones are very similar acoustically and lack surface contrast by definition. In other words, the apparent phonemic category effect may just be epiphenomenal, and not due to an active role of phonemic category in interlanguage perception. These findings in interlanguage phonology do not necessarily mean that there is no abstract phoneme, as an anonymous reviewer points out. It is possible that there are abstract phonemes but interlanguage perception and production simply refers to a representation less abstract and closer to a surface representation. In any case, the cases reviewed here fail to provide positive evidence for abstract phonemes and if anything points to a richer conception of

phonological representation (not an economical one devoid of redundancy), where non-contrastive phonetic details play an active role in interlanguage segmental mapping.

Acknowledgements

I would like to thank two anonymous reviewers and the editor for helpful comments and suggestions on an earlier version of this article.

Short Biography

Yoonjung Kang's areas of specialization are phonology and its interface with phonetics and morphology, with a special focus on Korean. Her current research focuses on the sound patterns of English and Japanese loanwords in Korean; specifically she is investigating the role of detailed phonetics in explaining the patterns of transformation the borrowed words undergo and she is also examining the diachronic changes in English loanwords in Korean over the last 100 years. She is also studying the dialectal variations in nominal and verbal inflections in Korean, investigating the theoretical implication of the dialectal data on the question of how inflectional paradigms are acquired and represented in the speakers' mental grammar. Kang is an assistant professor in linguistics in the Department of Humanities at the University of Toronto Scarborough and holds a graduate appointment in the Department of Linguistics at the University of Toronto. Previously, she has taught at Yale University and Stony Brook University. She holds a BA in English language and literature from Seoul National University and a PhD in Linguistics from MIT.

Notes

⋆ Correspondence address: Yoonjung Kang, Department of Humanities, University of Toronto at Scarborough, 1265 Military Trail, H521 Scarborough, Ontario, Canada M1C 1A4. Email: kang@utsc.utoronto.ca.

[1] Evidence shows that L2 learners are capable of learning contrasts involving acoustic dimensions that are not contrastive in their native phonology (Flege 1995). For example, Escudero and Boersma (2004) show that beginning Spanish learners indeed distinguish English /i/ and /ɪ/ based on spectral difference only, which is the contrastive phonetic dimension for their L1 vowel contrast. But, advanced learners show weighting of spectral and durational cues that closely resemble those found in the ambient English dialect indicating that L2 learners are capable of acquiring contrasts that employ an acoustic dimension that is not contrastive in their L1. So, the question addressed here is not whether the L2 learners are able to perceive the phonetic details of L2 sounds that are not contrastive in their L1, but rather to what extent such details are relevant in establishing L1–L2 segmental equivalence.

[2] Similarly, in an artificial learning experiment, Maye and Gerken (2001) found that adult speakers fail to extend a VOT contrast learned in one place of articulation to another place of articulation. A similar experiment conducted on infants, on the other hand, finds that infants do extend a VOT contrast learned in one place of articulation to a different place of articulation (Maye et al. 2002; Maye and Weiss 2003).

[3] An anonymous reviewer points out that English /s/ is not always mapped to /h/ in Maori: [kirihimete] 'Christmas' (De Lacy 2006). While it remains to be seen how this variable adaptation of /s/ can be handled in Herd's (2005) analysis, the crucial contrast between Maori and Hawaiian still stands, namely, given the choice between /h/ and /k/ as a substitute for English /s/, Maori chooses /h/ (at least optionally) while Hawaiian chooses /k/.

[4] Iverson and Salmons (1995) and Avery and Idsardi (2001) propose that the so-called voicing contrast of English stops is in fact an aspiration contrast. Under this alternative analysis, the English speakers' behavior found in Pater's (2003) study is predicted by contrastively underspecified representation and no recourse to non-contrastive phonetic information is necessary.

[5] LaCharité and Paradis (2005) claim that loanword adaptation is mainly phoneme based such that allophonic realizations in L1 are ignored and the mapping applies at the phonemic level. For example, Quebec French does not have a phoneme /h/ but has [h] as an optional allophonic variant of /ʃ/ or /ʒ/, as in /ʃãʒe/→ [h̲ãh̲e] 'to change' and English loanwords with /ʃ/ or /ʒ/ participate in this process as well, as in *shop* → [h̲ɔp]. It is argued that if the segmental mapping occurs at an allophonic level, English /h/ is expected to map to allophonic [h] of Quebec French but, in fact, English /h/ is systematically lost in Quebec French loan adaptation and in L2 English of Quebec francophones, as in *hamburger* → [ambəɹgəɹ] (Janda and Auger 1992; LaCharité and Prévost 1999; Goad and Mah 2007). However, the [h] variant of /ʃ/ and /ʒ/ in Quebec French is one extreme of a spectrum of variants ranging from [ʃ ʒ] to [ʃʰ ʒɦ χ] to [h ɦ] (April 2006 and references cited therein). Given such a range of variants, from the perspective of the exemplar theory, the phonetic category formed over this lenited allophone of sibilants may be quite distinct from /h/ in English, which does not carry any oral place of articulation. Therefore, it is plausible that even at the level of context-specific phonetic categories, English /h/ is more closely matched by ∅ in Quebec French than the lenited allophone of the sibilants.

[6] As an anonymous reviewer points out, [ɖ] and [t̄] are phonetically so close such that no language seems to phonemically contrast these two types of sounds.

Works Cited

Archangeli, Diana. 1988. Aspects of underspecification theory. Phonology 5.183–207.

Avery, Peter, and Keren Rice. 1989. Segment structure and coronal underspecification. Phonology 6.179–200.

Avery, Peter, and William J. Idsardi. 2001. Laryngeal dimensions, completion and enhancement. Studies in distinctive feature theory, ed. by T. A. Hall and U. Kleinhanz, 41–70. Berlin, Germany: Mouton de Gruyter.

Best, Catherine T. 1995. A direct realist perspective on cross-language speech perception. Speech perception and linguistic experience: issues in cross-language research, ed. by W. Strange, 171–204. Timonium, MD: York Press.

Brannen, Kathleen. 2002. The role of perception in differential substitution. Canadian Journal of Linguistics/Revue canadienne de linguistique 47.1–46.

Brown, Cynthia A. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. Second language acquisition and linguistic theory, ed. by John Archibald, 4–65. Oxford, UK: Blackwell.

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

Cho, Taehong, Sun-Ah Jun, and Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. Journal of Phonetics 30.193–228.

Clements, G. N. 2001. Representational economy in constraint-based phonology. Distinctive feature theory, ed. by T. Alan Hall, 71–146. Berlin, Germany: Mouton de Gruyter.

Curtin, Suzanne, Heather Goad, and Joe Pater. 1998. Phonological transfer and levels of representation: the perceptual acquisition of Thais voice and aspiration by English and French speakers. Second Language Research 14.389–405.

De Lacy, Paul. 2006. Markedness: Reduction and preservation in phonology. Cambridge Studies in Linguistics 112. Cambridge, UK: Cambridge University Press.

Dresher, B. Elan. 2003. On the acquisition of phonological contrasts. Proceedings of GALA 2003, vol. 1, ed. by J. van Kampen and S. Baauw, 27–46. Utrecht, The Netherlands: LOT.

Dupoux, E., C. Pallier, K. Kakehi, and J. Mehler. 1999. Epenthetic vowels in Japanese: a perceptual illusion? Journal of Experimental Psychology: Human Perception and Performance 25.1568–78.

Eckman, Fred R., Abdullah Elreyes, and Gregory K. Iverson. 2003. Some principles of second language phonology. Second Language Research 19.169–208.
Escudero, Paola, and Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. Studies in Second Language Acquisition 26.551–85.
Flege, James Emil. 1995. Second-language speech learning: theory, findings, and problems. Speech perception and linguistic experience: issues in cross-linguistic research, ed. by W. Strange, 233–77. Timonium, MD: York Press.
Flege, James E., and Robert F. Port. 1981. Cross-language phonetic interference: Arabic to English. Language and Speech 24.125–46.
Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1995. Effects of age of second-language learning on the production of English consonants. Speech Communication 16.1–26.
Goad, Heather, and Jennifer Mah. 2007. Rethinking English /h/. A paper presented at Montreal-Ottawa-Toronto Phonology Workshop.
Goldinger, Stephen D. 1997. Words and voices – perception and production in an episodic lexicon. Talker variability in speech processing, ed. by Keith Johnson and John W. Mullennix, 33–66. San Diego, CA: Academic Press.
Hall, Kathleen Currie, Amanda Boomershine, Elizabeth Hume, and Keith Johnson. 2004. The influence of contrast vs. allophony on perception. Poster presented at the 148th meeting of the Acoustical Society of America.
Halle, Morris, and G. N. Clements. 1983. Problem Book in Phonology. Cambridge, MA: MIT Press.
Hammerly, Hector. 1982. Contrastive phonology and error analysis. International Review of Applied Linguistics in Language Teaching 20.17–32.
Han, Jeong-Im. 1996. The phonetics and phonology of 'tense' and 'plain' consonants in Korean, dissertation. Ithaca, NY: Cornell University.
Hancin-Bhatt, Barbara Jean. 1994. Segment transfer: a consequence of a dynamic system. Second Language Research 10.241–69.
Heffernan, Kevin. 2005. The role of phonemic contrast in the formation of Sino-Japanese. Journal of East Asian Linguistics 16.61–86.
Herd, Jonathan. 2005. Loanword adaptation and the evaluation of similarity. Toronto working papers in linguistics: Special Issue on Similarity in Phonology, 24.65–116.
Ito, Chiyuki, Yoonjung Kang, and Michael Kenstowicz. 2006. The adaptation of Japanese loanwords into Korean. Studies in loanword phonology: MIT Working Papers in Linguistics 52, ed. by Feng-fan Hsieh and Michael Kenstowicz, 65–104.
Iverson, Gregory, and Joseph Salmons. 1995. Aspiration and laryngeal representation in Germanic. Phonology 12.369–96.
Janda, Richard D., and Julie Auger. 1992. Quantitative evidence, qualitative hypercorrection, sociolinguistic variables and French speakers' 'eadhaches with English h/ø. Language and Communication 12.195–236.
Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. Talker variability in speech processing. ed. by Keith Johnson and John W. Mullennix, 145–65. San Diego, CA: Academic Press.
de Jong, Kenneth. 1995. On the status of redundant features: the case of backing and rounding in American English. Phonology and phonetic evidence: papers in laboratory phonology IV, ed. by B. Connell and A. Arvaniti, 68–86. Cambridge, UK: Cambridge University Press.
Kager, Rene. 1999. Optimality Theory. Cambridge, UK: Cambridge University Press.
Kang, Hyunsook, and Soyoung Kang. 2006. Post-stop tensing rule in Korean: based on perception experiments. Harvard Studies on Korean Linguistics 11.171–82.
Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. Phonology 20.219–73.
Kirchner, Robert. 1997. Contrastiveness and faithfulness. Phonology 14.83–111.
LaCharité, Darlene, and Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. Linguistic Inquiry 36.223–58.
LaCharité, Darlene, and Philippe Prévost. 1999. The role of L1 and teaching in the acquisition of English sounds by francophones. Proceedings of the 23rd Annual Boston University Conference on Language Development, ed. by Annabel Greenhill, Heather Littlefield, and Cheryl Tano, 373–85. Somerville, MA: Cascadilla Press.

Lahiri, Aditi, and William D. Marslen-Wilson. 1991. The mental representation of lexical form: a phonological approach to the recognition lexicon. Cognition 38.245–94.

Maye, Jessica C. 2000. Learning speech sound categories from statistical information, dissertation. Tucson, AZ: University of Arizona.

Maye, Jessica, and LouAnn Gerken. 2000. Learning phoneme categories without minimal pairs. Proceedings of the 24th Annual Boston University Conference on Language Development, ed. by S. Catherine Howell, Sarah A. Fish and Thea Keith-Lucas, 522–33. Somerville, MA: Cascadilla Press.

——. 2001. Learning phonemes: How far can the input take us? Proceedings of the 25th Annual Boston University Conference on Language Development, ed. by Anna H.-J. Do, Laura Domínguez and Aimee Johansen, 480–90. Somerville, MA: Cascadilla Press.

Maye, Jessica, Janet Werker, and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82.B101–B11.

Maye, Jessica, and Daniel Weiss. 2003. Statistical cues facilitate infants' discrimination of difficult phonetic contrasts. Proceedings of the 27th Annual Boston University Conference on Language Development, ed. by Barbara Beachley, Amanda Brown and Frances Conlin, 508–18. Somerville, MA: Cascadilla Press.

Miller, Joanne L. 1994. On the internal structure of phonetic categories: a progress report. Cognition 50.271–85.

Nooteboom, Sieb G., Fred Weerman, and Frank Wijnen. 2002. Minimising or maximizing storage? An introduction. Storage and Computation in the Language Faculty, ed. by Sieb G. Nooteboom, Fred Weerman and Frank Wijnen, 1–19. Dordrecht, The Netherlands: Kluwers Academic Press.

Oh, Mira. 1996. Linguistic input to loanword phonology. Studies in Phonetics, Phonology and Morphology 2.117–26.

Ohala, John J., and Manjari Ohala. 1995. Speech perception and lexical representation: the role of vowel nasalization in Hindi and English. Papers in laboratory phonology IV: Phonology and phonetic evidence, ed. by B. Connell and A. Arvaniti, 41–60. Cambridge, UK: Cambridge University Press.

Pater, Joe. 2003. The perceptual acquisition of Thai phonology by English speakers: task and stimulus effect. Second Language Research 19.209–23.

Pegg, Judith E., and Janet F. Werker. 1997. Adult and infant perception of two English phones. Journal of the Acoustical Society of America 102.3742–53.

Peperkamp, Sharon, and Emmanuel Dupoux. 2003. Reinterpreting loanword adaptations: the role of perception. Proceedings of the 15th International Congress of Phonetic Sciences, 367–70. Adelaide, Australia: Causal Productions.

Peperkamp, Sharon, Michèle Pettinalo, and Emmanuel Dupoux. 2003. Allophonic variation and the acquisition of phoneme categories. Proceedings of the 27th Annual Boston University Conference on Language Development. ed. by Barbara Beachley, Amanda Brown and Frances Conlin, 650–61. Sommerville, MA: Cascadilla Press.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. Frequency Effects and the Emergence of Lexical Structure, ed. by Joan Bybee and Paul Hopper, 137–57. Amsterdam, The Netherlands: John Benjamins.

——. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. Language and Speech 46.115–54.

Port, Robert, and Jonathan Dalby. 1982. Consonant/vowel ratio as a cue for voicing in English. Perception and Psychophysics 32.141–52.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: constraint interaction in Generative Grammar. Rutgers University Center for Cognitive Science Technical Report 2.

Sánchez, Alicia Muñoz. 2003. The effect of phonological status on the acquisition of new contrasts: evidence from Spanish and Japanese L2 Learners of English, dissertation. San Diego, CA: University of California, San Diego.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. Phonology 9.289–328.

Steriade, Donca. 1987. Redundant values. Papers from the 23rd Annual Meetings of the Chicago Linguistics Society 23.339–62.

Sung, Eunkyung. 2003. Flaps in American English and Korean: production and perception, dissertation. Newark, DE: University of Delaware.

