External Sandhi in a Second Language:

The Phonetics and Phonology of Obstruent Nasalization in Korean-accented English

Elizabeth Zsiga

Dept. of Linguistics, Georgetown University

Washington, DC 20057

202-687-2238

zsigae@georgetown.edu

External Sandhi in a Second Language:

The Phonetics and Phonology of Obstruent Nasalization in Korean-accented English

Abstract

This paper presents the results of an acoustic study of nasal assimilation and gestural overlap at word boundaries in Korean and Korean-accented English. Twelve speakers of Seoul Korean recorded phrases containing obstruent#nasal and obstruent#obstruent sequences in both Korean and English. Nasalization of the word-final obstruent, predicted by the rules of Korean phonology, occurred in 93% of obstruent#nasal sequences in Korean and in 32% of such sequences in Korean-accented English, a rate of application higher than that reported in most other studies of external sandhi alternations in non-native speech. Acoustic analysis found categorical nasalization in the L1 Korean productions, but both categorical and gradient nasalization, along with a high degree of inter- and intra-speaker variation, in the L2 English productions. For a subset of speakers, there was a significant correlation between quantitative measures of nasalization in English and measures of consonant overlap in the English obstruent#obstruent sequences. An analysis in terms of articulatory gestures and the coupled-oscillator model of speech planning is supported, The analysis is based on the Articulatory Phonology model (Browman & Goldstein 1990, 1992, 2000, Goldstein et al. 2006), though with modifications. Implications for phonetic and phonological representations, and for speech planning in both L1 and L2, are explored.

* Thanks are due to Soojeong Eom and Kimberly Teague for help with materials preparation, data collection, and data analysis; to Hyouk-Keun Kim for collaboration on earlier work on this topic; and to two anonymous reviewers, whose advice greatly improved both theory and analysis.

1. Introduction. The term external sandhi (from the Sanskrit word for juncture) refers to phonological alternations that take place across a word boundary. In Korean, for example (Kim-Renaud 1991, Sohn 1999, Stuart & Shin 1999), a syllable-final obstruent becomes a nasal when the following syllable begins with a nasal, even when a word boundary intervenes, as in (1):

(1) Korean obstruent nasalization at word boundaries

[pap]	rice	[pam mekta]	eat rice
[ot]	clothes	[on man]	only clothes
[jak]	medicine	[jaŋ mekta]	take medicine

The study of external sandhi alternations in any language has the potential to shed light on issues such as the interaction of grammar and the lexicon, the nature of phonological representations, and phrasing and speech planning. When a native and non-native language (L1 and L2) differ in patterns of external sandhi, investigating how learners handle these junctures—following the L1 pattern, the L2 pattern, or an interlanguage-specific pattern—provides information not only on how phonological patterns are represented and speech plans are carried out, but how such cognitive structures are learned, how they change, and where intervention may or may not be necessary.

Most obviously, inappropriate application in the L2 of an L1 pattern of external sandhi impedes communication. Chu & Park (1978) and Kim (2000) note that a Korean speaker of English, applying the Korean pattern illustrated in (1), is likely to pronounce phrases such as Pick me up as Pi[ŋ] me up or I cut myself as I cu[n] myself, creating a 'major problem . . . for an average Korean learner of ESL [English as a Second Language]' (Chu & Park p. 1). Contrary to these reports, however, other studies of external sandhi in L2 speech (reviewed in more detail in section 3 below), have found that patterns of external sandhi rarely carry over from L1 to L2. Cebrian (2000) suggests in fact that L2 speakers generally obey a word-integrity constraint that 'prevents the synchronization of sounds belonging to different words' (p. 19), thus blocking external sandhi from applying. Yet the careful separation of words can also lead to problems in communication. If the target language requires close connection between words and the application of sandhi processes, learners who fail to connect their words appropriately will sound stilted. Rhythm and prosody are important aspects of (mis)understanding in L2 speech: correct phrasing not only signals fluency, but also aids understanding. A too-careful articulation may result in the listener incorrectly perceiving extra syllables and stronger phrase boundaries than the speaker intends, which may impede understanding as much as using an incorrect segmental allophone (see, e.g. Anderson-Hsieh, Johnson & Koehler, 1992; Anderson-Hsieh, Riney & Koehler 1994; Flege, Munro & MacKay 1995; Munro and Derwing 1995; Tajima, Port & Dalby 1997; Trofimovich & Baker 2006; see also Cutler, Dahan & van Donselaar 1997 on the general importance of prosody in comprehension).

Thus, the study of external sandhi in L2 is important from a practical standpoint. Although it is clear that inappropriate over- or under-application of sandhi will cause problems for the learner, few language pairs have been studied in this area, and results have been conflicting. Cebrian (2000) found that Catalan speakers of English failed to apply voicing assimilation at word boundaries even when it would aid communication, while Kim (2000) found that Korean speakers of English often applied nasal assimilation even when it impeded communication. The reason for these different conclusions is not clear (see the discussion in section 3), but conflicting findings in previous research point to the need for further study. In addition, underlying the practical considerations for L2 learners, and determining how they can and should be addressed, are important theoretical considerations. The study of external sandhi in L2 raises the question of what exactly it is that's being carried over (or not), and thus has the potential to shed light on questions of phonological and phonetic representation and speech processing.

External sandhi offers an opportunity to explore the productivity of phonological alternations--the ability of a speaker to generalize beyond a static set of learned examples--and the form such generalization takes. For word-internal alternations, it is not always easy to separate phonology from the lexicon, especially where an alternation is indicated orthographically. For example, a phonologist might argue that the basic form of one English negative prefix is /in-/, and that the /n/ undergoes a phonological change to [m] in words like impolite, and other phonologists might debate the representation of this rule. Yet it is also reasonable to propose that impolite is memorized as a complete lexical item, with the /mp/ sequence in place, and that there is thus no rule to argue about (see, e.g. Bybee 2000, 2002). However, because word combinations may be novel, there can be no stored representation of all external sandhi outcomes. This is not to say that no word pairs are stored: there is ample evidence that common word combinations such as I don't know and would you are stored as lexical units, and some evidence that less common combinations have a persistent mental representation as well (Bybee 2002, Erman & Warren 2000). However, to the extent that phonological alternations occur across word boundaries in novel phrases, this is evidence for the existence of a general rule that has been abstracted from the data and that exists independent of its specific instantiations.

The persistence of L1 sandhi processes in L2 is particularly strong in making this point: in such cases the application of the rule is completely divorced from the lexical items that originally gave rise to the generalization. If a Korean speaker of English says kee[m] Matt on the team instead of keep Matt on the team, it is unlikely that she is repeating a previously-heard or stored pronunciation. So the study of external sandhi has the potential to separately focus on the general principles or plans that govern pronunciation—that is, the grammar—apart from the lexicon.

The study of external sandhi has developed along several different dimensions. One area of research has focused on determining the domains over which external sandhi applies, and how those domains should be specified. The processes of external sandhi have served as the basis for the development of phonological theories of the prosodic hierarchy (e.g. Inkelas & Zec 1990, Kaisse 1985, Nespor & Vogel 1986, Selkirk 1984, 1986); and numerous phonetic studies have investigated the ways that prosodic structure influences the shape and timing of articulatory movements. The work of Cho, Keating, and colleagues, for example, (Cho 2002, 2007; Keating 2006; Keating, Cho, Fougeron & Hsu 2003) addresses the ways in which processes of coarticulation, lengthening and strengthening make reference to the domains and boundaries of the prosodic hierarchy. Another approach is found in the work of Byrd, Saltzman, and colleagues (e.g., Byrd 2006; Byrd, Kaun, Narayanan & Saltzman 2000; Byrd, Lee, & Campos-Astorika 2008; Byrd & Saltzman 2003, Saltzman, Löfqvist, Kinsella-Shaw, Kay & Rubin 1995; Saltzman, Nam, Krivokapic & Goldstein 2008), in which prosodic effects are modeled by modulation gestures that influence the timing of articulatory gestures that occur at or near prosodic boundaries. In this latter approach, different boundary effects are modeled not by imposing different kinds of category boundaries according to the prosodic hierarchy, but by varying the strength of the influence of the modulation gestures: greater overall slowing, for example, is perceived as a stronger boundary.

The present study focuses on a different (though obviously related) area of research: the nature and representation of the sandhi alternations themselves, and the ways in which external sandhi can provide insight into the nature of stored representations and generalizations. The theory of Articulatory Phonology (Browman & Goldstein 1986, 1990a,b, 1992, Goldstein, Byrd, & Saltzman 2006) argues for two hypotheses: 1) that phonological contrasts are represented in terms of articulatory gestures, not phonological features, and 2) that all external sandhi alternations are the result of changes in the timing and magnitude of these gestures. The evidence for gestural reorganization rather than feature change comes from phonetic studies that show that many external sandhi changes are gradient and non-neutralizing. For example, Browman & Goldstein (1990a) provide articulatory evidence that while phrases like in polite society may be perceived as identical to impolite society, the tongue tip closing gesture for the [n] in the former phrase is still present. They thus argue that the apparent change from /n/ to [m] is not the result of a featural change from [coronal] to [labial], in which case no trace of the underlying coronal would be expected to remain, but the perceptual result of pronouncing an [n] and [p] at the same time. For lexical changes such as impolite, for which there is no phonetic evidence of gradience, Browman & Goldstein argue that the [mp] is part of the stored lexical representation, as noted above. Thus the only productive phonology is gestural phonology.

Researchers within the theory of Articulatory Phonology have found many examples of external sandhi alternations that appear to be best described as gradient gestural overlap or reduction (e.g. Barry 1992, Browman & Goldstein 1992 and references therein, Chen 2003, Ellis & Hardcastle 2002, Kochetov & Pouplier 2008, J. Jun 1996, S.-A. Jun 1995, Zsiga 1995, 1997).

If the theory of Articulatory Phonology is correct in the claim that all external sandhi is the result of articulatory reorganization, then sandhi processes are not evidence for a feature-changing grammar. Rather, to the extent that there are rules for combining words, these rules consist of instructions for how articulatory gestures are to be coordinated, not how segments are to be changed. For the L2 learner and teacher, this sets an entirely different L2 target to be attained.

While the references cited above provide clear phonetic evidence that some external sandhi processes are non-neutralizing, a number of other cases of external sandhi that *are* apparently categorical and neutralizing have also been put in evidence (e.g, Bradley 2007, Ellis & Hardcastle 2002, Holst & Nolan 1995, Honorof 1999, Kochetov & Pouplier 2008, Ladd & Scobbie 2003, Scobbie & Wrench 2003). (The majority of sandhi alternations described in the literature have simply not been phonetically tested.) The existence of categorical external sandhi casts doubt on the claim that all productive phonology is a matter of gestural reorganization. Ladd & Scobbie (2003:16) conclude 'that gestural overlap is on the whole *not* a suitable model of most of the assimilatory external sandhi phenomena in Sardinian, and more generally that accounts of gestural overlap in some cases of English external sandhi cannot be carried over into all aspects of post-lexical phonology.' Ladd & Scobbie argue instead for an analysis of their Sardinian data in terms of autosegmental spreading.

In the Articulatory Phonology model, gestural reorganization is never categorical, and thus categorical external sandhi alternations are argued not to exist. Cases that appear to be categorical deletion or assimilation are argued to be outliers in the range of gestural variation: deletion being the limiting case of reduction and complete assimilation the limiting case of overlap (Browman 1995). For example, Son et al. (2007: 215) describe the change of word-medial /pk/ --> [kk] in Korean, for which they show the outcome to be indistinguishable from an underlying /kk/, as an extreme case of 'lip aperture reduction.'

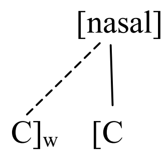
However, given the phonetic evidence for both kinds of process, Son & Pouplier (2008) make the point that speakers' linguistic competence must contain knowledge of both gradient and categorical processes (see also Scobbie 2007). The crucial question is whether there is a theory of gestural timing and organization that is both powerful enough to account for gradient changes, and constrained enough to account for changes that result in category neutralization (see the discussion in Ladd & Scobbie 2003, Zsiga 1997). It will be argued here that, given recent innovations in the specification of gestural timing (discussed in section 2), Articulatory Phonology has the resources to be such a theory, although further modifications, incorporating some of the capacities of Autosegmental Phonology, is required.

The phonetic and phonological study of external sandhi is crucial to the debate over phonological representation, because external sandhi is necessarily productive and non-lexical, as argued above, and because it has been shown to exhibit, in different instances, both gradient and categorical properties. The importance of the question is only intensified when the perspective of L2 learning is added. If the task of the L2 learner is to attain speech patterns that are like those of native speakers, it is crucial to ask what the set of possible temporal patterns may be, and how different patterns may be across languages.

The next section now turns to the description of theories of external sandhi alternations, beginning with L1 studies (section 2), focusing in particular on the similarities and differences between the autosegmental and articulatory approaches. Section 3 discusses research that has addressed external sandhi in L2. Experimental findings for the present study are reported in sections 4 and 5, and section 6 concludes.

2. External sandhi in L1: Autosegmental and Articulatory Phonology. Traditional phonological descriptions of external sandhi have in general appropriated the conventions used for word-internal phonological alternations, simply adding a description of the domain or boundary over which the rule applies. Thus, an approach to phonology that assumes distinctive features and autosegmental association would represent the Korean alternation exemplified in (1) with the rule in (2).

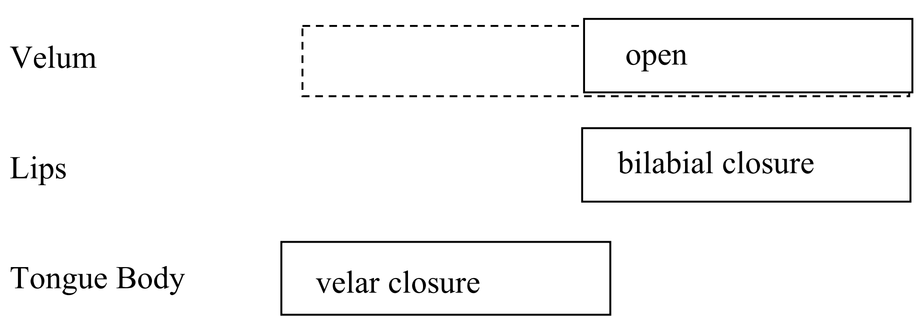
(2) Autosegmental rule for Korean obstruent nasalization at word boundaries



In this representation the nasal feature begins with an association to the second consonant in the cluster (solid association line), and then spreads to become associated with the preceding consonant (dotted association line), across a word boundary ($]_{w}$). The focus here is not on how the boundary is indicated, nor on whether the new association comes about as a result of rule application (Goldsmith 1976) or constraint interaction (Prince & Smolensky 2004). Rather, the point is that in this phonological approach, external sandhi alternations are represented in the same way as word-internal alternations: as distinctive features categorically associated and re-associated to segmental constituents.

An Articulatory Phonology approach to Korean nasalization would model the nasalization as gestural overlap between the word-final consonant and the velum opening gesture for the word-initial nasal. There are different ways to model increased overlap; one way would be to assume that the nasal gesture is extended in time, so that it is coextensive with both consonant gestures, as shown in the gestural score in (3). (This is similar, for example, to the gestural extension argued for by Zsiga 1997.) In a gestural score representation (Browman & Goldstein 1986), each gesture is indicated by a rectangle, whose length indicates the gesture's activation interval, a measure of its extent in time. Three gestures are shown: tongue body closing for [k], and labial closing and velum opening for [m]. A shorter nasal gesture (solid rectangle) would result in nasalization only on the word initial consonant (thus [jak#mekta]); a longer nasal gesture (dashed rectangle) would result in nasalization on both consonants ([jaŋ#mekta]). It would also be possible to model nasalization with greater overlap of the oral closing gestures and no change in the temporal extent of the velum opening gesture, but this would predict assimilation of place as well as nasality across word boundaries, which is contrary to what has been found in Korean (see Kochetov & Pouplier 2008; Son, Kochetov & Pouplier 2007, and results below).

(3) Hypothetical gestural score for Korean [k#m] pronounced as [ŋ#m]



A strength of the Articulatory Phonology approach is its simplicity. The same units, articulatory gestures, suffice for both the description of phonological contrast and the exact modeling of articulator movement (Browman & Goldstein 1989, 1990b, 1992; Goldstein, Byrd, & Saltzman 2006; Goldstein, Nam, Saltzman & Chitoran in press; Nam, Goldstein, & Saltzman 2009). In the lexicon, words contrast in the presence or absence of gestures (mad has a velum opening gesture that bad lacks) and in their relative timing (mad and ban have the same gestures, but differ in whether the velum opening gesture is associated with the labial or alveolar closure).

The details of articulatory trajectories arise as the abstract gestural targets are realized in specific articulations that unfold in space and over time. Different gestures may compete for control of a specific articulator, or may interfere with one another in various ways, resulting in the details of allophonic realization (such as vowel nasalization in ban) or the assimilations and deletions that have been attributed to rules of external sandhi. For example, phonetic studies (e.g. Barry 1985, Browman & Goldstein 1990, Byrd 1996, Zsiga 1994) show that in a C1#C2 sequence within a phrase in English, closure for C2 is reached before the release of C1 is effected. English speakers start producing the [p] in a phrase like hit parts before the closure for the [t] has been released, and they begin the [j] in a phrase like miss you while the [s] is still being articulated. Such overlap often causes the perception of assimilation: hit parts sounds like hip parts and miss you sounds like mish you.

Because gestures are units of both contrast and implementation, in this approach there is no phonology-to-phonetics translation, in which features or other abstract units must be mapped into corresponding physical parameters. Articulatory Phonology thus posits a single set of primitives (gestures), while Autosegmental Phonology must posit at least two (abstract distinctive features and their physical instantiations.)

However, as compared to Autosegmental Phonology, Articulatory Phonology as originally conceived (Browman & Goldstein 1986, 1992) has been argued to both overgenerate contrasts, by allowing too many possible timing relations between gestures, and to undergenerate alternations, in not allowing for any categorical change outside the lexicon. The issue of possible timing relations has been addressed in more recent work on gestural timing, as discussed below. The question of categorical alternations remains open, and is the main focus of the present paper.

In the original formulations of Articulatory Phonology (e.g. Browman & Goldstein 1989, 1990b, 1992; Saltzman 1986; Saltzman & Munhall 1989) the gesture was defined as a 360° cycle, and the co-ordination of different gestures (gestural phasing) was defined as the specification of any points within the cycles of two gestures as being simultaneous. For example, Zsiga (2000) proposes that the pattern of consonant overlap at word boundaries in English, such that the second consonant in a C1#C2 sequence achieves closure just before the first consonant closure is released, should be specified as an alignment constraint of the form: Align(C1,300°,C2,270°). Later proponents (e.g. Gafos 2002, Bradley 2007) suggest limiting possible coordinated points to five landmarks: onset, target, c-center, release, offset. While this reformulation limits the possible patterns of coordination, it is still many more degrees of freedom than traditional autosegmental phonology allows.

Autosegmental representations allow just two temporal relations to be defined: linear precedence of features or segments on a single tier, and lack of linear precedence (simultaneity) in features linked to a single root or class node (Sagey 1988, Zsiga 1997). As noted by Sagey (p. 112), multiple linking to a root node does not connote actual perfect simultaneity, only some unspecified degree of overlap: 'some instant' in the realization of one feature is simultaneous with some other instant in the realization of another feature. Exact degrees of overlap are not 'accessible to or manipulable by phonological processes,' and thus the number of possible contrasts is constrained. For example, velum opening precedes labial closing in the articulation of post-vocalic [m] (Krakow 1999), but representation of the nasal stop with unordered features [nasal] and [labial] prevents the phonology from referencing multiple contrastive degrees of nasalization.

Further refinements of the theory of gestural timing (Browman & Goldstein 2000, Goldstein 2008, Goldstein et al. 2006, Nam et al. 2009, Nam & Saltzman 2003, Saltzman et al. 2008) however, have brought Articulatory Phonology and Autosegmental Phonology closer together in terms of the types of contrasts predicted. These recent innovations have added a planning component to the gestural model, which limits the types of timing relationships that can be specified. The new model connects models of articulatory coordination to models of other types of coordinated movements, through the theory of coupled oscillators (Haken, Kelso, & Bunz 1985, Turvey 1990, Löfqvist & Gracco 1999). These sources argue that while complicated rhythmic patterns can be learned with skill and practice, as in drumming, there are only two natural and easily-acquired patterns of coordination: in-phase (simultaneous) and anti-phase (sequential), with simultaneous preferred. (The same point is made with reference to articulatory coordination by de Jong 2003.)

In the planning component, coordination between gestures may be specified only in terms of the two coupling modes, represented as an intergestural coupling graph, which resembles an autosegmental representation. More important than graphical resemblance, however, is the fact that the two coupling modes correspond to the two types of timing that are recognized in autosegmental phonological descriptions: simultaneity and precedence. In order to compute actual trajectories, a gestural score is generated from the coupling graph, such that gestural activations are triggered according to the modes specified in the coupling graph. Gestural phasing is thus derived from the couplings, not independently stipulated. The planning component then feeds into an implementation component, which in turn generates articulatory trajectories consistent with the active set of gestures, according to general dynamical laws (see Goldstein et al. 2006: 219).

The more detailed timing patterns of actual speech arise because multiple couplings may be specified, some of which may be incompatible with each other, and which then must compete. For example, articulatory studies (Browman & Goldstein 1988, 2000; Goldstein, Chitoran & Selkirk 2007, Goldstein et al. in press) have found different timing patterns for consonants in onsets and codas: the consonants in a [pl] sequence in an onset are more overlapped with the vowel than are the consonants in an [lp] sequence in a coda. In the coupled oscillator model, this asymmetry is explained in terms of competing couplings in the onset but not in the coda (Browman & Goldstein 2000). Consonants in the coda are coupled in the sequential mode only, and are thus realized sequentially. Consonants in the onset, however, have competing couplings: each consonant gesture is coupled in sequential mode with each other consonant gesture, but also in simultaneous mode with the vowel. In addition to accounting for the articulatory timing data, the fact that onset consonants have the preferred simultaneous coupling can explain, it is argued, the universal preference for CV over VC syllables. The proposed coupling relationships are represented in the coupling graph in (4). A simultaneous coupling is indicated with a solid line, a sequential coupling with a dotted line. (Note that this is different from an autosegmental diagram, in which a solid line indicates an underlying association and a dashed line a derived one.)

(4) Coupling graph for consonant sequences in English onsets and codas (following Browman & Goldstein 2000, Goldstein et al. 2006)

a. competing couplings in onset position b. non-competing couplings in coda position

In the onset case (4a), the specification for both consonants to be simultaneous with (that is, to begin at the same time as) the vowel competes with the specification for the second consonant to follow the first in sequential order. In order to compute actual trajectories, the two specifications are averaged, weighted according to their coupling strength, with the result that the two consonants are neither completely sequential nor completely simultaneous, but overlapped. Different degrees of overlap may arise when different coupling patterns are present, or when different coupling strengths apply.

The potential for correspondence with autosegmental representations is obvious. There is a close correspondence between the features of Autosegmental Phonology and the gestures of Articulatory Phonology: the features [labial] and [-continuant], for example, map straightforwardly into 'labial closing gesture' (see Zsiga 1997 for discussion). The in-phase couplings in a coupling graph will in a great many cases represent the same relations as the association lines of Autosegmental Phonology: for example linking velic or laryngeal gestures to oral closing gestures in a segment-sized unit. In larger domains, it may be hypothesized that the presence of a coupling, either simultaneous or sequential, indicates grouping within a prosodic domain, so that in (4) above C and V belong to the same syllable. (The idea of using couplings to represent prosodic constituency is further explored in Goldstein et al. in press, Nam

2007, Nam et al. 2009, Nava et al. 2008; see also Saltzman et al. 2008). Different coupling strengths may indicate different prosodic levels, with smaller domains imposing a tighter coordination (Byrd & Saltzman 2003).

The correspondence is not perfect. Not all features correspond clearly to gestures: [sonorant], for example, has no straightforward gestural correlate. Nor are the hypothesized hierarchical groupings (or constellations) always the same. For example, Articulatory Phonology does not posit the existence of the segment per se, though many of its constellations are segment-sized. Conversely, most models of Autosegmental Phonology do not recognize an onset constituent, which is modeled in Articulatory Phonology. A further complication is introduced by Nam (2007) who hypothesizes that couplings may differentiate between closure and release gestures for each oral constriction: different cross-linguistic timing patterns may require coupling of closure-to-closure vs. closure-to-release (see also Browman 1994). Most autosegmental representations do not represent closure and release as distinct nodes, but some do (e.g. Steriade 1999).

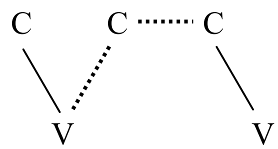
Crucially, given the introduction of the planning module and coupling graphs indicating only in-phase and anti-phase coupling, Articulatory Phonology and Autosegmental Phonology now posit the same two degrees of freedom in contrastive temporal relations. The newer version of Articulatory Phonology no longer overgenerates phonological contrast. With coupling graphs, many of the insights of Autosegmental Phonology can be expressed in gestural terms, and further generalizations (such as onset/coda asymmetries) are given a new explanation. Importantly, in adding the planning module, Articulatory Phonology has not lost its ability to account for the details of gradient phonetic implementation. However, the issue of categorical alternations remains.

As was noted above, Articulatory Phonology hypothesizes that all within-word allomorphy (as in i[m]polite) is part of the lexical specification of each word, and that all external sandhi is the result of gradient changes in gestural magnitude and overlap. It follows from these hypotheses that there is no need for any statement of categorical phonological alternation in the grammar. While Articulatory Phonology recognizes that language users make choices among phonological forms, the theory makes no provision for formalizing a category-changing alternation. Thus, to the extent that there are truly categorical, productive, external sandhi alternations (as has been argued by Ladd & Scobbie 2003 and other references cited above), Articulatory Phonology as currently implemented does not account for them.

According to Ladd & Scobbie (2003), two different representations are needed for different processes of external sandhi: feature spreading for categorical alternations and gestural overlap (or some other gradient phonetic implementation) for gradient changes. However, the addition of coupling graphs to the Articulatory Phonology model offers an opportunity to account for categorical as well as gradient alternations at word boundaries in terms of articulatory gestures. At present, the model (deliberately) offers no component in which categorical reorganization of gestures across word boundaries can take place. It is one goal of the present paper to argue that such a component can and should be added.

As the Articulatory Phonology model is currently formulated (e.g. Goldstein et al. 2006) gestural coupling graphs are created in the lexicon, indicating the relationships between gestures that create distinct words. When words are put together into phrases, couplings may be added to coordinate word-sized constellations as a whole, and modulation gestures serve as overall clocks to regulate prosody-based tempo (Saltzman et al. 2008). Nava et al. (2008) model the close coordination of English words within a foot by means of a sequential coupling between a word-final and word-initial consonant, as shown in (5).

(5) Coordination of two syllables across a word boundary, within a foot (Nava et al. 2008)

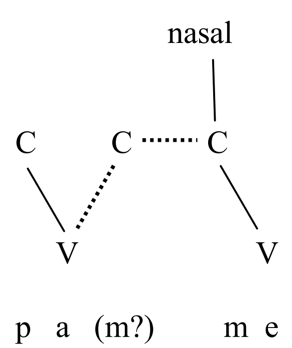


Following Browman and Goldstein (2000), onset consonants are modeled as simultaneous (completely overlapped) with the vowel of their syllable (solid lines in the graph), while coda consonants are sequential to the vowel, beginning when the vowel ends (dotted lines). In addition, a sequential coupling between the consonants of adjacent syllables is hypothesized, indicating membership within a domain (here, the foot). Syllables that were not members of the same prosodic domain would lack this consonant-to-consonant coupling.

In order to account for language-specific differences in degree of overlap (Kochetov, Pouplier & Son 2007, Nava et al. 2008, Yanagawa 2006, Zsiga 2000), it may be necessary to make reference to release gestures in addition to closing gestures (Nam 2007). The more overlapped pattern of English is modeled by closure-to-closure coupling: thus movement toward C2 closure begins during C1, and the release of C1 is obscured by the closure for C2. A less overlapped pattern, typical of most clusters in Russian and Georgian (Goldstein et al. 2007, in press), can be modeled by release-to-closure coupling: delaying the closure for C2 until the release for C1 has been effected, rendering the release audible. (The least overlap is found in back-to-front clusters, such as [k#p], where the release of C1 is most likely to be acoustically hidden in the case of overlap.) Section 6.2 returns to these language-specific details (after additional phonetic data has been presented).

If obstruent nasalization in Korean is the result of gestural overlap, as is predicted by Articulatory Phonology for all external sandhi, it might be represented with a coupling graph as in (6). The graph is the same as that in (5), except that an additional nasal gesture for the [m] has been added. Articulatory studies of consonant overlap in Korean (J. Jun 1995, 1996; Kochetov et al. 2007; Kochetov & Pouplier 2008; Son et al. 2007) agree in finding that Korean speakers produce close transitions between obstruents at a word boundary, similar to the English pattern, as long as the words fall within a single accentual phrase. Close transitions are also consistent with the findings on voicing and prosodic phrasing by S.A. Jun (1993, 1995, 1996) and Silva (1992). Kochetov et al. (2007) specifically find that Korean consonant sequences are typically more overlapped than sequences in Russian, though they test only a subset of possible consonant pairs, and note considerable variation.

(6) Korean nasal assimilation (/pap mekta/ --> [pam mekta] as in (1) above), as gradient gestural overlap

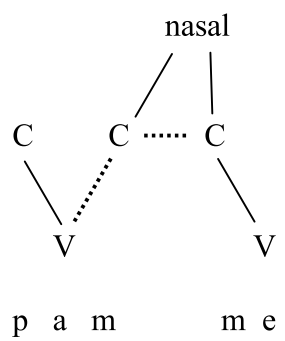


If the C-to-C coupling coordinates closure to closure, there will be no intervening release burst for the final consonant. Further, the close coordination of the two consonants may result in some variable and gradient nasalization being realized on the word-final consonant. The exact degree of overlap (and therefore nasalization) would depend on the various coupling strengths involved, on whether any other couplings might be relevant, possible addition of phrasal modulation gestures, and some level of noise inherent in any physical system (Nam & Saltzman 2003).

Crucially, in this representation, any nasalization of the word-final obstruent comes about not as the result of a rule or constraint targeting the nasal gesture per se, but as a byproduct of a more general pattern of word-to-word coordination. The general plan of coordination for Korean specifies closure-to-closure coordination for all consonant sequences. That pattern of overlap gives rise to the perception of nasalization on the word-final consonant just in case the configuration includes a word-initial nasal gesture. Thus the approach diagrammed in (6) follows the Articulatory Phonology analysis of place assimilation in English (Browman & Goldstein 1990a): there is no rule of assimilation, there is just the perceptual consequence of the language-specific pattern of overlap. If (6) is the case, nasal assimilation in Korean is expected to be partial and gradient, and traces of the underlying non-nasal specification of the coda consonant are expected to remain.

There is, however, currently no provision in Articulatory Phonology for any external sandhi alternation that is not partial and gradient. As the theory is currently formulated (Goldstein et al. 2006, Saltzman et al. 2008), couplings and modulation gestures added to the coupling graph at the phrasal level do not target specific gestures. However, there is no reason (in principle) why they could not. The addition or deletion of coupling relations, parallel to the addition or deletion of autosegmental association lines, could be used to model categorical alternations outside the lexicon. A hypothetical coupling graph for categorical obstruent nasalization in Korean is shown in (7).

(7) Categorical Korean nasal assimilation: /pap mekta/ --> [pam mekta] as in (1) above.



In (7) nasalization of the coda consonant comes about through the addition of a new coupling between the nasal specification and the coda consonant, capturing the same essential insight as autosegmental spreading: the nasal specification that begins as the property of the word-initial consonant is now coupled to both consonants in the sequence. The crucial innovation is that the new coupling is added outside the lexicon, across a word boundary, as the result of a grammatical generalization targeting a specific configuration.

The graph in (7) is hypothetical, and exact specifications remain to be worked out. The addition of the new coupling could cause the nasal gesture to stretch (as in (3) above), or it could force the two consonant gestures to overlap more (as in the onset consonant effect in (4a)): the outcome would depend on the specification of the inherent stiffness of the nasal gesture and the strengths assigned to each coupling. To date, studies of the implementation of competing couplings have focused on oral constriction gestures, but the acoustic findings of Bauer (2005) suggest that specifications for oral and nasal gestures will not always be parallel.

Crucially, however, the addition of the new coupling has created a categorical change in the word-final obstruent: it is now associated with a [nasal] specification that it did not start out with. In such an approach to external sandhi, the question isn't: is this pattern due to gestural overlap or feature spreading? (as the problem is defined by Ladd & Scobbie 2003, for example). Rather, the question is: is this pattern due to the addition of a new (categorical) gestural association across the word boundary, or does this pattern arise from more general patterns of word-boundary overlap in the language that happen to result in the perception of assimilation in this particular context? In such an approach, the question of whether the representation in (7) is a coupling graph or an autosegmental representation may not be important. The crucial insight is that the same relations are captured in both representations. Phonology (the planning module) supplies the couplings (the specification, addition, and deletion of associations) and phonetics (the implementation module) interprets the couplings (via the application of different coupling strengths and the computation of actual trajectories through averaging of competing constraints.)

The proposed model departs from the original hypotheses of Articulatory Phonology in allowing the addition and deletion (not just lexical specification) of categorical associations between specific gestures – the essential insight of Autosegmental Phonology. It differs from Autosegmental Phonology in adopting the gesture as its basic unit, and assuming the gestural constellations that Articulatory Phonology proposes. Such differences are important, and contradictions remain to be worked out. The large-scale correspondences, however, render the promise of a unified approach worth pursuing.

Of course, it remains to be seen whether either (6) or (7) is a plausible representation for external sandhi in general and Korean nasal assimilation in particular. An analysis of planned reorganization such as that in (7) rests on finding that obstruent nasalization is categorical. A gestural overlap analysis (6) requires two findings: first, that nasalization is variable and gradient, and second, that that the general pattern of word-boundary coordination in Korean is indeed that of overlap, similar to English (Byrd 1996, Zsiga 1994), rather than lag, similar, e.g., to Russian (Zsiga 2000, Kochetov & Goldstein 2001, Kochetov 2006b). These hypotheses are tested directly in the two experiments reported in the present paper: Experiment 1 examines whether obstruent nasalization is categorical or gradient, in both native Korean and in Korean English. Experiment 2 examines consonant overlap in non-nasal sequences for the same set of speakers, and seeks to relate the more general patterns of overlap to the nasalization results. It will be argued that representations of both gradient (6) and categorical (7) assimilation are needed.

At present, phonetic data on external sandhi in L1 Korean is preliminary, and further acoustic details are sought in the experiments reported here. Son et al. (2007) and Kochetov & Pouplier (2008) argue that (optional) place assimilation in Korean compounds is sometimes categorical, and is otherwise the result of gestural overlap. J. Jun (1995, 1996) finds a more important role for reduction of C1 in a C1+C2 cluster, and argues that assimilation is due to gradient reduction, not overlap. None of these articulatory studies examines cross-word nasalization, however, nor the other external sandhi processes of Korean. Still less is known about the interaction of consonant timing patterns in Korean-accented English, or in L2 speech at all.

3. External sandhi in L2 speech. Few studies have examined external sandhi in second language phonology. Those that have find conflicting results. This section reviews the findings of Cebrian (2000), which also summarizes some earlier studies, Zsiga (2003), Lléo & Vogel (2004), and for Korean English in particular Kim (2000) and Park (2005).

Cebrian (2000) examined the speech of Catalan learners of English, and found that the process of word-final devoicing carried over from Catalan to English much more often than the process of voicing assimilation across word boundaries. In Catalan, obstruents are pronounced as voiceless in phrase-final position, but assimilate to the voicing of a following stop when another word follows in the phrase, as shown in (8).

(8) Catalan word-final devoicing and voicing assimilation at word boundaries (Cebrian 2000)

	/vaz/		/gos/
[vas]	glass	[gos]	dog
[vazos]	glasses	[gosos]	dogs
[vas petit]	small glass	[gos petit]	small dog
[vaz gran]	big glass	[goz gran]	big dog

When speaking English, Catalan speakers were much more likely to apply (within word) devoicing than (cross-word) voicing assimilation, tending to pronounce phrases such as wise guy as wi[s g]uy. In Cebrian's data (Tables 9 and 10, pp. 12, 14), 97% of underlyingly voiced obstruents were devoiced preceding a pause or a voiceless consonant, but only 20% of voiceless obstruents were voiced before a voiced stop. The failure of voicing assimilation to apply is especially striking because in many cases application of the Catalan assimilation rule would have resulted in the correct English pronunciation, as in the case of wise guy. Based on the asymmetry in the rate of application of word-internal vs. external sandhi rules, Cebrian proposes the word integrity constraint, defined as 'an interlanguage prosodic constraint that treats every word as a separate unit and prevents the synchronization of sounds belonging to different words' (2000:19). Non-content words and clitics are argued to form a single phonological word with the host and are thus exempt, explaining the findings of higher rates of application at what were called word boundaries in previous similar studies (Altenberg & Vago 1983, Rubach 1984, Solé 1997). If each word is treated as a separate, unconnected unit, then no rules will apply across word boundaries, and there will be no gestural overlap. Thus, Cebrian predicts that external sandhi should be rare or nonexistent in L2 speech. In what follows, this prediction will be referred to as the Word Integrity hypothesis. Note, however, that even in Cebrian's data the Word Integrity hypothesis is sometimes violated, and assimilation does sometimes apply, indicating that other forces that conflict with Word Integrity will take precedence in some cases or for some speakers.

A further point worth noting in Cebrian's data is that there was an asymmetry between [+continuant] and [-continuant] segments: only 6% fricatives and affricates underwent voicing, compared to 34% of stops. Cebrian suggests that part of this difference may be due to the fact that analysis was by transcription, and that voicelessness is 'more clearly perceived' in fricatives (p. 13); that is, a partially voiced stop will be transcribed as voiced, but a partially voiced fricative will be transcribed as voiceless. The problems inherent in relying on transcription, especially in cases of partial assimilation, point to the need for instrumental phonetic study.

Two subsequent studies, Zsiga (2003) and Lléo & Vogel (2004) find general support for a principle of Word Integrity in L2 speech, but also note exceptions, and in addition have some design drawbacks.

Zsiga (2003) examined gestural overlap at word boundaries in Russian-English and English-Russian bilinguals. This study found, however, that English learners of Russian did not use the native pattern of overlap in the L2. They did not pronounce Boris Yeltsin as Borish Yeltsin, (comparable to English miss you as mish you) at least when speaking Russian. For the Russian speakers, however, the non-overlapped pattern typical of most clusters in the L1 did persist into their L2 English: Russian speakers tend to say hi[t^h] parts. The preference of both English and Russian speakers for a lagged pattern in L2, despite their different L1 patterns, suggests support for the Word Integrity hypothesis. Zsiga (2003) found one exception, however. Despite their general pattern of lag between consonants, Russian speakers produced overlap in front-to-back (e.g. [p#k]) sequences in both L1 and L2. Thus Zsiga (2003) concludes that while Word Integrity may be a preferred or default coordination pattern for L2 speech, it is not an inviolable constraint, and the pattern can be overridden by other, perhaps language-specific, preferences, such as different preferred coordination for front-to-back vs. back-to-front clusters (Goldstein et al. 2007, in press).

A drawback of the Zsiga (2003) study, however, is that a general preference for Word Integrity could not be definitively disentangled from the language-specific pattern for Russian. The English speakers of Russian may have been applying Word Integrity, or they may have been attempting to mimic native Russian (albeit overgeneralizing in some cases). The Russian speakers of English may have been applying Word Integrity to back-to-front sequences, or they may have been applying the pattern from their L1. Thus the present study turns to English and Korean, both languages where overlap and assimilation (though of different kinds) is hypothesized to be the preferred L1 pattern. If Word Integrity is found in the speech of Korean learners of English, it must be due to emergence of a pattern specific to interlanguage speech, not due to influence of the L1 or semi-successful acquisition of the L2.

Lléo & Vogel (2004) studied Spanish learners of German. They argue that these two languages differ in their basic prosodic structures: Spanish phonology tends to 'obscure the edges of the smaller phonological constituents' (p. 85) through the application of external sandhi processes, while such processes are absent from German, keeping prosodic constituents more clearly delineated. While Lléo & Vogel do not adopt an Articulatory Phonology perspective, the distinction invites comparison between languages with and without gestural overlap at word boundaries. Lléo & Vogel specifically test for influence from Spanish to German of two external sandhi processes, spirantization and nasal assimilation. Examples are shown in (9):

(9) Spanish spirantization and nasal place assimilation at word boundaries (Lléo & Vogel 2004)

a. Spirantization

niebla [ð]ensa	thick fog	son [d]ensos	they are thick
cuatro [ɣ]atos	four cats	son [g]atos	they are cats

b. Nasal place assimilation

tiene[m p]ocos libros	they have few books
tiene[n d]iez libros	they have 10 books
tiene[ŋ g]randes ideas	they have great ideas

In a reading task, spirantization applied in (approximately) 41% of eligible sequences within a word, but only 13% of eligible sequences over a word boundary (data from Figure 3B, p. 96). The results for nasal place assimilation were slightly different. Nasal place assimilation within words was not tested, as assimilation applies obligatorily to all within-word clusters in both German and Spanish, and is indicated orthographically. Instead, application across two different prosodic boundaries was tested: at the boundary of two words within a phonological phrase, and two words across an intonational phrase boundary. At the word boundary within a phrase, nasal place assimilation was found to apply in 35% of cases. Across the intonational phrase boundary, assimilation applied in only 11% of cases (data from Figure 4B, p. 97). These results support the Word Integrity hypothesis in the conclusion that the greater the boundary strength, the less likely it is that external sandhi will take place. However, the difference between rates of spirantization and nasal assimilation at word boundaries within a phrase remains unexplained. It may be that nasal assimilation applied at a higher rate because this process is also optional in German. Another problem may be perceptibility: nasal place distinctions in pre-consonantal position are notoriously difficult to perceive (Steriade 2001). Differences in perceptibility may have affected the performance of the speakers, in that they may have been more aware of spirantization, which is easier to hear, and thus worked harder to avoid it. Perceptibility may also have been an issue for the transcribers who marked each utterance as correct or incorrect: it may have been difficult to decide whether place assimilation had applied or not. Transcription would be especially problematic if place assimilation was partial and gradient, the result of gestural overlap between words, as predicted by Articulatory Phonology. Again, the need for instrumental study is supported.

Another point worth noting is that Lléo & Vogel found an effect of level of instruction: intermediate-level students were significantly more likely than either beginning- or advanced-level students to apply both spirantization and nasal place assimilation in German. One interpretation of this finding is that intermediate-level speakers may be advanced enough to plan utterances in units larger than the word, but not advanced enough to coordinate juncture in the way typical of the target language. Cebrian (2000), on the other hand, found no effect of level of instruction. Possible differences due to level of instruction are tested in the present study.

As a whole, studies that have investigated the persistence of L1 external sandhi in L2 have found that it applies infrequently, generally supporting the hypothesis that learners will not coordinate articulations across word boundaries. An exception has been studies of Korean learners of English. Korean has numerous examples of processes of external sandhi (Kim-Renaud 1991, Stuart & Shin 1999) including those listed in (10). (The formalism is intended to be simply descriptive.)

(10) Some processes of external sandhi in Korean

a. nasalization: [-nas] → [+nas] / ___ [+nas]

b. lateralization: /n/ → [l]/ __ [l]

c. spirantization: /t/ → [ʃ] / __ [i]

d. voicing: lax [p,t,k] → [+voice] / [+sonorant] ____ [+sonorant]

Several previous studies (Chu & Park 1978, Kim 2000, Zsiga & Kim 2005, Park 2005) have found that these external sandhi processes do persist from Korean to Korean-accented English. For example, Park (2005), in a study of 32 intermediate-level Korean speakers of English, found that nasalization across a word boundary (that is, phrases like Bob noticed pronounced as Bo[m] noticed) applied in 42% of eligible sequences (p. 64). Within the subset of voiceless stop followed by homorganic nasal (e.g. group marriage), 59% of sequences were nasalized (p. 65). Kim (2000) found a somewhat lower rate of nasalization (16% for voiced stops and 23% for voiceless stops (p. 122)), but found that intersonorant voicing assimilation (e.g., cap is pronounced as cab is) applied 50% of the time. Chu & Park (1978) did not conduct a controlled study, but note many examples of nasalization, spirantization, and lateralization in Korean-accented English, and state that these rules "often transfer" (pp. 12-15).

The lower rate of nasalization in Kim's (2000) study compared to Park's (2005) study may have been due to a number of factors. Kim's study included word-final velars while Park's did not. (Both studies found that rates of nasalization were highest for labial consonants.) Kim's study included advanced-level speakers, while Park's study included only intermediate-level speakers. Direct comparison between speakers in the two studies is difficult, since subject groups were defined in different ways. Park classified speakers as "intermediate" based on raters' judgments of a short speech sample. Kim classified speakers based on placement in ESL classes: those in lower-intermediate and intermediate classes were labeled "less advanced" and those in higher-intermediate and advanced classes were labeled "more advanced." Kim did find a significant difference between the two groups in his study (though both groups nasalized at lower rates than the speakers in Park's study). In Kim's study the rate of nasalization was 28% for the less advanced group (35% for voiceless consonants and 21% for voiced), and 10% for the more advanced group (no effect of voicing). Kim found no effect of level of instruction on the rate of voicing assimilation.

Finally, Kim's study made greater use of phonetic analysis than did Park's. Park's study relied primarily on transcription, making reference to waveforms and spectrograms only 'in the case of unclear segments and/or disagreements' (p. 57). Further, Park does not seem to have distinguished between voicing and nasalization in pre-nasal obstruents. Park placed his sequences into only four categories: correct (English-like), nasalized, epenthesis, and 'other' (which included frication and de-nasalization but not voicing). Park further notes (p. 62) that 'obvious voicing' in the spectrogram or waveform (with or without nasal formants) was taken to indicate the presence of nasalization. Kim (2000) counted pre-nasal voiced stops as a separate category. Based on these descriptions, it may be that Park was counting as nasalized cases that Kim was counting as voiced.

In summary, previous research has found that carry-over of external sandhi from L1 to L2 is possible, but is less common than carry-over of word-internal alternations. Studies looking at different languages and processes have found very different rates of persistent errors: as low as 13% for spirantization in Spanish German, and as high as 59% for nasalization in Korean English. It is not clear whether such differences are due to the nature of the processes studied, the languages targeted, or aspects of experimental design. Some studies (Lléo & Vogel 2004, Kim 2000) have found an effect of level of instruction, such that intermediate-level speakers are more likely than advanced-level speakers to carry over external sandhi; but other studies (Cebrian 2000, Zsiga 2003) did not find a proficiency effect. Kim (2000) found a proficiency effect for nasalization but not voicing.

Most previous studies have relied on transcription rather than instrumental analysis, so the effects of partial vs. complete assimilation have not been controlled for (specifically, the possibility of categorial perception of a gradient articulation), and may account for some of the differences between studies. For example, Cebrian notes that the difference in rate of application he found by manner of articulation may have been due to the fact that the same percentage of voicing is more perceptible in stops vs. fricatives: a partially voiced stop sounds voiced, but a partially voiceless fricative does not. Instrumental analysis is needed to diagnose the effects of gradient gestural overlap and categorical gestural reorganization.

Thus, the present study proposes an acoustic phonetic analysis of nasal assimilation at word boundaries in Korean and Korean-accented English. The study is divided into two experiments. Experiment 1 examines acoustic data on the realization of obstruent#nasal sequences in Korean, American English, and Korean-accented English, with the goals of determining the prevalence of cross-word nasalization, and considering whether the process, when it applies, is best characterized as planned phonological assimilation (2 or 7 above) or as the unintended consequence of a more general pattern of gestural organization at word boundaries (6 above). Experiment 2 investigates whether the higher-than-expected incidence of cross-word assimilation that has been reported for Korean speakers of English is related to a high degree of gestural overlap at word boundaries in general for these speakers. Overall, the two experiments provide data that allow for testing of different models of external sandhi and speech planning in both L1 and L2.

4. Experiment I: Obstruent#Nasal Sequences in Korean, American English, and Korean-accented English

4.1. Hypotheses. Five hypotheses are tested in Experiment I. The first four address the occurrence of cross-word nasalization; the fifth addresses its phonetic characteristics.

Hypothesis 1: Consistent with descriptions in the literature (e.g. Kim-Renaud 1991), Korean speakers speaking their native language will nasalize word-final voiceless stops preceding a word-initial nasal.

/kimpap mekta/ sushi eat → [kimpam mekta]

Hypothesis 2: Consistent with descriptions in the literature (e.g. Jones 1956, Pierrehumbert 1995, Cohn 1993, Huffman 2005), English speakers speaking their native language will produce word-final voiceless stops preceding a word-initial nasal as either voiceless unreleased, glottalized, or glottal stops.

/stap mæt/ stop Matt → [stap˺ mæt] or [stapˀ mæt]

/eɪt maɪn/ ate mine → [eɪt˺ maɪn] or [eɪtˀ maɪn] or [eɪʔ maɪn]

Hypothesis 3: Consistent with previous studies of Korean speakers of English (e.g. Kim 2000, Park 2005), but contra the Word Integrity hypothesis (Cebrian 2000), Korean speakers will carry over obstruent nasalization to L2 English in at least 30% to

40% of cases. (The average percent nasalization of the Kim and Park studies would be 33%).

/stap mæt/ <u>stop Matt</u> → [stam mæt]

Hypothesis 4: Consistent with previous studies that have found an effect of level of instruction on the rate of application of external sandhi in L2 (Kim 2000, Lléo & Vogel 2004), intermediate-level speakers of English will carry over nasalization at a higher rate than advanced-level speakers.

Hypothesis 5a: If nasalization is a categorical association of [nasal] with a consonant that is underlyingly [-voice, -sonorant] (as in (2) or (7) above), then obstruent#nasal sequences will fall clearly into two categories: either a derived nasal sequence that is indistinguishable from an underlying nasal sequence (nasal assimilation applies), or an unaltered sequence of voiceless stop coda followed by nasal consonant onset (nasal assimilation does not apply).

Hypothesis 5b: If nasalization is due to gestural overlap rather than creation of a new categorical assimilation (as in (6) above), the distribution of nasalization will not be bimodal. Rather, different degrees of overlap in the coordination of oral, laryngeal, and velic gestures will give rise to multiple intermediate realizations with partial nasalization and partial voicing.

4.2. Participants. Twelve speakers of Seoul Korean, residing in Washington, D.C. at the time of the study, participated. Some details of the participants' English language background are reported in Table 1. All speakers received English-language instruction in Korea, beginning in late elementary or middle school, and continuing, for some participants, through college or graduate study.

Five speakers were classified as advanced English speakers. At the time of the recording session, these speakers were enrolled in a graduate program at Georgetown University, but were not currently taking ESL classes. They described their own proficiency as 'advanced' (K1, K3, K4, K5) or 'near-native' (K2). Four of the five had completed an undergraduate or graduate degree at an American university, and had lived in the U.S. for four to six years. The other advanced speaker (K5) had attended university in Korea, and had been in the U.S. for less than two years.

Seven speakers were classified as intermediate. While each reported years of English instruction in Korea, all seven were enrolled in intermediate-level ESL classes at Georgetown University, with level determined via placement testing, and they rated their own proficiency as 'intermediate' (K6, K9, K10, K11, K12) or 'high intermediate' (K7, K8). Each had lived in the U.S. for one year or less. The difference between advanced and intermediate speakers in length of instruction and length of U.S. residence was significant by t-test ($t(10) = 3.03$, $p = .013$ and $t(10) = 5.72$, $p < .01$ respectively), but the difference in age at first instruction was not ($t(10) = .737$, n.s.).

Table 1 about here

In addition to the Korean-speaking participants, three native speakers of English participated as controls. All three grew up in the Northeast U.S. (Connecticut, Long Island N.Y. and Pennsylvania), and at the time of the study had been residents of the greater Washington D.C. area for fifteen years or more.

4.3. Materials and recording procedures. Materials for the study consisted of two-word phrases in Korean and in English. To keep syntactic structure as consistent as possible while respecting the basic word order restrictions of the two languages (SOV in Korean and SVO in English), phrases were always Object#Verb in Korean and Verb#Object in English. For English, there were two different phrases for each cluster, with differing vowel qualities. Phrases to test for nasalization (Table 2) consisted of word-final [p, t, k, n] followed by word-initial [m, n] in each language. The obstruent#nasal sequences provide the appropriate environment for nasalization to apply, while nasal#nasal sequences are included as controls. The materials for Experiment II were recorded in the same session. They are discussed in section 5.3 below.

Table 2 about here

Each two-word phrase was placed in a sentence, which was printed on a 4x6 card. Korean sentences were written in Korean orthography, and the English sentences in English orthography. Cards were randomized for each participant, mixing together the phrases targeting nasalization (Experiment 1: Table 2) and those targeting obstruent sequences (Experiment 2: see Table 7 below), but keeping the two languages separate.

The Korean-speaking participants were recorded directly to disk in a sound-treated room at Georgetown University in Washington, D.C. Recording was done by a Korean-speaking research assistant, who gave instructions in Korean. Each speaker read each sentence three times in succession. These 12 participants read the set of Korean sentences first, then the set of English sentences.

The English-speaking participants were recorded directly to disk in a quiet room. Recording was done by the author. They read only the sentences containing the English phrases in Table 2, repeating each three times.

4.4. Analysis. All target phrases were transcribed, independently, by two researchers trained in phonetics. For the Korean phrases, one transcriber was a native speaker of Korean; the other transcribers were native speakers of English. Agreement between the two transcribers was 100% on the transcription of the native Korean phrases, and 83% on the transcription of the Korean English phrases. (In 9% of the Korean English tokens transcribers differed on voiced vs. voiceless; in 6% they differed in voiced vs. nasal; in 2% of tokens one transcribed a voiceless stop and the other a nasal stop.) Disagreements were resolved by consensus.

The following acoustic measures were made for each of the target obstruent#nasal and nasal#nasal sequences. All measures were based on waveform and spectrographic analysis using Praat (Boersma & Weenik 2007). First, total cluster duration was measured, from the end of the preceding vowel to release for the following vowel. A release burst, if any, between the two consonants was included in the total duration, but release into the following vowel was not included. Within this total cluster duration, the following subparts were marked:

1. Duration of (oral) closure voicing. Voicing was defined by low-frequency pulses visible in waveform and spectrogram, but no higher-level formant structure.
2. When present, closure voicing was characterized as modal (smooth, regular pulses) or glottalized (longer, irregular pulses), based on visual inspection of the waveform and spectrogram, following Huffman (2005) and Redi & Shattuck-Huffnagel (2001).
3. Duration of voiceless closure.
4. Duration of (inter-consonantal) release burst.

5. Duration of nasal closure. Nasality was distinguished from oral voicing by increased amplitude, the presence of higher-level formant structure, and corresponding increased complexity in the waveform.

Depending on the acoustic realization of the cluster, any of measurements 1 through 5 might equal zero. Some tokens had a portion of voiceless closure and some did not, for example, and only a few tokens had a release burst. In all cases, however, durations of voiced closure + voiceless closure + release burst + nasal closure always summed to total cluster duration.

Cases where there was a stumble or pause (operationalized as more than 350 ms of silence) between the two words were excluded from further analysis. A total of 16 of 420 tokens (4%) of obstruent#nasal sequences in Korean English were excluded due to pausing; no pauses at the crucial juncture occurred in native Korean or native English.

4.5. Results for Hypotheses 1 and 2: Native Speakers. Hypothesis 1 was confirmed. Consistent with previous descriptions of Korean, the Korean speakers pronounced 93% of the obstruent#nasal sequences in their native language as fully nasal (201 of 216). Nasal formants, and amplitude consistent with a sonorant articulation, were evident throughout the closure. An example sequence is shown in Figure 1.

Figure 1 about here

In each case of nasal assimilation, underlying place of articulation was preserved, e.g. /t#m/ was pronounced as [nm], not [mm]. Note, in Figure 1, that the formant transitions into the nasal are indicative of an alveolar, and those out of the nasal are indicative of a labial. (Son et al. 2007 also found no application of place assimilation across word boundaries in native Korean.) Of the 15 stop tokens in the native Korean data that occurred in the nasal environment but were not nasalized, 8 were fully voiced (e.g. /t#m/ pronounced as [dm]), 4 were fully voiceless (e.g. /t#m/ pronounced as [tm]), and 3 were partially voiced, evidencing modal voicing for about half the duration of the oral closure. There was no pattern evident in the distribution of these non-nasalized tokens: Seven of the 12 speakers produced one to four non-nasalized tokens each, and all three final consonants and both initial nasals were represented.

Table 3 gives the durations of the derived and underlying nasal sequences that can be matched for place of articulation. Two-tailed t-tests, pooling over subjects, found no significant difference in duration between derived and underlying /n#m/ or derived and underlying /n#n/ in native Korean ($t(68) = 1.1$, $p = .279$ and $t(69) = 1.5$, $p = .140$, respectively).

Table 3 about here

Hypothesis 2 was also confirmed. There was more token-to-token and speaker-to-speaker variation for the English speakers than for the Korean speakers, but all word-final, pre-nasal obstruents in native English were produced as one of three allophones: either voiceless with audible release (e.g. /t^hn/), voiceless with no release (e.g. [t̚n], [$t^ʔ$n] or possibly [ʔn]) or creaky voiced [t̰n]. Distribution of the allophones for each speaker is shown in Table 4.

Table 4 about here

Results are consistent with those reported in previous literature on stop#nasal sequences in English, specifically Cohn (1993) and Huffman (2005). Voiceless unreleased was by far the most common allophone. There was some tendency for perseverative voicing from the vowel into the closure in these tokens, but this lasted only for a small percentage of the oral closure (on average 6.9 % for E1, 4.4% for E2, 9.6% for E3), and was usually glottalized.

Two examples of a native English /t#m/ sequence are shown in Figures 2 and 3. Both are taken from speaker E2's pronunciation of the sentence <u>Jill didn't have any cake of her own, so</u>

she ate mine when I wasn't looking. Figure 2 (repetition 1), shows a long voiceless closure, with glottalization at the end of the preceding vowel. This pattern is similar to Huffman's Figure 1 (2005:336), and Cohn's Figure 5b (1993:62). Figure 3 (repetition 3) shows a short oral closure with creaky voicing throughout, similar to Huffman's Figure 3 (2005:342) and Cohn's Figure 9 (1993:75). This creaky-voiced realization was most common with word-final /t/, and as shown in Table 4, occurred most often with speaker E2. Note also the downturn in F2 at the end of the vowel in both Figures 2 and 3, greater with the shorter closure, indicative of the overlap between C1 and C2 that has been found to be typical of English: movement toward the labial closure (C2) is beginning before closure for C1 has been achieved (Zsiga 1994). Both Cohn (1993, based on measures of oral and nasal airflow) and Moll & Daniloff (1971, based on cinefluorographic films) find that velum opening in C#N sequences in English generally begins during the obstruent closure. Such early opening, however, will not result in nasal resonance until the laryngeal configuration and supra- and sub-glottal pressure differential are such as to allow voicing.

Figures 2 and 3 about here

No instances of Korean-like full nasalization (Figure 1) were found for the English speakers. Cohn (1993) reports the possibility of complete nasalization of the coda obstruent in a /t#n/ sequence in English, but this is always accompanied by glottalization and decreased airflow during the coda portion. (A similar pattern of coda nasalization plus glottalization is reported for German by Kohler (1994).) In the present data set, however, all native English tokens were produced with some period of non-nasal closure, either voiceless or creaky voiced. The length of the non-nasal closure varied considerably, however, ranging from just 23% percent of the total sequence duration, to 76% percent. These native English and native Korean results serve as a baseline to which the Korean English tokens can be compared.

4.6. Results for Hypotheses 3 and 4: Nasalization by Korean Speakers of English.

Hypothesis 3 predicted that obstruent nasalization across word boundaries would carry over from Korean to Korean English, with a predicted rate of at least 30%. This hypothesis was confirmed. Overall, as shown in Figure 4, pre-nasal obstruents were transcribed as nasalized in 32% of tokens in Korean English. This rate is higher than the results for Kim (2000) and lower than the results from Park (2005), but is very close to the average of the two studies, and is also consistent with Lléo & Vogel's (2004) finding of nasal place assimilation in 35% of tokens in Spanish German. The rate is higher than Lléo & Vogel's finding of spirantization at 13% in Spanish German, and Cebrian's finding of voicing at 20% in Catalan English. Possible reasons for these differences in rate of application are discussed in sections 5.8 and 6.1.

Figure 4 about here

While nasalization occurs at a much lower rate in Korean-accented English than in Korean, 32% nasalization is higher than predicted by the Word Integrity hypothesis. Another source of evidence against the Word Integrity hypothesis is the low incidence of audibly released final stops (that is, tokens where a release burst, sometimes accompanied by aspiration, was visible in waveform and spectrogram). Only 14% of tokens in this dataset had an audibly-released word-final consonant. An example is shown in Figure 5: the phrase ate mine as produced by Speaker K8.[i] A released final stop is a clear indication of a lag between the articulations of the obstruent and the nasal: a release burst will only be audible if there is a period of time between the closures for the two consonants in sequence. Released final consonants in this context are typical of neither native Korean nor native English (Kim-Renaud 1991, Huffman 2005), but are predicted for L2 by the Word Integrity hypothesis. Zsiga (2003) suggests that release of word-final stops may be the unmarked state for L2 productions; however, such a realization was the minority in these data.

Figure 5 about here

The most common realization (49%) for pre-nasal stops in this dataset was that of an unreleased oral stop. Unreleased stops were also the most common realization for native English speakers. As discussed in Section 4.7 below, however, these unreleased stops were not necessarily English-like. Before moving on to the phonetic realization of these unreleased tokens, which speaks to the question of whether nasalization is categorical or gradient (Hypothesis 5), three further points about the incidence of nasalization must be noted.

First, the stops classified as nasal in Figure 4 were in fact of two types. Most of the nasal tokens (22% of the entire dataset, as indicated by the solid black bar) were full nasals, with nasal formants and sonorant amplitude extending across the entire closure. These full nasals were indistinguishable from underlying nasals. Figure 6 shows an underlying /n#m/ sequence (K12's production of the nasal#nasal sequence in the phrase train Matt), and Figure 7 shows a full nasal derived from underlying /t#m/ (K12's production of the phrase ate mine).

Figures 6 and 7 about here

As with the native Korean data, the durations of underlying and derived full nasal sequences in English, matched for place of articulation, did not differ in duration. (For derived vs. underlying [n#m], $t(83) = .106$, $p = .92$; and for derived vs. underlying [n#n], $t(81) = .221$, $p = .825$.) Table 5 gives mean durations and standard deviations, including only those obstruent#nasal sequences that were realized as fully nasal. (The number of tokens counted for underlying /n#m/ and /n#n/ is less than 72 due to pauses at the crucial juncture in 5 tokens.) That the English sequences are slightly longer and slightly more variable than corresponding sequences in native Korean is not unexpected. The facts that derived and underlying sequences do not differ in duration, and that durations in English are longer than those in Korean supports the analysis of these tokens as having undergone assimilation rather than deletion.[ii]

Table 5 about here

An additional 9% of tokens classed as nasal in Figure 4 (patterned portion of the bar) were transcribed as nasal, and were therefore put into the nasal category, but were acoustically different. These tokens might best be described as pre-oralized. During the consonant closure, there was a short period of oral voicing (between 10% and 30% of closure duration), followed by increasing nasalization. An example is shown in Figure 8. No pre-oralized tokens were found in the native Korean productions. There are similarities, however, between these pre-oralized tokens and some of the native English tokens with short oral closure (Figure 3). An important difference, however, is that the native English tokens showed evidence of glottalization, while the Korean English tokens did not. These pre-oralized tokens are discussed further in section 4.7 below.

Figure 8 about here

The second point to be noted about the incidence of nasalization is that it varied greatly from speaker to speaker. Figure 9 shows the realizations of each speaker. Full nasals (black bar) and pre-oralized nasals (patterned) are distinguished as in Figure 4.

Figure 9 about here

For four speakers, nasalization was the predominant realization (either majority or plurality of tokens transcribed as nasal): 42% for K1 and 75% for K5, both advanced speakers, and 50% for K7 and 89% for K12, both intermediate-level speakers. (Recall that speakers K1 – K5 are advanced and speakers K6 – K12 are intermediate.) For the other speakers, the percentage of tokens nasalized ranged from 0% (K2) to 39% (K9). For three speakers (K4, K6, K8), all the tokens transcribed as nasal were in fact pre-oralized. Speakers also varied considerably in whether or not they produced audibly-released final stops, ranging from 50% for K8 (her plurality realization) to 0% for K3, K5, K7, K11, and K12. There is also a tendency for those speakers who produced the most audibly-released stops to produce the fewest nasals, and vice versa. For most speakers (K2, K3, K4, K6, K9, K10, K11), an unreleased oral stop was the majority or plurality realization.

The category 'other' in Figures 4 and 9 includes mostly pauses (silence of greater than 350 ms between the two target words), 4% of the data, and a few unexpected realizations. In one token, stop Nat was pronounced as sto[b d]at, with voicing throughout but no nasalization. In three cases, a [k#n] sequence was pronounced with a voiceless velar nasal (breathy voicing visible on spectrogram) in place of the [k]. These tokens are discussed further in section 4.8.

While there were large inter-speaker differences, Hypothesis 4 was not supported. There were no significant differences between the group of advanced English speakers and the group of intermediate English speakers. The overall rate of nasalization was 31% for advanced speakers and 32% for intermediate ($t(10) = .071$, $p = .945$). The overall rate of audible release was 14% for advanced speakers, and 15% for intermediate ($t(10) = .155$, $p = .880$). Nor was there any significant correlation between rate of nasalization and age of first English instruction, years of study, nor time in the U.S. It is true that speaker K12, who has the highest rate of nasalization, was also the speaker who began learning English the latest (age 15). But Speaker K4, who began learning English the next-latest, at age 14, has among the lowest rates.

In sum, Hypothesis 3 was supported: nasalization did occur frequently in these data (32% of tokens), but was not as prevalent as in native Korean (93% of tokens). Speakers nasalized at very different rates, ranging from 0% to 89%, suggesting, if nothing else, that speakers are using different strategies of coordination across words. Hypothesis 4 was not supported: there was no significant effect of level of instruction.

4.7. Results for Hypothesis 5: Is nasalization gradient? Hypothesis 5 addresses the question of whether nasalization is a categorical phenomenon, requiring a close examination of the phonetics of these sequences. Hypothesis 5a predicts that a pre-nasal voiceless obstruent is either nasalized ([+nasal, +voice]) or unchanged ([-nasal, -voice]), supporting the analysis of nasalization as a category-changing alternation. Hypothesis 5b predicts gradient nasalization, with variable and intermediate realizations, supporting the analysis of nasalization as variable gestural overlap. Close examination of the data in fact supports both kinds of process in different cases.

There was little or no evidence that nasalization in native Korean (section 4.5) was anything other than a categorical change: in 93% of tokens, nasalization was complete across the closure duration, and underlying and derived nasal sequences were indistinguishable. The few non-nasal tokens appeared to either have simply not undergone the rule, or to have perhaps been subject to the rule of intersonorant voicing instead, but there are too few non-nasal tokens to draw any definitive conclusion. In native English there was evidence of gestural overlap resulting in varying durations of oral closure, but no cases of full nasalization.

On the other hand, as was evident in section 4.6, Korean speakers of English produced a variety of different realizations of pre-nasal stops: audibly-released stop (Figure 5), full nasal

(Figure 7), pre-oralized nasal (Figure 8), and unreleased stop (variations exemplified and discussed below).

It is not enough to show variability, however. There will always be some variability from token to token and speaker to speaker, whether or not a phonological rule has applied. The question is whether the tokens fall into two different categories or not. That is, is the variation normally distributed about a single mode, indicating a single phonological target that may be subject to more or less phonetic undershoot and overshoot? Or is there a bi-modal distribution of the variation, indicating two separate phonological targets? (See Pierrehumbert & Gross 2003, Scobbie 2007 for discussion.)

A reasonable variable to look at for this question is the percent of the cluster duration that is nasalized. Examining percent nasalization rather than duration in ms allows for variation in speaking rate, and probably corresponds most closely to listeners' perceptions. If nasalization is the result of a phonological assimilation (Hypothesis 5a) we expect two different categories: [-nasal]#[+nasal] corresponding to non-application and [+nasal]#[+nasal] corresponding to application. For the first category, nasalization should extend for approximately 50% of the cluster duration. For the second category, nasalization should extend throughout the cluster duration. If nasalization is the result of variation in gestural overlap (Hypothesis 5b), we expect a single distribution, whose mode will differ from speaker to speaker.

Figure 10 graphs, for each speaker, the percent of sequence duration that is nasalized in each English-language token. The x-axis gives the speaker numbers: both Korean L2 and English L1 speakers are included. Each symbol corresponds to a single token. Triangles indicate tokens transcribed as nasal; circles indicate tokens in which an obstruent coda was transcribed. Since the fully nasal tokens are superimposed at 100% nasalized, the numbers at the top of each column indicate how many fully nasal tokens were articulated by the given speaker. The one 0% nasalized token for speaker K1 corresponds to the pronunciation of stop Nat as stob dat.

Figure 10 about here

Tokens with nasalization less than 30% are generally those with an audible release (as in Figure 5): the percentage of nasalization is low because the duration of the release is counted as part of the overall sequence duration. Phrases with a pause (> 350 ms) at the crucial juncture, and underlying nasal#nasal sequences, are not included. The total number of data points for each column is 36.

Figure 10 reveals several points. First, as noted above, complete nasalization occurred only for the Korean speakers, never for the English speakers. The inter-speaker differences among the Korean speakers in number of fully nasal tokens produced are also evident. Further, both the native English speakers and the Korean English speakers were variable in terms of how much of the cluster was actually nasalized. If the fully nasal tokens are excluded, the ranges of variation are in fact quite similar, though slightly larger for Korean English. The mean of the native English distribution is 49%, with a standard deviation of 13%. The mean of the Korean English distribution (excluding the full nasals) is 48%, with a standard deviation of 15%. The means and standard deviations for percent duration nasalized for each speaker are given in Table 6.

Table 6 about here

It is clear from Figure 10 that the distributions for the native English speakers are not bimodal. The tokens for each speaker are distributed about a mean near 50%, exactly what is predicted by variable gestural overlap between the two consonants at a word boundary, but no phonological category change.

What about the Korean English speakers? The pre-oralized nasals (open triangles with nasalization less than 100%), although they were transcribed as nasal#nasal clusters, appear to be just the tails of the [-nasal]#[+nasal] distribution. A few tokens stand out by themselves (see Speakers K6 and K10), but outliers at the tail of a distribution are to be expected. There is no clear evidence that the pre-oralized tokens constitute a separate category from the other oral#nasal sequences: they just have a shorter than average oral portion and longer than average nasal portion. Consistent with the predictions of Articulatory Phonology, if gestural overlap is great enough there can be a perception of category change, even though a remnant of the underlying articulation (here, a portion of the closure that is oral) can be detected by phonetic analysis. It is interesting to note that clusters that are 30% oral and 70% nasal were transcribed as being fully nasalized in Korean English but not in native English. The reason for this difference in perception appears to be due to differences in voicing during the oral portion of the closure, discussed below.

However, the fully nasal tokens do not appear to be just even more extreme examples of gestural overlap. For most of the speakers, there is a gap between the most nasalized pre-oralized token and the group of fully nasal tokens. This is especially evident for Speakers 1, 7, 9, and 11: none has any pre-oralized token that is more than 80% nasalized, but each has a fairly large number of 100% nasalized tokens: a clear second spike in the distribution. For these speakers, it would appear that the full nasals result from a categorically different gestural organization.[iii] In order to produce full nasalization across both consonants at consistent amplitude, the velum must be fully open at the point of consonant closure, and thus the opening movement must be timed to begin during the preceding vowel (Cohn 1993, Krakow 1999). The short oral closure and increasing nasalization of the pre-oralized cases is consistent with the velum opening beginning later, at some point during the consonant sequence (Cohn 1993, Moll & Daniloff 1971).

The conclusion is that two types of nasalization are found in this data, and that both Hypotheses 5a and 5b are supported, though for different tokens. For native Korean and the full nasals in Korean English, there appears to have been a categorical reorganization, such that the velum is timed to be fully open at the time closure for the coda consonant is reached: an oral segment has become [nasal]. But a second type of nasalization is found as well. For the pre-oralized tokens, the actual extent of nasalization is no greater than that found for some native English tokens, with velum opening at some point after consonant closure has been achieved. In these cases, the perception of a categorical change arises, even though a portion of the cluster remains oral.[iv] Timing of the oral and velum gestures in these clusters appears to be very similar in native English and Korean English.

But why do the more overlapped Korean English tokens sound nasalized when the native English tokens do not? The answer probably lies in differences in voicing.

As shown by the graphs in Figures 4 and 9, most stop#nasal sequences in these Korean English data were pronounced with some portion of oral closure and no release burst. The incidence of this realization (including the pre-oralized tokens) was 58% overall, and this was the most common realization for 9 of the 12 speakers. (As noted above, one speaker, K8, produced more released stops than any other realization, and two speakers, K5 and K12, produced mostly full nasals.)

An unreleased oral stop followed by a nasal was also the most common realization for the native English speakers (89%). However, the Korean English productions were not necessarily English-like. A first difference is that the productions of the Korean speakers were only seldom glottalized. Second, the Korean speakers were much more variable in terms of the extent of voicing. Figures 11 – 14 provide examples.

Figures 11, 12, 13, 14 about here

Figure 11 shows a token with an unreleased stop that is almost completely voiceless. This token is very similar to those produced by the native English speakers, though without evident glottalization. Figures 12 and 13 show increasing amounts of voicing, and Figure 14 shows a token where voicing extends throughout the obstruent closure. It is entirely possible that in cases such as that in Figure 14, where voicing extends throughout the closure, that the voicing is abetted, if not caused, by decreased supraglottal pressure as the velum begins to open. The difference in amplitude, and the lack of nasal formants during the first half of the cluster, however, indicate that the velum is not open enough to produce strong nasal resonance, and the perception is that of a voiced stop. The phrase in Figure 14 sounds like keeb Matt.

Further, voicing of underlyingly voiceless stops in Korean English is not restricted to the pre-nasal context. Voicing in other contexts was not systematically measured, but voicing of obstruents in post-vocalic position, especially when the following consonant was also a sonorant, was often noted. Figure 15 shows a typical example, from Speaker K3. The test sentence was Maybe that ticket will stop Nan from speeding. The target sequence /p#n/ was pronounced by this speakers as [b#n], but, as shown in Figure 15, there was also partial voicing on the syllable-initial /t/ and /k/ in ticket, and complete voicing of the word-final /t/ preceding /w/.

Figure 15 about here

Figure 16, using a format parallel to that of Figure 10, graphs the duration of oral voicing for each token. Duration in ms is used here rather than the percentage of oral closure that is voiced, because of the variation in length of oral closure: from less than 20 ms to more than 150 ms. With that range of variation, 20 ms of voicing would correspond to anywhere from 100% to 13% of the oral closure duration. Only tokens that had some measurable oral closure and that had no audible release are included; thus the total number of tokens for each speaker differs depending on how often they used these different realizations. In Figure 16, tokens where voicing was clearly glottalized are symbolized with +. Otherwise, as in Figure 10, tokens are shown as open triangles (pre-oralized nasals) or open circles. Also parallel to Figure 10, the number of tokens for each speaker with no measurable voicing (superimposed at 0%), is indicated at the bottom of each column.

Figure 16 about here

Unlike the results for percent nasalized (Figure 10), there is a clear difference in patterns of voicing for Korean English and native English speakers. There are differences in the extent of voicing, and in the use of glottalization. For the native English speakers, 56% of unreleased tokens showed no oral closure voicing at all, 14% had obvious creaky voicing, and 28% had a short period of modal voicing. Mean duration of modal closure voicing (when present) for the native English speakers is 16 ms. Especially for speakers E1 and E2, longer periods of voicing are glottalized. These tokens (+'s in Figure 16) presumably represent a weak glottal constriction, sufficient to affect the quality of voicing, but not to cut off airflow (see Huffman 2005:344—345). Mean duration of glottalized voicing is 43 ms.

For the Korean English speakers, only 9% of unreleased tokens were completely voiceless, 6% had obvious creaky voicing, and 85% of had a measurable period of modal voicing. Of the 13 creaky-voiced tokens, 6 were from one speaker, K11, whose voice quality was somewhat creaky overall. As can be seen in Figure 16, voicing for the Korean speakers was more variable than for the English speakers. There are four speakers (K2, K5, K9, K10) whose range of variation was similar to that of the native English speakers (at least to E3), but for the other speakers voicing extends for much longer, and this longer period of voicing is modal, not glottalized. Mean duration of measurable modal phonation for the Korean speakers is 42 ms into the oral closure, nearly 3 times longer than the modal voicing duration for the English speakers. The difference in voicing duration between the Korean and English speakers is highly significant by t-test ($t(279) = 12.12$, $p < .001$).

For the completely voiceless stops in these data, the acoustic analysis undertaken here cannot determine whether voicelessness is the result of glottal adduction or abduction. The overall pattern, however, is consistent with other reports in the literature that point to glottal adduction as the usual pattern for English speakers, and glottal abduction as the usual pattern for Korean speakers. While glottalization is a common feature of voiceless coda stops in native English (Huffman 2005), in Korean glottalization is associated with the realization of fortis onset stops, not lax coda stops (Cho, Jun & Ladefoged 2002, Hirose, Lee, & Ushijima 1974). The variable voicing of intersonorant obstruents in the Korean English tokens in this dataset is consistent with descriptions of the realization of intersonorant lax stops in native Korean (Cho et al. 2002, S.-A. Jun 1995, Silva 1992). These sources provide phonetic evidence that lax coda stops in Korean are realized with a small glottal opening that is subject to varying degrees of reduction, which often results in voicing continuing throughout the oral closure when the stop is surrounded by sonorants. In native Korean, of course, this voicing is not reported in the pre-nasal context, because nasalization occurs there, although a few tokens that were voiced rather than nasalized were found in the native Korean data recorded for this experiment (section 4.5 above). In some cases (more so for some speakers than others) the Korean speakers were able to suppress the tendency toward intersonorant voicing, at least in some cases through the use of glottal adduction. In many cases, however, even when nasal assimilation did not carry over from L1 to L2, intersonorant voicing did.

4.8. Summary and Discussion of Results, Experiment I. Hypotheses 1 and 2 were supported, establishing patterns typical of each L1. In native Korean, there is full nasalization of pre-nasal obstruents across a word boundary. Full nasalization means that sonorant amplitude and nasal resonance are maintained throughout the consonant closure. In order for this pattern of resonance to be achieved, the velum must be fully open at the point of consonant closure, and therefore the opening gesture must begin during the preceding vowel. Since the trigger for nasalization is the onset of the second word, full nasalization clearly indicates that the two words are planned as a unit: the nasal gesture lexically associated with the second word begins in the middle of the first word. This extension of nasalization occurs without concomitant extension of the oral closure constriction for C2: there is nasal assimilation, but not place assimilation, in these tokens. There were a few tokens (7%) where nasalization did not apply, but there was otherwise very little variation in native Korean. Derived and underlying nasals were not acoustically distinct. Each of these facts argues for a categorical alternation in native Korean: nasalization of the word-final stop is planned at the phonological level (as represented in (2) or (7)).

In native English, on the other hand, there were no cases of full nasalization. Instead, the data is consistent with variable, sometimes extensive, overlap between an obstruent coda (often glottalized) and a nasal onset. This timing pattern is consistent with the graph in (5).

For Korean English, the data is much more variable. Hypothesis 3 was supported overall: nasalization was perceived to have applied in 32% of tokens. The high incidence of nasalization and voicing, and low incidence of released final stops, do not support the Word Integrity hypothesis. In many cases, the articulation of the two words in the verb#object phrase did appear to be planned as a unit.

Hypothesis 4 was not supported: there was no overall group difference between the advanced and intermediate speakers in the rate of nasalization.

While there was no overall group difference, there a great deal of inter-speaker variation. Rates of nasalization for each speaker ranged from 0% of tokens to 89% of tokens. A wide variety of realizations were attested, as illustrated by Figures 5, 7, 8 and 11—14, and by the variation graphed in Figures 9, 10, and 16. Each speaker had a preferred realization: speakers K5 and K12 produced the majority of clusters as fully nasal. K8 had a majority of tokens with audible release. The other speakers produced mostly unreleased final stops, with varying degrees of voicing and nasality, some of which were transcribed as nasal, some as voiced, some as voiceless. The graphs reveal intra-speaker variation as well. No speaker produced all tokens in the same way. It was not in fact unusual for a single speaker to produce the same phrase three different ways in three consecutive repetitions. This high degree of inter-speaker variability may be a cause of variation across studies on L2 external sandhi: results are highly dependent on the selection of the set of speakers for the subject pool. One would also expect variability due to the task and other variables, which also differ from study to study.

At a more theoretical level, the variability suggests that there are different word-to-word timing strategies available to L2 speakers. Speaker K8 appears to be following the Word Integrity hypothesis most of the time. This speaker usually produced final stops with an audible release and seldom evidenced assimilation across the word boundary, consistent with a strategy of planning each word as an independent unit. Since this pattern is typical of neither native Korean nor native English, it would appear to be emerging as a specific strategy for simplifying articulatory planning in a non-native language, consistent with the predictions of Cebrian (2000) and Zsiga (2003). The pattern seems to be typical of only one speaker among the 12, however. Speaker K2 also had a high percentage of audibly-released final stops (42%), and evidenced no nasalization across the word boundary (0%), although the majority realization for this speaker was voiceless unreleased (58%).

Two speakers, K5 and K12, produced mostly full nasals in their L2 productions. K1 and K7 also produced a high proportion of fully nasal tokens (31% and 42% respectively), and 8 of the 12 speakers produced full nasals at least some of the time. The categorical nature of the change in these tokens is consistent with a planned association of nasality with the coda consonant: that is, carryover of Korean phonological planning from the L1 to the L2. The existence of full nasal tokens in the L2 data supports Hypothesis 5a, and representations (2) or (7) above.

Most of the Korean English clusters, however, were produced with no audible release and some period of oral closure, sometimes very long, sometimes very short, and with varying durations of voicing. The tokens with a very short oral portion were often transcribed as nasals, and the acoustics of these pre-oralized tokens suggest that velum opening begins during the consonant closure rather than during the vowel. This pattern of partial nasalization is consistent with gestural overlap (close transition) between the first and second word, and thus the planning of the two words as a prosodic unit, but not necessarily a planning decision to extend the nasal gesture. Overall, the percentage of the closure that was nasalized in Korean English was in fact very similar to the percentage of nasalization in the native English tokens, suggesting a similar pattern of overlap between an obstruent coda and nasal onset. The existence of varying degrees of nasalization, particularly the pre-oralized tokens, supports Hypothesis 5b, and the representation in (6).

The Korean speakers differed from the English speakers, however, in the laryngeal gestures employed in these sequences. The English speakers tended to varying degrees of glottal adduction in their coda obstruents, leading to glottal stops or creaky voice, while the Korean speakers tended to varying degrees of glottal abduction in their coda obstruents, often leading to modal voicing extending throughout the closure, similar to the patterns found in native Korean intersonorant lax stops. Thus even when speakers did not apply the typical Korean timing and realization of nasal gestures, they did apply the typical Korean timing and realization of glottal gestures.

Overall, then, making a smooth articulatory transition from voiceless obstruent to nasal sonorant is complicated, involving precise coordination of oral constriction, velum opening, and laryngeal adduction or abduction. When the Korean L2 speakers produce the two oral constrictions in close transition, they vary in how they coordinate the velic and glottal gestures. Further evidence of variation in the coordination of these gestures is found in the unusual realizations that occurred from time to time: a completely denasalized sequence (stob dat for stop Nat), where the velum opening gesture was missing (or so reduced that no nasal resonance resulted), and a few cases of voiceless nasals, where velum opening and glottal opening coincided.

The data from Experiment I thus support an analysis of three qualitatively different timing patterns being used to implement obstruent#nasal sequences in Korean English: 1) Word Integrity, with the two words being executed as separate units, 2) a categorical nasalization planned at the phonological level, comparable to the categorical alternation in native Korean, and 3) consonant overlap at word boundaries similar to that seen in native English, which sometimes results in the perception of nasal assimilation.

From these findings, several predictions follow. It is hypothesized (following the tenets of Articulatory Phonology), that the gradient nasalization found in many tokens in Korean English follows from a pattern of gestural overlap at word boundaries. In fact, any cross-word assimilation, whether categorical or gradient, follows from some degree of coordination between the two words: representation (6) (gradient nasalization) assumes a general pattern of close transition between consonants, and representation (7) (categorical assimilation) assumes a general pattern of close transition, plus planned extension of the nasal gesture.[v] It is predicted that close transition in obstruent#nasal clusters is not specific to this environment, but follows from a more general pattern of between-word temporal coordination: a pattern of close transition should also be evident in other consonant clusters, not just those in which C2 is a nasal. It is further predicted, based on previous studies and on the results reported above, that this pattern of consonant overlap should be evident in both L1 Korean and L2 English.

It is also predicted that at least some of the inter-subject variability in nasalization, voicing, and word-final audible release reported above is due to each speaker having a typical (though not completely consistent) temporal pattern in L2 speech, a pattern that will be evident in other clusters, not just obstruent#nasal clusters. Speakers who showed the most audible releases in obstruent#nasal sequences will show a similar pattern in obstruent#obstruent sequences. Conversely, those who produced nasal clusters in close transition will also produce obstruent clusters in close transition. Finally, if gradient nasalization is the result of increased gestural overlap, then the degree of gradient nasalization should correlate with the degree of overlap in non-nasal clusters. Categorical assimilation should be associated with a close transition (no audible release), but not necessarily with more extensive overlap.

Experiment II tests these predictions explicitly, by measuring gestural overlap in obstruent#obstruent sequences for the same set of Korean English speakers.

5. EXPERIMENT II: Gestural overlap in Korean and Korean-accented English

5.1. Hypotheses. Three hypotheses are tested in Experiment II.

> Hypothesis 6: Consistent with previous articulatory studies (Kochetov et al. 2007) Korean speakers speaking their L1 will produce close transitions between word-final and word-initial obstruents, with evidence of between-word gestural overlap as measured by duration ratio (Zsiga 2000, 2003, defined below).

> Hypothesis 7: Consistent with the findings reported above, but contra the Word Integrity hypothesis, general patterns of consonant overlap in the L1 will carry over to the L2: duration ratios in native Korean and Korean English will be similar.

> Hypothesis 8: The L2 timing patterns evidenced by different speakers in obstruent#nasal sequences will also emerge in obstruent#obstruent sequences.

8a: The speakers with the most audible releases in obstruent#nasal clusters (K2 and K8) will evidence a similar pattern of lag and audible release in obstruent#obstruent clusters.

8b: Those speakers who typically produced categorical nasalization (K12 and K5) will typically produce obstruent#obstruent clusters in close transition (no lag, no audible release).

8c: For speakers who typically produced clusters in close transition with gradient nasalization, there should be a correlation between the degree of nasalization and the amount of overlap in obstruent clusters as measured by duration ratio.

5.2. Participants and Recording Procedures. The Korean speakers who participated in Experiment I also participated in Experiment II. As noted above, the test phrases for Experiments I and II were recorded in the same session, and were interspersed. The native English speakers did not participate in Experiment II. Similar data from a previously-analyzed group of 6 English speakers and 6 Russian speakers (Zsiga 2000, 2003) are used for comparison.

5.3. Materials. For Experiment II, phrases were designed to test for general patterns of gestural overlap at word boundaries. These phrases (Table 7) are modeled on the pattern developed in Zsiga (2000, 2003), with modifications to remove extraneous factors such as consonant clusters that might prove problematic for Korean speakers. The phrases are designed so that the closure duration in a C1#C2 obstruent cluster can be compared to the sum of the durations of word-final C1 and word-initial C2 surrounded by more open articulations. (Details for computing this measure of duration ratio are given in section 5.4 below.) The clusters consist of word-final [p, t, k] followed by word-initial [p^h, t^h, k^h]. As was the case in Experiment I, consonants were chosen to sample different places of articulation while respecting the phonotactic constraints of both languages. (Additional Korean sentences, with fortis and lenis initial stops, were also recorded, but are not analyzed here.) Singleton measurements are taken of word-initial [p^h, t^h, k^h] preceded by a word ending in a vowel, and word-final [p, t, k] followed by a word beginning with [h]. Word-initial [h] is used rather than a vowel to avoid the resyllabification or reduction likely to occur in intervocalic contexts.

5.4. Phonetic analysis of obstruent clusters and singletons. All measures were based on waveform and spectrographic analysis using Praat (Boersma & Weenik 2007). For the obstruent clusters and singletons, closure duration was measured, from the end of the preceding vowel to release for the following vowel or [h]. A release burst, if any, between the two consonants was included in the total duration, but release into the following vowel or [h] was not included. Cases where there was a stumble or pause (> 350 ms of silence) between the two words were excluded from further analysis. A total of 21 cases (2% of the obstruent phrases) in English were excluded due to pausing; no pauses at the crucial juncture occurred in Korean.

In order to quantify the degree of overlap in a cluster, an average duration ratio was computed for each cluster in each language for each speaker. The formula for duration ratio (Zsiga 2000, 2003) is given in (11).

(11) duration ratio for C1#C2 =

$$\frac{\text{mean closure duration [C1\#C2]}}{\text{mean closure duration [C1\#h] + mean closure duration [V\#C2]}}$$

Considering a ratio rather than raw duration in milliseconds controls for inherent differences due to place of articulation, and inter-speaker differences in speaking rate. Duration ratio less than 1 is interpreted to indicate overlap between the consonants: the cluster is shorter than the sum of its parts. Duration ratio greater than 1 is interpreted as lag between the consonants: the cluster is longer than the sum of its parts.

Figures 17 and 18 illustrate the computation of duration ratio for a [t#k^h] cluster. Figure17 shows tokens from the English pronunciations of speaker K8. The figure shows 400 ms waveforms extracted from the utterances ate cake, ate ham, and saw Ken. The lines above the waveforms show how closure duration was measured: in this case 220 ms for the [t#k^h] cluster (including the intervening release), 84 ms for word-final [t] preceding [h], and 98 ms for word-initial [k^h] following a vowel. Duration ratio is then computed as the duration of the sequence divided by the sum of the closure durations of the word-final and word-initial consonants: 220 / (84 + 98) = 1.21. As can be seen in the diagram below the waveforms, the cluster is 121% as long as the sum of its parts, and the lag between the two consonant closures is evident. For illustrative purposes in these figures, only one token is shown for each sequence, but in the actual analysis the mean duration of all tokens of a given sequence for each speaker (usually 6) was used to compute the ratio.

Figures 17 and 18 about here

In tokens like that in Figure 17, with an audible intervening release, the relationship between the two consonants in the sequence is obvious. This is not the case when there is no intervening release, as illustrated in Figure 18. This figure shows tokens of the same three utterances, ate cake, ate ham, and saw Ken, from the English pronunciations of K9. Again, each extract is 400 ms long. In this case, there is one long closure for the [t#k^h] sequence, and visual inspection of the waveform does not reveal which part of the closure can be attributed to the [t] and which part to the [k^h]. In the measure of duration ratio, comparison to the durations of [t] and [k^h] between more open articulations is used to estimate their contributions to the cluster. In Figure 18, durations are measured, and duration ratio computed in the same way as in Figure 17. In this case, however, the cluster is found to be only 73% as long as the sum of its parts. As shown in the diagram, overlap between the two articulations can be inferred.

An acoustic measure, duration ratio cannot provide the details of articulatory movements that can be discovered with direct articulatory measures such as electro-magnetic articulography (EMA) (e.g. as was used by Kochetov et al. 2007), although analyses using duration ratio have been found to be consistent with the results of articulatory studies (see Zsiga 2003). Byrd et al. (2008), analyzing EMA data on consonant sequences in English, found that syllable position and adjacency to boundaries of varying strength significantly affected the shape and time course of consonant closing gestures, but that the identity of preceding and following segments across the boundary did not. Direct articulatory measures, however, are expensive, time-consuming, and intrusive enough to inhibit natural speech, especially for language learners who are not confident in their speech skills under the best of circumstances. For the present study in particular, duration ratio cannot differentiate effects of gestural overlap and gestural reduction. For duration ratios less than one, it may be that reduction of the coda consonant is playing a role in addition to that of overlap between the two gestures. As noted above, J. Jun (1995) found coda reduction to be an important factor in obstruent clusters in Korean, while Son et al. (2007) did not. For the purposes of this study, however, duration ratio serves as a measure of how closely coordinated the transition between two words may be: whether the words are run together in a close transition or articulated separately in an open transition.

5.5. Statistical analysis for hypothesis testing. Hypotheses 6 predicts that Korean speakers will show evidence of overlap rather than lag in obstruent#obstruent sequences in their native language, as measured by duration ratio. In order to test Hypothesis 6, an analysis of variance was conducted to compare duration ratio in native Korean to previously analyzed data on duration ratio in native English and native Russian. This control data is from the study reported in Zsiga (2000), which used materials and an experimental design exactly parallel to that used in the present study. Zsiga (2000) found duration ratios in English consistent with consonant overlap, and duration ratios in Russian consistent with consonant lag, findings that were independently confirmed by articulatory studies (Kochetov & Goldstein 2001, Kochetov et al. 2007). In the present study the hypothesis of overlap in Korean is supported if duration ratio in Korean is more like English (< 1), and lag is supported if duration ratio is more like Russian (≥ 1). In this analysis, the independent variable is native language, and the dependent variable is duration ratio in the native language, with each data point corresponding to the duration ratio for a specific cluster for a specific speaker in each language: n= 54 for English and for Russian (6 speakers * 9 clusters) and n= 108 for Korean (12 speakers * 9 clusters).

Hypothesis 7 predicts that the pattern of consonant coordination will carry over from L1 Korean to L2 English. Two analyses of variance were used to test this prediction. First, a repeated measures analysis of variance compares duration ratio in native Korean with duration ratio in Korean English. The independent variables are language spoken (within subjects) and

L2 level of instruction (between subjects). The second analysis of variance compares Korean English with native English and Russian English, using data from Zsiga (2003). (Again, data was collected and analyzed using the same design as Zsiga 2000 and the present study.) Independent and dependent variables are defined as in the first anova (hypothesis 6), except that the language spoken in each case is English.

Hypothesis 8 predicts that typical gestural overlap patterns found for each subject in obstruent#nasal sequences will also emerge in obstruent#obstruent sequences, and that, for a subset of speakers, those who exhibit gradient assimilation, there will be a correlation between the degree of assimilation in L2 English (Experiment I), and the degree of overlap in L2 English (Experiment II). The quantitative measure of gestural overlap is the mean duration ratio for each subject in L2 English. Four different quantitative measures of assimilation for each speaker are considered:

1. Mean percentage of closure duration that was nasalized, averaging across the set of obstruent#nasal tokens for each speaker:

    a. including all tokens

    b. excluding the full nasals

2. Percent of tokens for each speaker that were transcribed as nasalized:

    a. including all tokens

    b. excluding the full nasals

Analyses both with and without the full nasals are conducted because while categorical nasalization is hypothesized to be associated with close transitions between words, it is not hypothesized to be the direct result of increased gestural overlap.

5.6. Results for Hypotheses 6 and 7: Duration ratio in Korean and Korean English. Figure 19 graphs mean duration ratio for the three groups of speakers, speaking their L1 (dark gray bars) and L2 English (light gray bars). Error bars indicate 95% confidence intervals.

Figure 19 about here

The analysis of variance on L1 Korean, English and Russian found a significant effect of native language ($F(2,213) = 35.04$, $p < .001$). As shown in Figure 19, duration ratio was .722 for Korean, .797 for English, and .982 for Russian. A Tukey post-hoc test found that native Korean and native English were not significantly different from each other, but both were significantly different from native Russian.[vi] Hypothesis 6 is thus confirmed: Korean and English tend toward overlap at word boundaries, unlike Russian, which tends toward a lag. (The overall duration ratio in Russian is slightly less than 1.0, since, as noted above, some clusters in Russian do show overlap, which brings down the overall average.)

Hypothesis 7 is not confirmed for the Korean speakers as a group, however: duration ratio is not the same in L1 and L2. In L2 English, the mean duration ratio for Korean speakers rises to .918: significantly different from native Korean by the repeated measures analysis ($F(1,106) = 73.66$, $p < .001$). Again, the effect of Level of Instruction was not significant ($F(1,106) = 1.791$, $p > .1$), nor was the interaction of the two variables. The mean duration ratio in English for advanced speakers (.900) was slightly lower than that for intermediate speakers (.936), but the difference was not significant. No effect of level of instruction in English was expected for pronunciation in native Korean, and none was found.

Further, the analysis of variance comparing Korean English (.918), native English (.799), and Russian English (.982) found a significant effect of native language ($F(2,213) = 11.547$, $p < .001$). Post-hoc analysis by Tukey test found that duration ratio for the Korean speakers was significantly higher than that of native speakers of English, but was not significantly different from Russian speakers of English. Thus, the evidence does not support transfer of the L1 pattern of overlap to the L2. Rather, L2 speakers are found to average duration ratios close to 1, more consistent with Word Integrity hypothesis.

The results for duration ratio thus far seem partly inconsistent with the findings on nasal assimilation in Experiment I. In native Korean, nasalization was found to be categorical, which is consistent with the close transition indicated by the low duration ratios in Experiment II. However, the high duration ratio for Korean English in Experiment II does not seem consistent with the high rates of assimilation found in Experiment I. The solution sorts itself out when individual speaker differences in both nasalization and duration ratio are considered.

5.7. Results for Hypothesis 8: Duration ratio and nasal assimilation

As was the case with both voicing and nasalization in Experiment I, the overall means in duration ratio hide important differences between speakers. Figure 20 shows the mean duration ratio for each speaker in both L1 Korean and L2 English. For a first approximation in relating duration ratio and patterns of nasalization, speakers are coded according to their majority realization of obstruent#nasal sequences (Figure 9). Filled diamonds indicate speakers who produced full nasals more often than not (K5, K12). Asterisks indicate a preference for released pre-nasal consonants (K2, K8). Open circles indicate mostly unreleased obstruents, with varying degrees of nasalization and voicing.

Figure 20 about here

Figure 20 shows that all speakers had a duration ratio well below 1.0 in L1 Korean, confirming hypothesis 6. Also, all speakers increased duration ratios from Korean to English, contra hypothesis 7, but some did so much more than others. Figure 20 also shows that hypotheses 8a and 8b are confirmed. The two speakers who typically produced released stops in Experiment 1 (K2 and K8) had duration ratios above 1.0 in Experiment 2, while the two speakers who typically produced full nasals (K5 and K12) had duration ratios below 1.0. Beyond that, however, duration ratio does not seem to be a good predictor of nasalization. In particular, speaker K12, who had the highest rate of nasalization in Experiment 1, was found to have a duration ratio of .891 in Experiment 2, very close to the mean. Also note that Speaker K10 was found in Experiment 2 to have a very high duration ratio (1.132), but this speaker was not characterized by audible release in clusters (in either experiment). Rather, this speaker tended to produce long voiceless closures without audible release.

The more systematic test of the predictive value of duration ratio is to test the correlation between duration ratio and quantitative measures of assimilation in Experiment I. The measures for each subject are given in Table 8. The correlation of each of the measures of assimilation and that of duration ratio are given at the bottom of Table 8, and are graphed in Figures 21 and 22.

Table 8 about here

Figures 21 and 22 about here

The graphs in Figure 21A and B show the correlation between duration ratio and the quantitative measures of nasal assimilation for each speaker: mean percent of cluster duration nasalized, and percent of tokens transcribed as having undergone nasal assimilation, with all tokens included. These measures give a global characterization of the tendency of each speaker toward more or less assimilation. The two measures are not unrelated to each other, and thus the graphs look similar. Both show a trend (a negative correlation) in the predicted direction: lower duration ratios in obstruent clusters are associated with higher rates of assimilation in nasal clusters. When all speakers are included, however, the trend is not significant in either case, although it does approach significance ($p = .067$) in the case of the analysis by tokens.

The graphs in Figure 21 show why the correlations are low: It is evident from the graphs that three speakers don't fit the pattern. (They graphed with a triangle (K12), or squares (K8, K10)). When these speakers are removed from the correlation, it reaches significance in both cases. These three speakers, however, are among those who are not predicted to show a significant correlation between nasalization and duration ratio. Speakers K8 and K10 are characterized by separation between words (duration ratios > 1.1), and thus are not predicted to show overlap-related nasalization (though they may occasionally produce nasalized tokens). Speaker K12 was characterized by a very high rate of categorical nasalization (75%). The finding that K12 does not have a correspondingly low duration ratio supports the hypothesis that his extensive nasalization is not the product of extensive overlap. (By comparison, it can also be noted that Speakers K4 and K6 also had an average duration ratio, but have very low rates of assimilation. These two speakers each had 0 cases of complete nasalization.) Speaker K5 had a high rate of categorical nasalization (61%), and also had a low duration ratio.

Taken together, the subject-specific findings indicate that duration ratio only partly predicts nasalization. Speakers with duration ratios greater than 1 (K2, K8, K10), indicating lag between words, had little or no nasalization. Speakers with very low duration ratios (less than .8, K5 and K7), had high rates of nasalization. With intermediate duration ratios, however (greater than .8 but less than 1), rates of nasalization varied considerably, and that variation is associated with the incidence of complete assimilation. K12 has a high rate of nasalization because this speaker often used complete assimilation, and K4 and K6 have a low rate of nasalization because these speakers never used complete assimilation. The findings thus support categorical nasalization as a separate process from gestural overlap.

Another way to attempt to tease apart the effects of categorical and gradient assimilation is to test for the correlation of duration ratio and degree of nasalization with the completely nasal tokens excluded. These analyses are shown in the last two columns of Table 8. While the trends are in the right direction (greater overlap associated with more nasalization), they do not reach significance. In large part, this may be due to the fact that with the smaller ranges of variation that result when the more extreme values are removed, one overall number for duration ratio per speaker is too gross a measure to capture fine-grained differences in gradient overlap. It is also the case, however, that more is going on, even in non-categorical cases, than consonant overlap.

The most successful correlation, that of duration ratio with percent of tokens transcribed as nasal, excluding speakers K12, K8, and K10, is graphed in Figure 22.

Figure 22 about here

This graph shows the relative rate at which these speakers produced pre-oralized nasals – hypothesized to result from extensive overlap – as related to amount of overlap in obstruent clusters. The correlation is very close to linear for a subset of speakers. The spoiler in this case is speaker K3, who had an unexpectedly high rate of non-categorical nasalization (29%), with a fairly high duration ratio (.917). (If this speaker is also removed, leaving only 8 of the original 12 speakers in the dataset, the correlation just misses significance ($r = -.693$, $p = .057$)). While the anomalous results for this speaker might be due to nothing more than the imprecision of using duration ratio to characterize a speaker's overall speech behavior, it is also interesting to note that K3 had the longest duration of modal voicing of any speaker (Figure 16). It may be that stronger voicing led to measurable nasal resonance occurring relatively earlier for this speaker compared to the others.

5.8. Summary and Discussion of Results, Experiment II.

Hypothesis 6 was confirmed. L1 Korean speakers were found to have a duration ratio of less than 1.0, indicative of overlap or close coordination between words, more similar to native speakers of English than to native speakers of Russian. This result is consistent with the articulatory data of Kochetov et al. (2007). Exact comparison between the acoustic and articulatory studies is not possible, as different obstruent clusters were considered in the two studies, and the translation of plateau duration as defined in articulatory terms and closure duration as defined in acoustic terms is not exact. Nonetheless, the studies agree that Korean is similar to English in having a general pattern of overlap in consonant clusters at word boundaries, significantly different from Russian, which tends toward a lag.

Hypothesis 7 was not confirmed for the set of speakers as a whole. All speakers increased duration ratio from L1 Korean to L2 English, with the result that mean duration ratio for the Korean English speakers was significantly higher than duration ratio for native English speakers, and not significantly different from duration ratio for Russian English speakers. The evidence does not support a simple transfer of between-word articulatory coordination from L1 to L2.

However, as was the case with the realization of obstruent#nasal sequences, there were important inter-speaker differences. The increase in average duration ratio in Korean English was driven primarily by three speakers, two of whom had a duration ratio of greater than 1.1, indicating a long lag. The other nine speakers had duration ratios at or below .9, higher than their values in Korean, but similar to values typical for English, and indicative of some degree of overlap between consonants.

Hypothesis 8 was supported, though with caveats: in general, the L2 timing patterns evidenced by the different speakers in obstruent#nasal sequences (Experiment 1) were also evident in obstruent#obstruent sequences (Experiment 2). Hypothesis 8a predicted that the speakers who had the most audible releases in Experiment 1 (K2 and K8) would have the highest duration ratios in Experiment 2. This was confirmed in that K2 and K8 did have duration ratios > 1. A third speaker, however (K10), also had a very high duration ratio (1.132), even though this high duration ratio did not result in audible release, but in a long voiceless closure. Hypothesis 8b predicted that speakers characterized by categorical nasalization (K5 and K12) would have duration ratios less than 1, and this also was confirmed. Hypothesis 8c predicted a correlation between degree of nasalization in Experiment 1 and degree of overlap in Experiment 2, but only for the subset of speakers characterized by gradient nasalization. When all speakers were considered, there was a non-significant trend in the direction predicted: lower duration ratios (indicative of more overlap) are associated with more assimilation. For a subset of the speakers (8 of 12), however, a significant or near-significant correlation was found between duration ratio and three measures of nasalization: the mean percent of closure duration nasalized, the percent of tokens transcribed as nasal, and the percent of pre-oralized tokens (Table 8 and Figures 21, 22). Duration ratio was not a good predictor of nasalization for speakers K8 and K10, who had duration ratios greater than 1.1, for speaker K12, who had the highest rate of complete nasalization, and (to a lesser extent) for speaker K3, who evidenced the most intersonorant voicing. Given the approximate nature of duration ratio, and the difficulty in assigning a single number to a speaker's pattern of articulation, high correlations might not be expected. However, Hypothesis 8 is in general confirmed: Overlap partly predicts assimilation in this data.

The correlation data can be interpreted as further evidence of two kinds of assimilation: phonetic assimilation due to overlap, and phonological assimilation due to planned reorganization. Those speakers for whom duration ratio and nasalization are correlated exhibit partial and gradient nasalization due to overlap. Those for whom duration ratio and nasalization are not correlated exhibit either a pattern of non-overlap (K2, K8, K10), consistent with Word Integrity, or a pattern of categorical assimilation (K12), consistent with planned reorganization.

6. General discussion. This study has examined in detail just one pattern of external sandhi in a second language – nasalization in Korean English – but the findings have relevance for several larger points. The introduction to this paper suggested three areas or interest: external sandhi as a second language phenomenon, external sandhi as it relates to speech planning, and external sandhi as it relates to phonetic and phonological representation. This concluding discussion addresses these topics in order.

6.1. External sandhi in L2. Beginning with the most specific conclusions, this study found a high rate of the application of nasalization: on average, 32% of obstruent#nasal sequences in Korean English were pronounced as nasalized. This is higher than has been found for many previous studies of L2 external sandhi (e.g. Cebrian 2000, Lléo & Vogel 2004), but consistent with rates of transfer found in previous studies of Korean English (Kim 2000, Park 2005). Overall, this shows that the Word Integrity hypothesis does not necessarily or generally hold of

L2 speech: at least some learners often coordinate the articulation of separate words, such that external sandhi processes can and do carry over from L1 to L2.

It remains unclear, however, whether this higher rate of application is a fact about Korean, about the combination of Korean and English, or whether this pattern is typical of a broader range of L1 and L2 combinations that do not happen to have been investigated yet. One hypothesis to be pursued is that the high application rate reflects the importance of the accentual phrase as a planning unit of Korean. The accentual phrase in Korean, which typically groups together an object and verb, or other words in close syntactic relation, has a single pitch pattern spread over the words in the phrase, with a single prominence (S.-A. Jun 1993, 1998). As was noted above (10), there are many sandhi processes that take place at the boundary between words within a phrase. It may be that the tight coordination that is typical within words in other languages extends to the accentual phrase in Korean. If so, this would explain why rates of assimilation between words in Korean English are more similar to rates of assimilation within words in other L1 and L2 combinations.

To address this hypothesis, more data on more L1 and L2 language pairs are needed. In this vein, Lléo & Vogel's (2004) study on Spanish learners of German is welcome, and other studies of external sandhi that pair languages with different prosodic characteristics are needed. An obvious next step is testing of Korean learners of Russian and Russian learners of Korean, which would fill out the L1 x L2 paradigm begun in Zsiga (2003) and continued here, and such a study is planned for the future.

While the Word Integrity hypothesis was not supported overall, there were three speakers of twelve who did produce L2 speech with a lag between words. Separation between words is not typical of either native English or native Korean, so its emergence for these speakers cannot be attributed to either carryover from the L1 or acquisition of the target pattern. The finding of a lagged pattern for at least some speakers supports the treatment of Word Integrity as a default or unmarked pattern which may emerge in L2 speech, but which may be overridden by language-particular considerations (Zsiga 2003). Nothing in the data collected here suggests why these three speakers are different from the others—no pattern in terms of length or level of instruction or age of exposure to English—so the difference at this point must be attributed to idiosyncrasy. The inter-subject differences do emphasize, however, that different speech planning modes are available, from which L2 speakers can and do choose (see section 6.2 below). Further research might target other characteristics that might predispose a learner to one pattern or the other. It also remains to be determined which pattern – lag with its extra releases and stronger boundaries or overlap with its unexpected sandhi – is more disruptive to comprehension by native speakers. Only after such factors have been explored can effective interventions be planned and evaluated. This study only begins by documenting the different patterns that occur, but it is hoped that further classroom-based studies might build on this baseline.

The present study found individual variation, but no overall effect of level of instruction. Since level of instruction and time in the U.S. were not independent in this sample of speakers, there is thus also no overall effect of time spent in an English-speaking environment. This lack of a level effect suggests at least two conclusions. First, habits of articulation from the L1, in word-to-word coordination as well as in segmental articulation, can and do persist even after years in the L2 environment. Secondly, the lack of a level effect emphasizes that general tests for overall proficiency and classroom placement are not good predictors of the phonetic characteristics of learners' speech. Studies where learners are divided into groups based on a more direct rating of accentedness may be more likely to find significant effects, though there is a danger of circularity in that approach. It may also be the case, however, that given the degree of inter-subject variation in this dataset, that the sample of 12 speakers was just too small for an effect of level of instruction or length of residence to emerge as statistically significant.

Finally, the difference between fully nasalized tokens and pre-oralized tokens, a difference only evident through acoustic analysis, emerged as an important distinction in this study. It is certainly important to transcribe L2 speech as it is perceived by a native listener. But phonetic analysis is also imperative for a fuller understanding of the speech patterns that underlie the perceptions.

6.2. External sandhi and speech planning. A further goal of the present study has been to investigate general patterns of coordination of consonant gestures at word boundaries, in both L1 and L2, and to consider how they might be handled in a framework based on coupled oscillators (Browman & Goldstein 2000, Goldstein et al. 2006). Four distinct patterns have been identified:

1. Word Integrity: lack of coordination between words, found for some L2 speakers.

2. Sequencing of words, optimized to foster audible release of word-final consonants in certain sequences, but allowing overlap in others (L1 Russian).

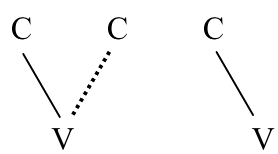
3. Overlap between words, leading under the right conditions to the perception of assimilation in consonant clusters (L1 English, many cases of Korean English).

4. Overlap between words, with the addition of planned reorganization leading to category-changing assimilation (L1 Korean, some cases of Korean English).

Each of these receives a straightforward (though at this point tentative) interpretation in terms of articulatory coupling. These are graphed in examples (12—15).

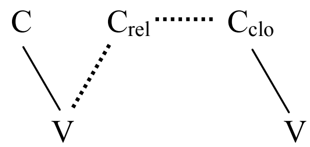
The first pattern, Word Integrity, corresponds to the lack of any temporal coordination between words. Each word is computed as a separate planning unit, with no specific strategy for transitioning between words. This is illustrated in (12). This plan would certainly be the simplest for L2 speakers to use, and so could be expected to emerge as the unmarked case. The lack of any timing relation between the two words can be hypothesized to be equivalent to the statement that each word forms its own phonological phrase (though the details of such a correspondence certainly remain to be worked out).

(12) Coupling graph for Word Integrity: no coordination across the word boundary, audible release may or may not occur



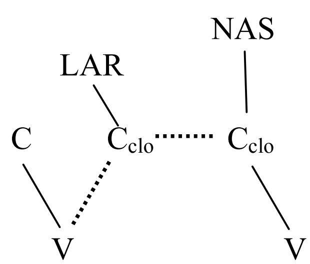
The second pattern, typical of native Russian (Zsiga 2000, Kochetov et al. 2007), involves coordination that takes into account characteristics of the word-final consonant, and ensures an audible release, thereby fostering clearer production of an important perceptual cue to the differing places of articulation (Kochetov 2006a). Back-to-front clusters ([k#p]) are lagged while front-to-back clusters are overlapped ([p#k]). This pattern requires a lag in just those clusters where the audible release of C1 is likely to be obscured by C2. The required lag could be achieved by timing the closure of the second consonant to begin sequential to the release of the first, as shown in (13). This approach follows Nam (2007) and Goldstein et al. (2007, in press) in allowing coupling graphs to reference closure and release gestures independently.

(13) Coupling graph for Russian back-to-front clusters: release is audible



It has already been proposed (5 above) that consonant overlap (typical of native English, some clusters in native Russian, and some speakers of Korean English) can be modeled in terms of sequential coupling of the two consonant's closure gestures (Nava et al. 2008). This coupling is shown in (14), based on (5), but with the addition of further detail, specifying the closure-to-closure timing and additional laryngeal and velic gestures. For simplicity, laryngeal and velic gestures were not indicated in (12) and (13), though they are assumed to be present; but such gestures become relevant in cases where overlap and assimilation become issues.

(14) Closure-to-closure coupling results in consonant overlap and gradient assimilation.

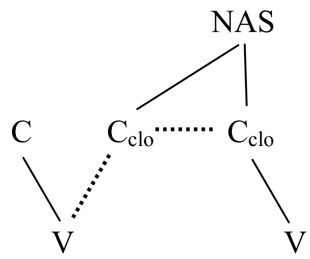


As was discussed in section 2 above, a closure-to-closure coupling will result in close transition between the coda and onset consonants, with variable overlap. With this pattern of overlap, whether or not assimilation is perceived will depend in large part on the specific oral gestures involved (Byrd 1992), the accompanying laryngeal and supralaryngeal gestures, the strengths of the different couplings, and some degree of random variation (Nam & Saltzman 2003).

In (14), an obstruent#nasal sequence is illustrated. In such a sequence, the velum opening gesture will begin at some point early in the consonant cluster (Cohn 1993), though exactly when nasal resonance will begin to be perceived will depend in large part on the laryngeal state. For English obstruent#nasal clusters, the laryngeal gesture is assumed to be one of adduction (Huffman 2005). If the adduction gesture is strong, there will be complete voiceless closure in the first half of the cluster (see Figure 2 above.) If the adduction gesture is weakened, the duration of voiceless closure will vary (Figure 16), or creaky voicing may result (Figure 3).

Example (15) shows the hypothesized pattern for native Korean. The vowel-to-vowel and consonant overlap coordinations are the same as in English: both languages show overlap in obstruent#obstruent clusters (Kochetov et al 2007; Figure 19 above). It may be the case that there is more reduction of C1 in Korean (J. Jun 1996), but the data is inconclusive (Kochetov & Pouplier 2008, Son et al 2007). In the case of an obstruent#nasal cluster, however, Korean utilizes a phonological reorganization across the word boundary, symbolized here by the addition of a new coupling. With this additional coupling, the coda consonant is now planned as part of a unit that is fully nasal. Velum opening will be planned to begin during the first vowel, so that full nasal resonance is achieved by the time of closure for the first consonant. The exact phasing details remain to be worked out, and would depend, as was noted in section 2 above, on factors such as coupling strength and the inherent dynamics of the velic gesture. It is also assumed that the laryngeal gesture is deleted, though it would be weak and variable in any case.

(15) Planned reorganization results in categorical assimilation.



For Korean English, it is hypothesized that speakers may be using the pattern either in (12), (14), or (15). Pattern (12), Word Integrity, is typical of speakers K2, K8, and K10. They show very little assimilation between words, and may or may not produce a coda consonant with audible release. Pattern (15), phonological assimilation, is especially typical of speakers K5 and K12, but other speakers also use this pattern, although less frequently. Pattern (15) represents the transfer of Korean phonology to L2 English. Pattern (14), overlap without phonological assimilation, is the one found most frequently, and, it is hypothesized, the reason for the correlation found in this study between overlap and perceived assimilation.

Korean English speakers using pattern (14) are hypothesized to differ from English speakers primarily in that the laryngeal gesture they use is typically one of weak abduction rather than adduction. Native English and Korean English speakers have similar ranges in terms of percent of the cluster that is nasalized (Figure 10), but Korean speakers have significantly more modal voicing (Figure 16). This pattern results in gradient nasalization (the pre-oralized tokens) and frequent voicing of the coda stop portion.

Given the assumption that consonant overlap is typical of both L1 Korean and L1 English, the data analyzed here can not distinguish whether the Korean speakers of English using pattern (14) are transferring to English the pattern of Korean articulatory timing (that is, Korean phonetics) without the Korean phonology, or whether they are successfully adopting the English pattern of overlap (without the English tendency to glottalize). It is hoped that a study of Korean learners of Russian may disentangle these two factors: if Korean learners also show gradient nasalization in pronouncing Russian phrases, that will support the hypothesis of transfer, and be further evidence for the importance of the phrase as a timing unit in Korean.

Overall, the patterns illustrated in (12) through (15) would seem to cover the range of predicted possibilities for consonant-to-consonant sequential timing in a theory of coupled oscillators, that is:

- no coupling (12)
- sequential coupling of release-to-closure (13)
- sequential coupling of closure-to-closure (14)
- sequential coupling of closure-to-closure with phonological assimilation (15)

It is not clear if simultaneous coupling of consonants across a word boundary is ever attested. In an alternative but plausible approach, one might model consonant overlap (14) with a simultaneous coupling between the two consonants, and lag (13) with a sequential coupling, thereby avoiding reference to closure and release gestures. While the study of Nava et al. (2008) suggests that closure-to-closure coupling best models the English pattern, further modeling studies with more extensive data from different languages are needed. Such studies are also needed to investigate whether patterns such as relase-to-release, which are predicted by Nam's (2007) model, are attested.

Another missing pattern is release-to-closure sequential coupling with phonological assimilation. This may be universally ruled out; evidence in favor of such a patterning would come from a case of consonant-to-consonant categorical assimilation across word boundaries (similar to Korean nasalization), but where final consonants are necessarily released even when assimilation has applied. Also, since the present study has concentrated on patterns of coda consonant assimilation, vowel-to-vowel timing has not been a focus. Vowel-to-vowel sequential timing is almost surely attested in languages that lack codas, and perhaps in some languages that allow them. (See Smith 1992 for articulatory evidence that languages may differ in whether syllable-to-syllable coordination takes into account intervening consonants or not .) Marin (2007) uses simultaneous coupling of vowels to model different realizations of the /ea/ dipthong in Romanian; this pattern might also be instantiated in languages with vowel coalescence at word boundaries (perhaps similar to Igbo (Zsiga 1997)). In general, much further research, both phonological and phonetic, is needed to test whether a typology of coupled oscillators, with and without independent closure and release gestures, and with and without categorical reorganization turns out to be useful in categorizing cross-linguistic patterns of word-to-word timing and external sandhi.

6.3. External sandhi and phonetic and phonological representation. Some further general points are noted in conclusion.

The high incidence of nasalization in Korean English attests to its robust productivity and independence from the Korean lexicon. Korean speakers have generalized a pattern that they apply to eligible sequences not only in new words, but in a new language. The question to be addressed by linguistic theory is how this pattern is best to be represented.

It has been argued here that the data support two different kinds of nasalization in Korean English, one that might be termed phonological and the other phonetic (see Cohn e.g. 1990, 1993). In the one case, full nasalization, a coda segment that is underlyingly not nasal receives a nasal specification (a nasal "target" in Cohn's terms), which changes its phonological category, and correspondingly alters the speech production plan. In the second case, pre-oralized nasals, the perception of nasalization arises as a result of the previously specified, more general, speech production plan (in this case consonant overlap) being carried out.

The case of Korean external sandhi, in L1 and especially in L2, strengthens evidence that category-changing phonology can apply across word boundaries. There is a level of phonological planning and change that is on the one hand independent of lexical specification and on the other hand different from an unintended consequence of coarticulation. There is a real phonology in between the lexicon and speech production.

Calling the category-changing assimilation "phonology" and the non-categorical assimilation "phonetics" recognizes that there are two different outcomes. Nonetheless, these terms may place too much of a wall between them, setting too firmly in stone the idea that two separate approaches, analyses, and mechanisms are required. Incorporating the theory of coupled oscillators, as part of Articulatory Phonology, offers a new perspective. Limiting couplings to sequential and simultaneous captures the essential insight of Autosegmental Phonology: that phonological units may either be linked to a single hierarchical node or sequenced along a featural tier, sometimes in complex patterns, but never requiring more detailed specification in order to represent the possibilities of categorical contrast and alternation. On the other hand, the presence of competing couplings and variable coupling strength allows for the rich computation of language-specific variability in realization. The possibility of specifying different combinations of couplings, as in (12—15), allows the beginnings of a typology to emerge.

The claim is not that there is no difference between phonology and phonetics: exactly the opposite has been shown in the demonstration of two kinds of nasalization in this data. Rather, the claim is that there may be a single object, call it a gestural constellation, that both phonology and phonetics may operate on. The phonology works by adding and deleting gestural associations, and the phonetics works by computing and adjusting the associations supplied by the phonology. (The wording here assumes a precedence relationship between phonological and phonetic operations, but that is a separate topic.)

Much further work remains to be done. The relationship between prosody and articulatory timing is complex, as noted in section 1 above. It is beyond the scope of this paper to address a comprehensive theory of prosody and timing, especially because boundary strength was not varied in this study, and the graphs in examples (12—15) only sketch a series of possibilities. Much more work is needed, both in data analysis and in computational modeling, to determine whether or not the couplings required to determine patterns of overlap will correspond with traditional understandings of prosodic constituency. The important role of perception in both phonology and phonetics has only briefly been touched on here. Much work remains to be done, but it is hoped that the data provided here will prove useful in working toward an improved theory of prosody, articulatory timing, and assimilation, in both L1 and L2.

References

Altenberg, Evelyn and Robert Vago. 1983. Theoretical implications of an error analysis of second language phonology production. Language Learning 33.427-448.

Anderson-Hsieh, Janet, Ruth Johnson, and Kenneth Koehler. 1992. The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable-structure. Language Learning 42.529-555.

Anderson-Hsieh, Janet, Timothy Riney, and Kenneth Koehler. 1994. Connected speech modifications in the English of Japanese ESL learners. Issues and Developments in English and Applied Linguistics 7.31-51.

Barry, Martin C. 1985. A palatographic study of connected speech processes. Cambridge Papers in Phonetics and Experimental Linguistics 4/1985.1-16.

Barry, Martin C. 1992. Palatalisation, assimilation and gestural weakening in connected speech. Speech Communication 11.393-400.

Bauer, Matthew. 2005. Prosodic strengthening of consonants in Iron Range English. Washington, DC: Georgetown University PhD dissertation.

Boersma, Paul, and David Weenink. 2007. Praat: Doing phonetics by computer (Version 4.6.21) [Computer program]. Retrieved September 2007, from http://www.praat.org/.

Bradley, Travis. 2007. Morphological derived-environment effects in gestural coordination: A case study of Norwegian clusters. Lingua 117(6).950-985.

Browman, Catherine. 1994. Lip aperture and consonant releases. Phonological structure and phonetic form: Papers in laboratory phonology III, ed. by P. Keating, 331-353. Cambridge: Cambridge University Press.

Browman, Catherine. 1995. Assimilation as gestural overlap: comments on Holst and Nolan. Phonology and phonetic evidence. Papers in laboratory phonology IV, ed. by B. Connell and A. Arvaniti, 334-342. Cambridge: Cambridge University Press.

Browman, Catherine, and Louis Goldstein. 1986. Towards an articulatory phonology. Phonology Yearbook 3, ed. by C. Ewen and J. Anderson, 219-252. Cambridge: Cambridge University Press.

Browman, Catherine, and Louis Goldstein. 1988. Some notes on syllable structure in articulatory phonology. Phonetica 45.140–155.

Browman, Catherine, and Louis Goldstein. 1989. Articulatory gestures as phonological units. Phonology 6.201-251.

Browman, Catherine, and Louis Goldstein. 1990a. Tiers in articulatory phonology, with some implications for casual speech. Papers in laboratory phonology I: Between the grammar and physics of speech, ed. by J. Kingston and M. Beckman, 341-376. New York: Cambridge University Press.

Browman, Catherine, and Louis Goldstein.1990b. Gestural specification using dynamicalIy defined articulatory structures. Journal of Phonetics 18.299-320.

Browman, Catherine, and Louis Goldstein. 1992. Articulatory phonology: An overview. Phonetica 49.155-180.

Browman, Catherine, and Louis Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. Bulletin de la Communication Parleé 5.25-34.

Bybee, Joan. 2000. Phonology of the lexicon. Usage-based models of language, ed. by M. Barlow and S. Kemmer, 65-85. Stanford, CA: CSLI Publications.

Bybee, Joan. 2002. Phonological evidence for exemplar storage of multiword seqeunces. Studies in Second Language Acquisition 24.215-221.

Byrd, Dani. 1992. Perception of assimilation in consonant clusters: a gestural model. Phonetica 49.1-24.

Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. Journal of Phonetics 24.209-244.

Byrd, Dani. 2006. Relating prosody and dynamic events: Commentary on the papers by Cho, Navas, and Smiljanic. Papers in Laboratory Phonology 8, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine Best, 549-561. Berlin, New York: Mouton de Gruyter.

Byrd, Dani, Abigail Kaun, Shrikanth Narayanan, and Elliot Saltzman. 2000. Phrasal signatures in articulation. Papers in Laboratory Phonology 5. Acquisition and the Lexicon, ed. by M. B. Broe and J. Pierrehumbert, 70-87. Cambridge,: Cambridge University Press.

Byrd, Dani, and Sungbok Lee, and Rebeka Campos-Astorika. 2008. Phrase boundary effects on the temporal kinematics of sequential tongue tip consonants. Journal of the Acoustical Society of America 123(6).4456-4465.

Byrd, Dani, and Elliot Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. Journal of Phonetics 31.149-180.

Cebrian, Juli. 2000. Transferability and productivity of L1 rules in Catalan-English interlanguage. Studies in Second Language Acquisition 22.1-26.

Chen, Larissa. 2003. Evidence for the role of gestural overlap in consonant place assimilation. Proceedings of the XVth ICPhS, Barcelona, 2821-2824.

Cho, Taehong. 2002. The Effects of Prosody on Articulation in English. New York, NY: Routledge.

Cho, Taehong. 2007. Prosodic strengthening in speech production and perception: The current issues. Speech Sciences 14(4).7-24. [Published by The Korean Association of Speech Sciences, KASS].

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

Chu, Harold, and Young-Hee Park. 1978. A contrastive analysis between Korean and English for ESL teachers. Arlington, VA: Bilingual Education Project, Arlington Public Schools.

Cohn, Abigail. 1990. Phonetic and Phonological Rules of Nasalization. Los Angeles, CA: UCLA PhD dissertation.

Cohn, Abigail. 1993. Nasalisation in English: Phonology or phonetics. Phonology 10.43-81.

Cutler, Anne, Delphine Dahan, and Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. Language and Speech 40.141-201.

Ellis, Andrew W., and William Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. Journal of Phonetics 30.373-396.

Erman, Britt, and Beatrice Warren. 2000. The Idiom Principle and the Open Choice Principle. Text 20.29–62

Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. Journal of the Acoustical Society of America 97.3125–3134.

Gafos, Diamandis. 2002. A grammar of gestural coordination. Natural Language and Linguistic Theory 20.269-337.

Goldsmith, John. 1976. Autosegmental phonology. Ph.D. dissertation, MIT. (Published 1979, New York: Garland.)

Goldstein, Louis. 2008. A coupled oscillator model of speech production planning. Paper presented at the Fifth International Workshop on Language Production, Annapolis, MD.

Goldstein, Louis, Dani Byrd, and Elliot Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. From Action to Language: The Mirror Neuron System, ed. by M. Arbib, pp. 215-249. Cambridge: Cambridge University Press.

Goldstein, Louis, Ioana Chitoran, and Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. Proceedings of the 16th Congress of Phonetic Sciences, 241-244.

Goldstein, Louis, Hosung Nam, Elliot Saltzman, and Ioana Chitoran. In press. Coupled oscillator planning model of speech timing and syllable structure. *Chinese Journal of Phonetics*. (url: http://sail.usc.edu/~lgoldste/ArtPhon/Papers/Week%2012/LG_final.pdf).

Haken, Hermann, J. A. Scott Kelso, H. Bunz. 1985. A theoretical model of phase transitions in human hand movements. Biological Cybernetics 51.347-356

Hirose, Hajime, Charles Y. Lee, and Tatsujiro Ushijima. 1974. Laryngeal control in Korean stop production. Journal of Phonetics 2.145-152.

Huffman, Marie. 2005. Segmental and prosodic effects on coda glottalization. Journal of Phonetics 33.335-362.

Holst, Tara, and Francis Nolan. 1995. The influence of syntactic structure on [s] to [ʃ] assimilation. Phonology and phonetic evidence. Papers in laboratory phonology 4, ed. by B. Connell and A. Arvaniti, 315-333. Cambridge: Cambridge University Press.

Inkelas, Sharon and Draga Zec. 1990. The phonology-syntax connection. Chicago: University of Chicago Press

Jones, Daniel. 1956. An outline of English phonetics. New York: E. P. Dutton & Co.

Jun, Jongho. 1995. Perceptual and articulatory factors in place assimilation: An optimality theoretic approach. Los Angeles, CA: UCLA PhD dissertation.

Jun, Jongho. 1996. Place assimilation is not the result of gestural overlap: Evidence from Korean and English. Phonology 13(3).377-407.

Jun, Sun-Ah. 1993. The phonetics and phonology of Korean prosody. Columbus, OH: Ohio State University PhD dissertation.

Jun, Sun-Ah. 1995. Asymmetrical prosodic effects on the laryngeal gesture in Korean. Papers in laboratory phonology IV: Phonology and phonetic evidence, ed. by B. Connell and A. Arvaniti, 235-253. Cambridge: Cambridge University Press.

Jun, Sun-Ah. 1998. The accentual phrase in the Korean prosodic hierarchy. Phonology 15(2).189-226.

Kaisse, Ellen. 1985. Connected speech: The interaction of syntax and phonology. Orlando, FL: Academic Press.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic wordfinal stops in Korean. Phonology 20.219-274.

Keating, Patricia. 2006. Phonetic encoding of prosodic structure. Speech production: Models, phonetic processes, and techniques, ed. by J. Harrington and M. Tabain, 167-186. Macquarie Monographs in Cognitive Science. New York, Hove: Psychology Press.

Keating, Patricia, Taehong Cho, Cécile Fougeron, and Chai-Shune Hsu. 2004. Domain-initial strengthening in four languages. Laboratory Phonology 6, ed. by J. Local, R. Ogden, and R. Temple, 145-163. Cambridge: Cambridge University Press.

Kim, Hyouk-Keun. 2000. A study of the interlanguage phonology of Korean learners of English. Washington, DC: Georgetown University PhD dissertation.

Kim-Renaud, Young-key. 1991. Korean consonantal phonology. Seoul: Hanshin.

Kochetov, Alexei. 2006a. Testing Licensing by Cue: A case of Russian palatalized coronals. Phonetica 63.113-148.

Kochetov, Alexei. 2006b. Syllable position effects and gestural organization: Evidence from Russian. Papers in Laboratory Phonology VIII, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine Best, 565-588. Phonology and Phonetics Series 4-2. Berlin, New York: Mouton de Gruyter.

Kochetov, Alexei, and Louis Goldstein. 2001. Competing recoverability factors and inter-gestural phasing in Russian stop clusters. Paper presented at the 75th meeting of the Linguistic Society of America, Washington, DC.

Kochetov, Alexei, and Marianne Pouplier. 2008. Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. Phonology 25(3).1-33.

Kochetov, Alexei, Marianne Pouplier, and Minjung Son. 2007. Cross-language differences in overlap and assimilation patterns in Korean and Russian. Proceedings of the 16th International Congress of Phonetics Science, 1361-1364.

Krakow, Rena. 1999. Physiological organization of syllables: A review. Journal of Phonetics 27.23-54.

Kühnert, B. 1993. Some kinematic aspects of alveolar-velar assimilations. Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 31.263-272.

Kohler, Klaus. 1994. Glottal stops and glottalization in German. Phonetica 51.38–51.

Ladd, D. Robert and James M. Scobbie. 2003. External sandhi as gestural overlap? Counter-evidence from Sardinian. Papers in Laboratory Phonology 6, ed. by J. Local, R. Ogden, and R. Temple, 162-180. Cambridge: Cambridge University Press.

Lléo, Conxita and Irene Vogel. 2004. Learning new segments and reducing domains in German L2 phonology: The role of the Prosodic Hierarchy. International Journal of Bilingualism 8.79-104.

Löfqvist, Anders, and Vincent Gracco. 1999. Interarticulator programming in VCV sequences: lip and tongue movements. Journal of the Acoustic Society of America 105.1854-1876.

Marin, Stefania. 2007. An articulatory modeling of Romanian dipthong alternations. Proceedings of the International Congress of Phonetic Sciences XVI.453-456.

Moll, Kenneth L., and Raymond Daniloff. 1971. Investigation of the timing of velar movements during speech. Journal of the Acoustical Society of America 50(2).678-684.

Munro, Murray J., and Tracey M. Derwing. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. Language Learning 45.73–97.

Nam, Hosung. 2007. A competitive, coupled oscillator model of moraic structure: Split- gesture dynamics focusing on positional asymmetry. Papers in Laboratory Phonology IX, ed. by J. Cole. Berlin: Mouton de Gruyter.

Nam, Hosung, Louis Goldstein, and Elliot Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. Approaches to phonological complexity, ed. F. Pellegrino, E. Marsico, I. Chitoran, C. Coupé. Berlin: Mouton de Gruyter.

Nam, Hosung, and Eliot Saltzman. 2003. A competitive coupled oscillator model of syllable structure. Proceedings of the 15th International Congress of Phonetic Sciences. 2252-2256.

Nava, Emily, Louis Goldstein, Elliot Saltzman, Hosung Nam, and Michael Proctor. 2008. Modeling prosodic rhythm: Evidence from L2 speech. Poster presented at the 156th Meeting of the Acoustical Society of America. Miami, Florida. November 2008. (url: http://artphon.usc.edu/LG/documents/ASA%20poster%20FINAL.pdf)

Nespor, Marina and Irene Vogel. 1986. Prosodic phonology. Dorcrecht: Foris.

Pierrehumbert, Janet. 1995. Prosodic effects on glottal allophones. Vocal fold physiology 8, ed. by Osamu Fujimura, 39-60. San Diego: Singular Press.

Pierrehumbert, Janet and Paul Gross. 2003. Community Phonology. Chicago Linguistics Society.

Prince, Alan and Paul Smolensky. 2004. Optimality Theory: Constraint interaction in Generative Grammar. Cambridge, MA: MIT Press.

Redi, Laura, and Stefanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. Journal of Phonetics 29.407-429.

Rubach, Jerzy. 1984. Rule typology and phonological interference. Theoretical issues in contrastive phonology: Studies in descriptive linguistics, ed. by S. Eliasson, 37-50. Heidelberg, Germany: Julius Groos.

Sagey, Elizabeth. 1988. On the ill-formedness of crossing association lines. Linguistic Inquiry 19.109-118.

Saltzman, Elliot. 1986. Task dynamic coordination of the speech articulators: A preliminary model. Experimental brain research series 15, ed. by Heuer Fromm, 129-144. New York:

Springer.

Saltzman, Elliot, Anders Löfqvist, Jeff Kinsella-Shaw, Bruce Kay, and Philip Rubin. 1995. On the dynamics of temporal patterning in speech. Studies in speech production: A Festschrift for Katherine Safford Harris, ed. by F. Bell-Berti and L. Raphael. Woodbury, NY: American Institute of Physics.

Saltzman, Elliot, and Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. Ecological Psychology I. 333-382.

Saltzman, Elliot, Hosung Nam, Jelena Krivokapic, and Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. Proceedings of the 4th International Conference on Speech Prosody (Speech Prosody 2008), ed. by P. A. Barbosa, S. Madureira, and C. Reis. Campinas, Brazil

Scobbie, James. 2007. Interface and overlap in phonetics and phonology. The Oxford handbook of linguistic interfaces, ed. by Gillian Ramchand and Charles Reiss, 17-52. Oxford: Oxford University Press.

Scobbie, James M., and Alan A. Wrench. 2003. An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona.

Selkirk, Lisa. 1984. Phonology and syntax: The relation between sound and structure. Cambridge, MA: MIT Press.

Selkirk, Lisa. 1986. On derived domains in sentence phonology. Phonology Yearbook 3.371-405.

Silva, David. 1992. The phonetics and phonology of stop lenition in Korean. Ithaca, NY: Cornell University PhD dissertation.

Smith, Caroline. 1992. The temporal organization of vowel and consonant gestures. New Haven, CT. Yale University PhD dissertation.

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

Solé, Maria-José. 1997. Timing patterns in the production of a foreign language. Views on the acquisition and use of a second language: Proceedings of the European Second Language Association 7th Annual Conference, ed. by L. Dîaz and C. Perez, 539-551. Barcelona: Universitat Pompeu Fabra.

Son, Minjung, Alexei Kochetov, and Marianne Pouplier. 2007. The role of gestural overlap in perceptual place assimilation: Evidence from Korean. Papers in Laboratory Phonology IX, ed. by J. Cole and J. I. Hualde, 507-534. Berlin: Mouton de Gruyter.

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. Proceedings of the 1998 Linguistics and Phonetics Conference, ed. by O. Fujimura, B. Joseph and B. Palek. Prague: Karolinum Press.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. The role of speech perception in phonology, ed. by E. Hume and K. Johnson, 219-250. New York: Academic Press.

Stuart, Davis, and Seung-Hoon Shin. 1999. The syllable contact constraint in Korean: An Optimality-Theoretic Analysis. Journal of East Asian Linguistics 8(4).285-331.

Tajima, Keiichi, Robert Port, and Jonathan Dalby. 1997. Effects of temporal correction on intellligibility of foreign-accented English. Journal of Phonetics 25(1).1-24.

Trofimovich, Pavel, and Wendy Baker. 2006. Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. Studies in Second Language Acquisition. 28.1-30.

Turvey, Michael. T. 1990. Coordination. American Psychologist 45.938-953.

Zsiga, Elizabeth C. 1994. Acoustic evidence for gestural overlap in consonant sequences. Journal of Phonetics 22.121-140.

Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV, ed. by B. Connell and A. Arvaniti, 282-302. Cambridge: Cambridge University Press.

Zsiga, Elizabeth C. 1997. Features, gestures, and Igbo vowels: An approach to the phonology/phonetics mapping. Language 73.227-224.

Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. Journal of Phonetics 28.69-102.

Zsiga, Elizabeth C. 2003. Relearning consonant timing. Studies in Second Language Acquisition 25(3).399-432.

Zsiga, Elizabeth C., and Hyouk-Keun Kim. 2005. What transfers? Word-integrity and assimilation in Korean/English interlanguage. Proceedings of the Boston University Conference on Language Development 29.674-685.

	Age at first English instruction	Years of English instruction	Length of residence in U.S.	Other languages spoken
Advanced speakers				
K1	13	20	6 yrs	
K2	13	10	5 yrs	
K3	11	13	4 yrs	
K4	14	20	4 yrs	
K5	11	14	1.5 yrs	Spanish (adv.)
Mean:	12.4	15.4	4.1 yrs	
Intermediate speakers				
K6	13	6	4 mos	Japanese (beg.)
K7	12	7	6 mos	Chinese (beg.)
K8	11	7	3 mos	
K9	8	16	6 mos	
K10	10	10	6 mos	
K11	12	6	12 mos	Japanese (beg.)
K12	15	6	4 mos	Japanese (beg.)
Mean:	11.6	8.3	5.9 mos	

Table 1. Korean participant background information (self-reported).

Sequence	Korean	English
p#m	/kimpap mekta/ *eat sushi*	keep Matt stop Matt
t#m	/ot mantulta/[1] *make clothes*	ate mine bought mine
k#m	/jak mekta/ *take medicine*	take Matt pick Matt
p#n	/cikap neta/ *put a purse in*	keep Nat stop Nan
t#n	/ot neta/ *put clothes in*	ate nine bought nine
k#n	/c^haik neta/ *put a book in*	take Nat pick Nat
n#m	/jen mantulta/ *make a kite*	train Matt can Matt
n#n	/sinm^wun neta/ *put a newspaper in*	train Nat can Nat

Table 2. Phrases used to test for nasalization.

[1] The Korean word meaning clothes is lexically /os/, but /s/ is realized as [t] in the syllable coda.

	Duration (ms)	
	mean	std dev
n#m	126.4	25.2
t#m --> n#m	119.9	24.6
n#n	111.3	25.2
t#n --> n#n	119.2	22.8

Table 3. Duration of underlying and derived nasal sequences in native Korean.

	E1	E2	E3
Voiceless released	7	3	2
Voiceless unreleased	24	25	33
Creaky voiced	5	8	1

Table 4. Pronunciation of pre-nasal voiceless stops (n = 36) by the native English speakers.

		Duration (ms)	
	# of tokens	mean	std. dev.
n#m	69	149	29
t#m --> n#m	16	149	32
n#n	70	126	27
t#n --> n#n	13	128	27

Table 5. Duration of underlying and derived nasal sequences in Korean English.

	Full nasals included		Full nasals not included	
	mean	std dev	mean	std dev
K1	66%	27%	50%	16%
K2	44%	11%	44%	11%
K3	59%	19%	56%	17%
K4	37%	10%	37%	10%
K5	82%	26%	53%	20%
K6	44%	18%	44%	18%
K7	74%	26%	53%	15%
K8	45%	13%	45%	13%
K9	62%	29%	43%	13%
K10	53%	19%	48%	13%
K11	43%	29%	31%	14%
K12	93%	16%	66%	21%
Korean average	**58%**	**20%**	**48%**	**15%**
E1			49%	15%
E2			51%	11%
E3			46%	13%
English average			**49%**	**13%**

Table 6. Percent consonant duration that is nasalized: English obstruent#nasal clusters

Sequence	Korean	English
$p\#p^h$	/cip p^halta/ *sell a house*	keep Pat stop Pam
$p\#t^h$	/pap t^hajwuta/ *burn the rice*	keep Tom stop Ted
$p\#k^h$	/nap k^hajta/ *dig up lead*	keep cats stop Ken
$t\#p^h$	/k'ot p^halta/ *sell flowers*	ate pie bought pie
$t\#t^h$	/ot t^hajwuta/ *have cloth burnt*	ate ten bought ten
$t\#k^h$	/k'ot k^hiwuta/ *raise flowers*	ate cake bought cake
$k\#p^h$	/jak p^halta/ *sell medicine*	make Pam pick Pam
$k\#t^h$	/c^wuk t^hajwuta/ *burn the porridge*	make Tom pick Tom
$k\#k^h$	/hopak k^hiwuta/ *raise pumpkins*	make Ken pick Ken
$V\#p^h$	/sakwa p^halta/ *sell apples*	saw Pam
$V\#t^h$	/tampaj t^hajwuta/ *smoke a cigarette*	saw Tom
$V\#k^h$	/kaj k^hiwuta/ *raise a dog*	saw Ken
p#h	/cip helta/ *demolish a house*	keep hats stop horses
t#h	/k'ot huntulta/ *wave flowers*	ate ham bought ham
k#h	/p^jek helta/ *demolish a wall*	make hats pick ham

Table 7. Phrases used to test for gestural overlap.

		Full nasals included		Full nasals not included	
Speaker	Duration Ratio L2 English	Mean percent of closure duration nasalized	Percent of tokens transcribed as nasal	Mean percent of closure duration nasalized	Percent of tokens transcribed as nasal
1	0.888	66%	42%	50%	17%
2	1.024	44%	0%	44%	0%
3	0.917	59%	33%	56%	29%
4	0.890	37%	3%	37%	3%
5	0.779	82%	75%	53%	36%
6	0.857	44%	6%	44%	6%
7	0.761	74%	50%	53%	16%
8	1.16	45%	11%	45%	11%
9	0.816	62%	39%	43%	8%
10	1.132	53%	11%	48%	3%
11	0.93	43%	17%	31%	0%
12	0.891	93%	89%	66%	71%
Correlation with Duration Ratio (all speakers):					
Pearson's rho:		-.487	-.545	-.169	-.293
p		.108	.067	.599	.356
Correlation with Duration Ratio (excl. K12 and dur ratio > 1.1):					
Pearson's rho:		-.705	-.731	-341	-.498
p		*.034	*.025	.370	.172

Table 8. Measures of overlap and assimilation.

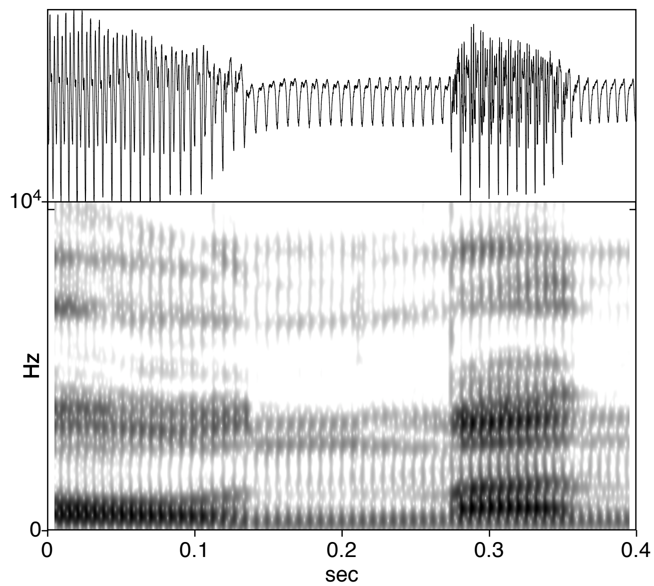


Figure 1. A derived nasal sequence in Korean: Speaker K12: [on man] from /os mantulta/, make clothes.

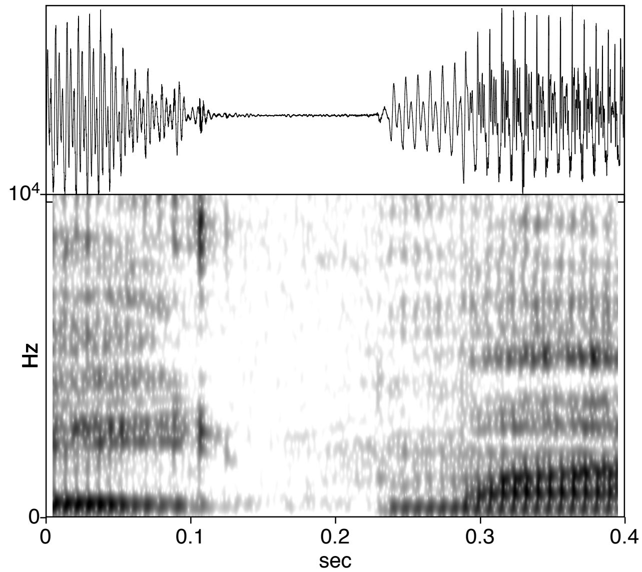


Figure 2. Unreleased voiceless stop, with glottalization. Speaker E2, ate mine (repetition 1).

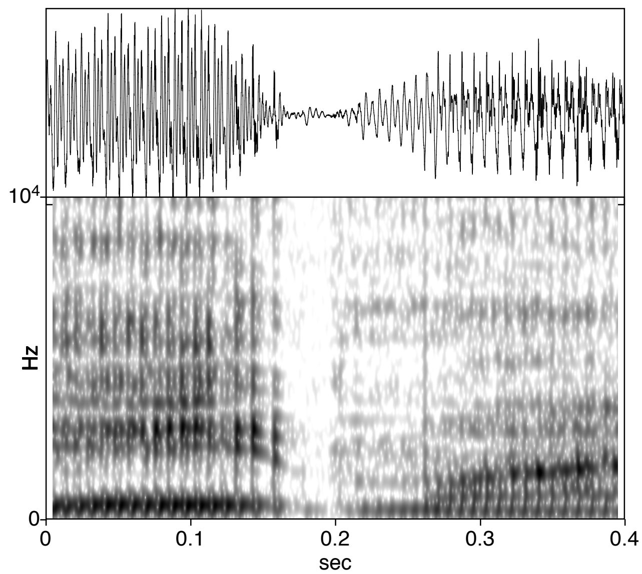


Figure 3. Short closure with creaky voice. Speaker E2, ate mine, repetition 3.

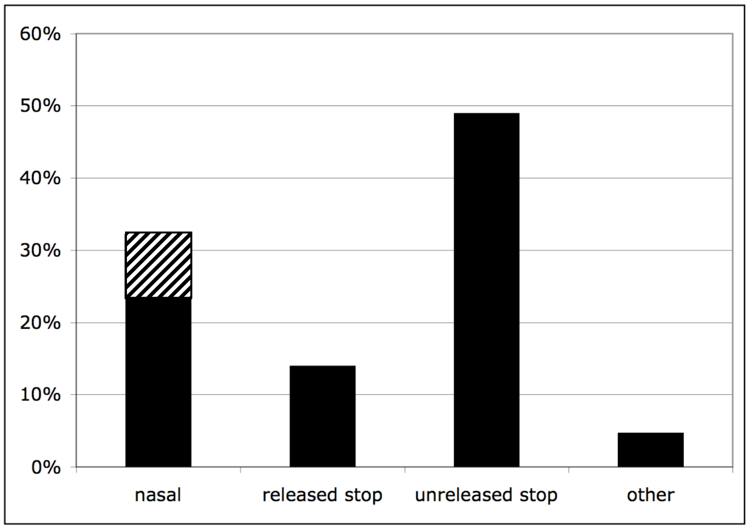


Figure 4. Realization of pre-nasal stops by Korean speakers of English. The patterned bar indicates pre-oralized nasals: final stops that were transcribed as nasals, but for which acoustic analysis found some oral closure duration. See text for discussion.

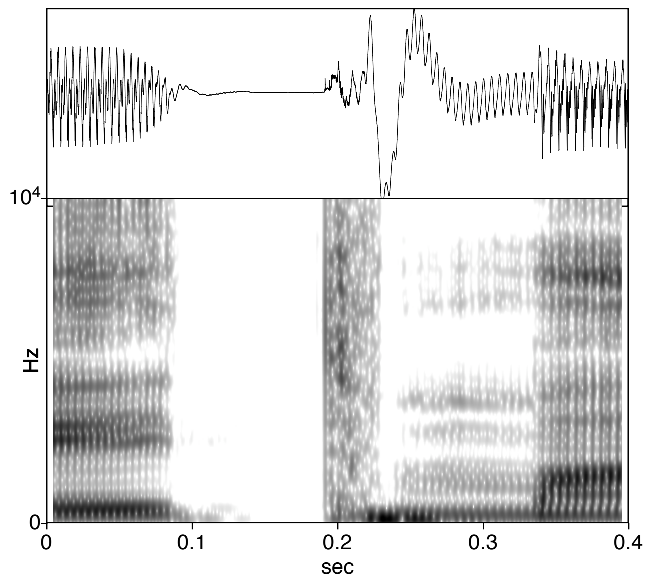


Figure 5. Voiceless coda with release. Speaker K8: ate mine.

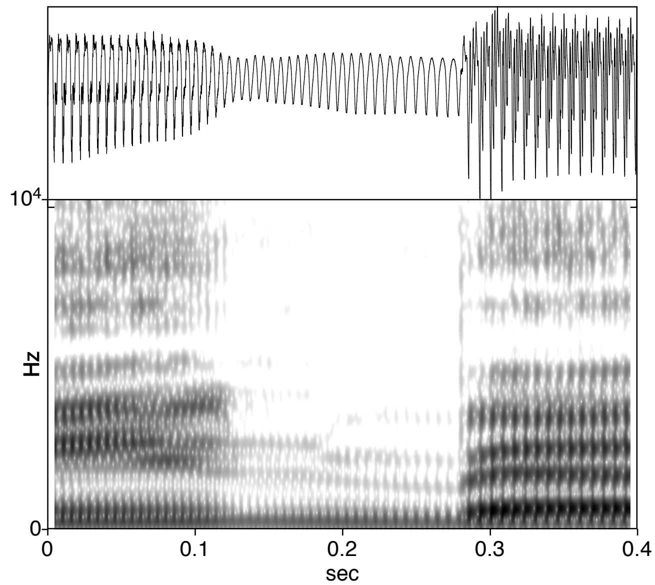


Figure 6. An underlying nasal sequence in Korean English. Speaker K12, from train Matt.

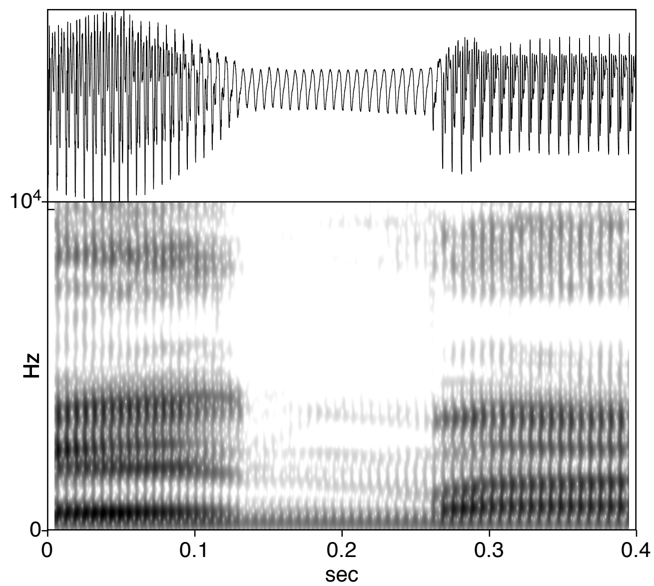


Figure 7. Full nasalization in Korean English. From Speaker K12: ate mine.

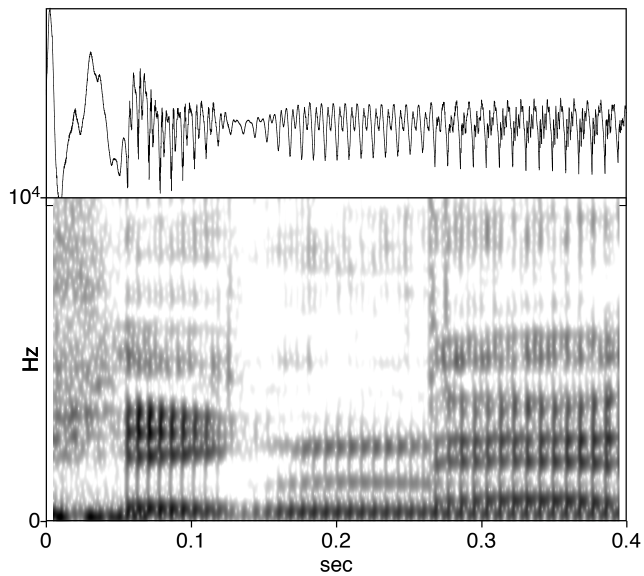


Figure 8. Pre-oralized nasal: Speaker 11: pick Nat.

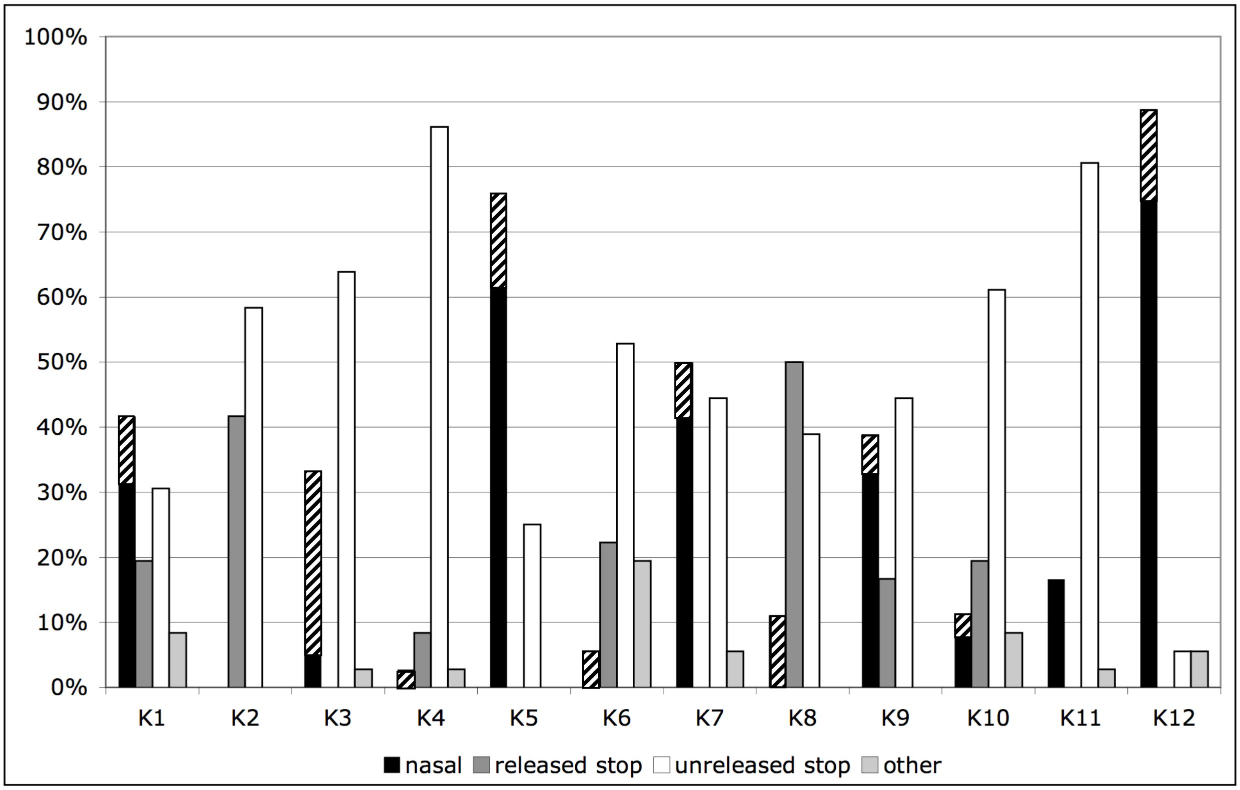


Figure 9. Realization of pre-nasal final stops for each speaker. The patterned bars indicate pre-oralized nasals: final stops that were transcribed as nasals, but for which acoustic analysis found some oral closure duration.

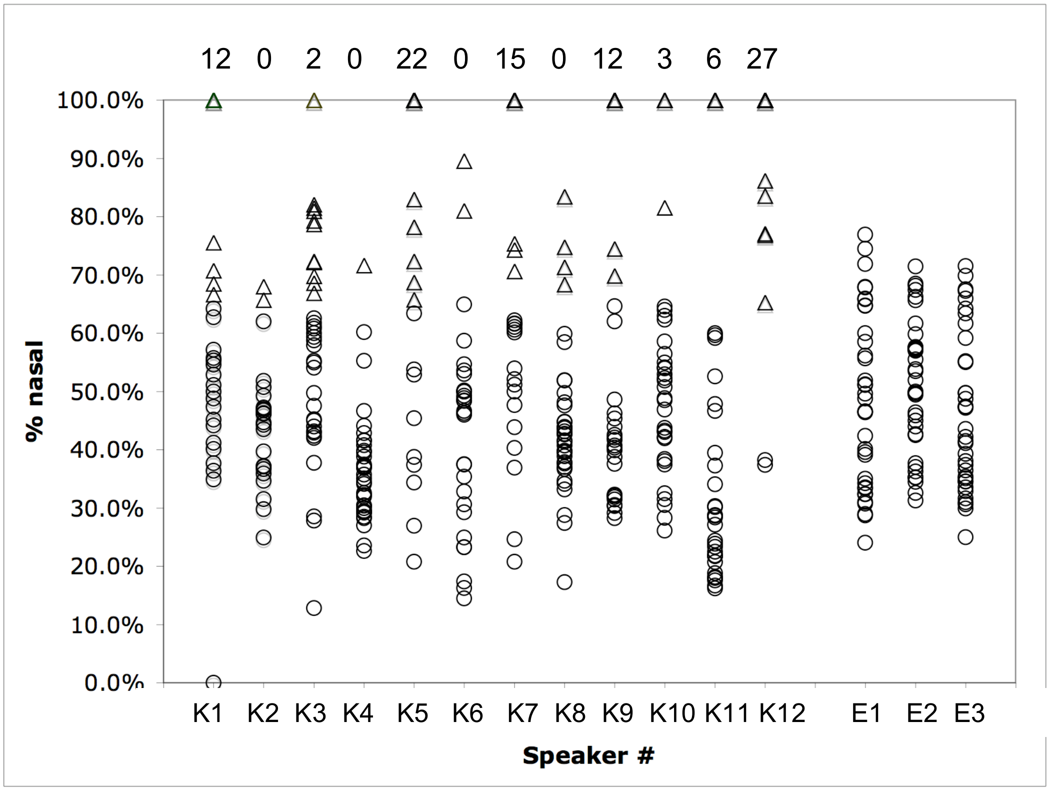


Figure 10. Percent nasalization for each Korean and English speaker. Circles = tokens transcribed as having word-final obstruents; triangles = tokens transcribed as having word-final nasals. Numbers at the top of the graph indicate the number of tokens for each subject where acoustic analysis found 100% nasalization (number of data points superimposed at 100%).

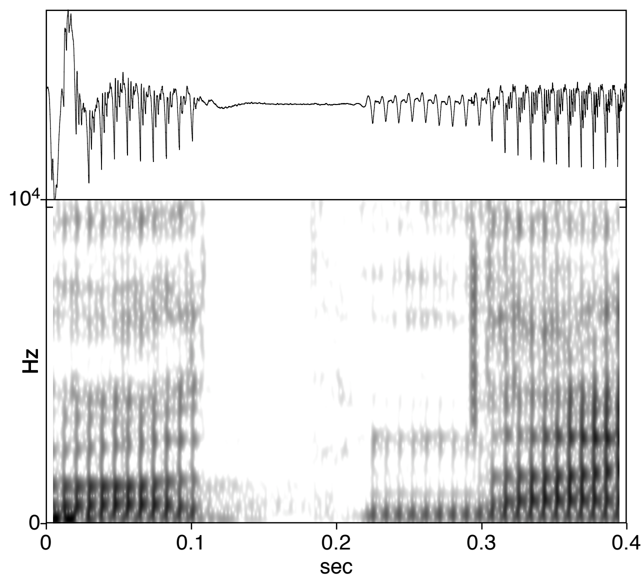


Figure 11. Unreleased voiceless coda in Korean English. Speaker K9: <u>bought nine</u>

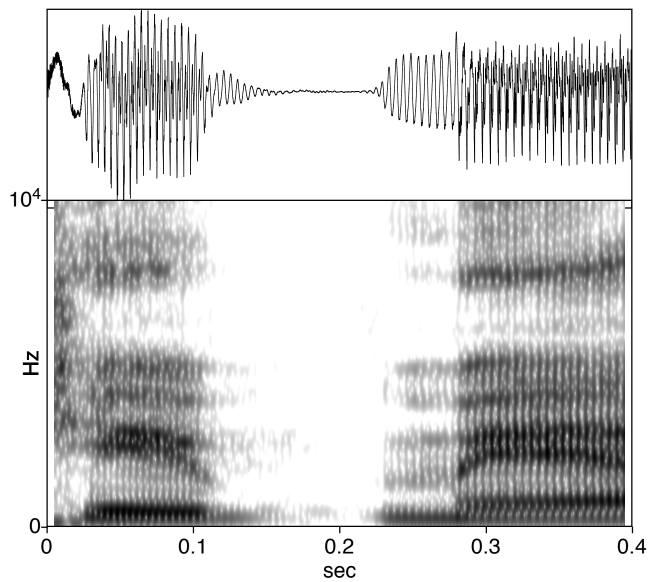


Figure 12. Partially voiced coda in Korean English. Speaker K5: keep Matt

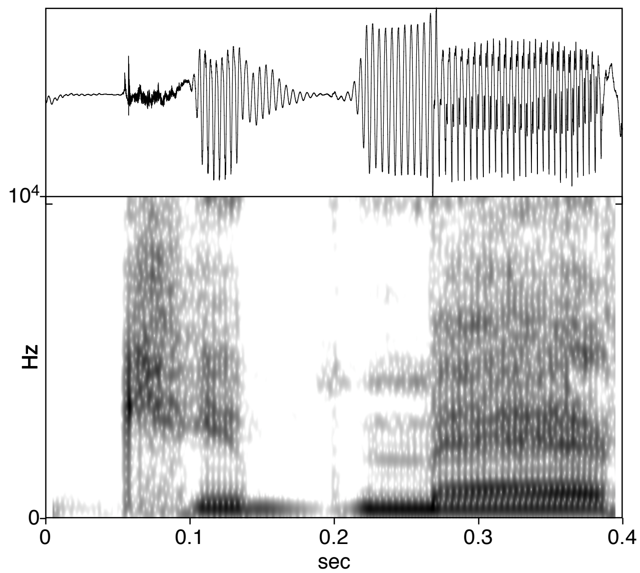


Figure 13. Partially voiced coda in Korean English. Speaker K2: keep Matt

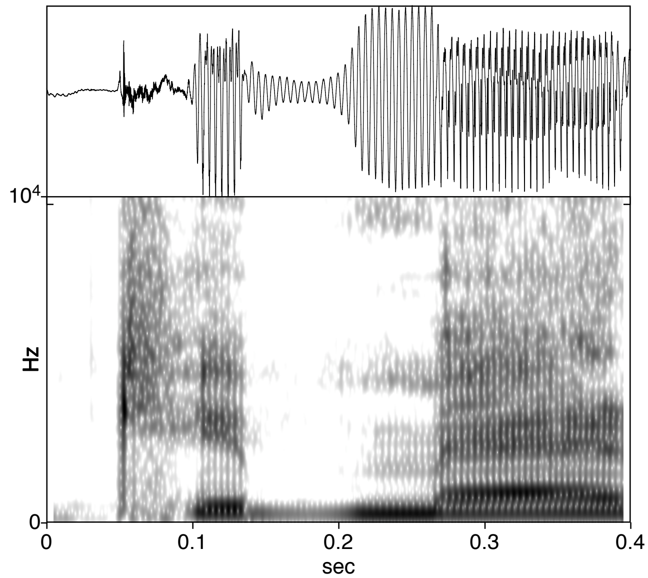


Figure 14. Fully voiced obstruent coda in Korean English, Speaker K2: <u>keep Matt</u>

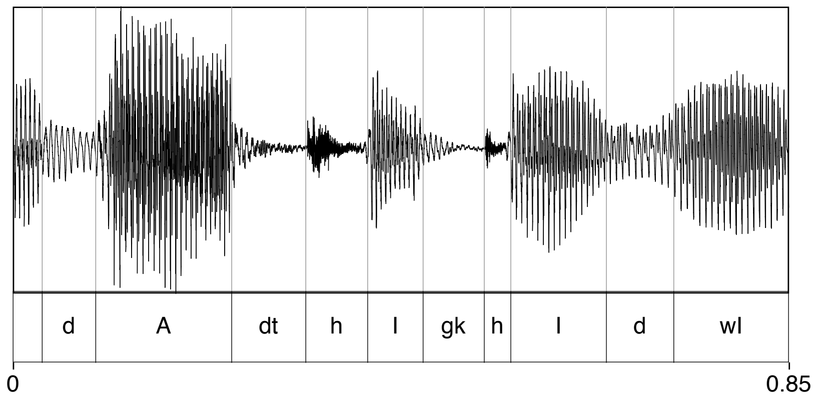


Figure 15. Intersonorant voicing in a non-nasal context: Speaker K3's pronunciation of <u>that ticket will</u>.

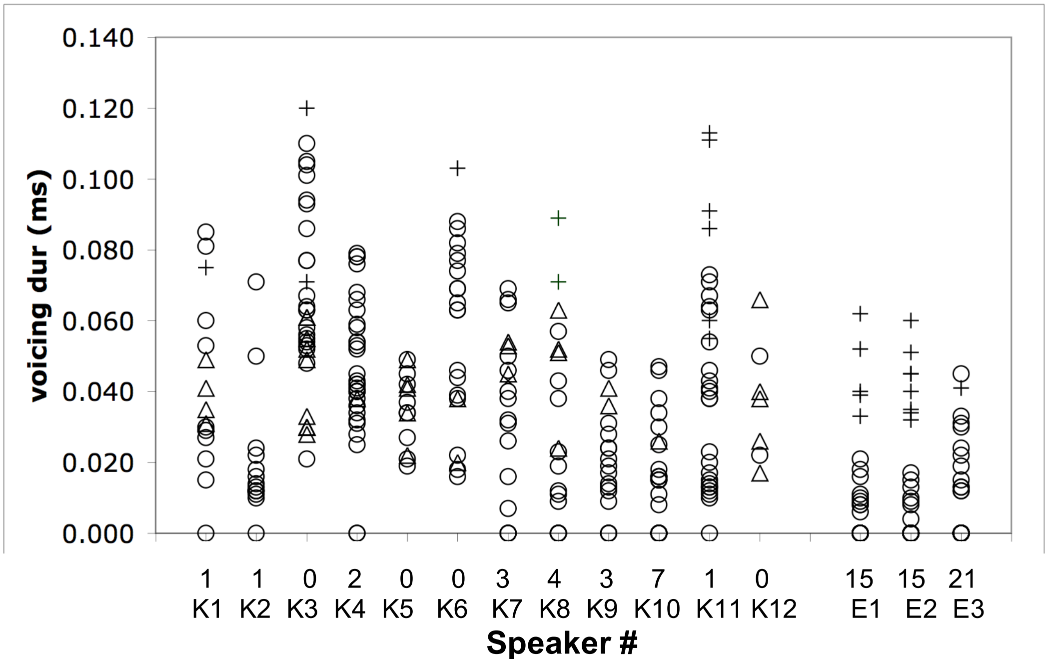


Figure 16. Duration of oral voicing for each Korean and English speaker, including only tokens with some oral closure and no audible release of C1. Circles = tokens with modal voicing; crosses = tokens with creaky voicing; triangles = tokens transcribed as having word-final nasals. Numbers at the bottom of the graph indicate the number of tokens for each subject where acoustic analysis found no voicing during the closure (number of data points superimposed at 0 ms).

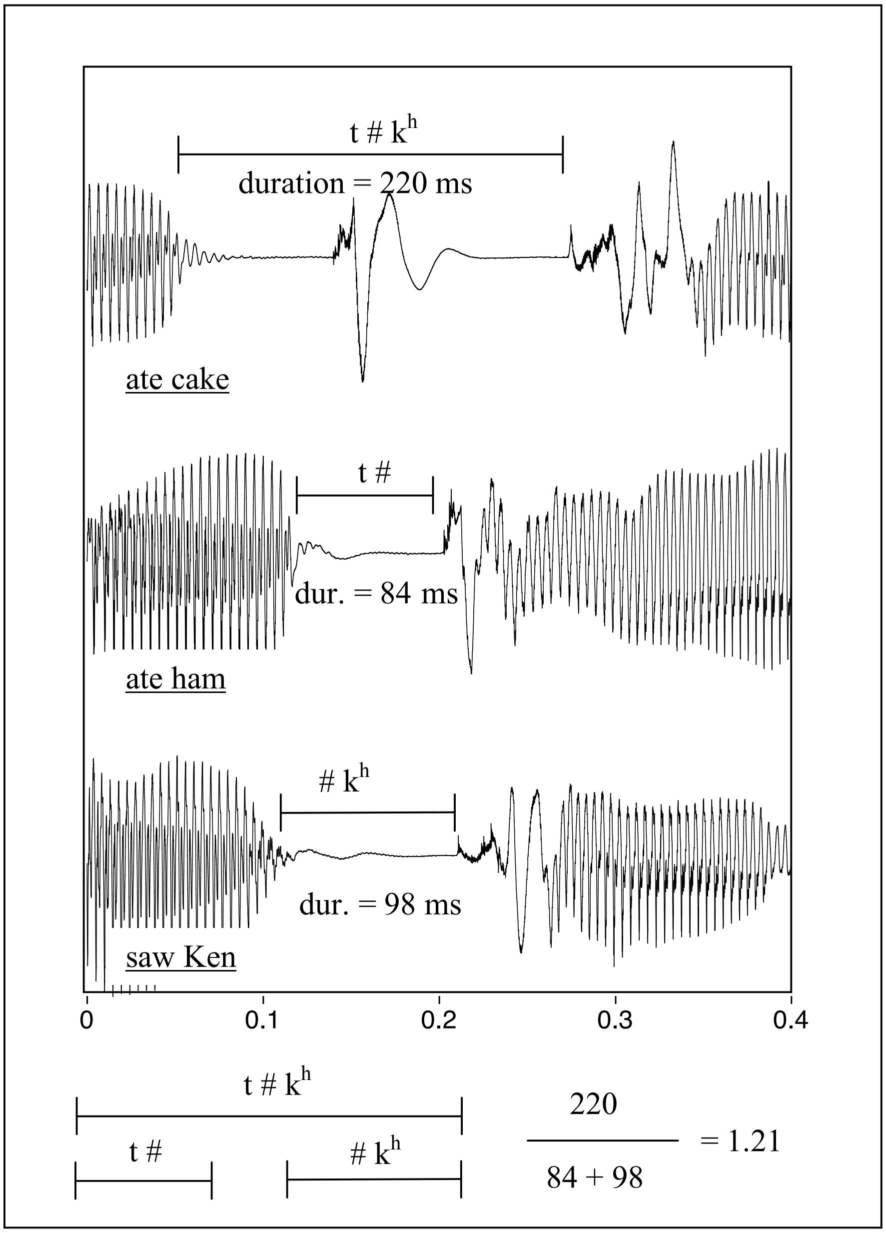


Figure 17. Illustration of duration ratio for a cluster with an open transition (gestural lag). Example phrases from speaker K8. The duration of the cluster is longer than the duration of its parts.

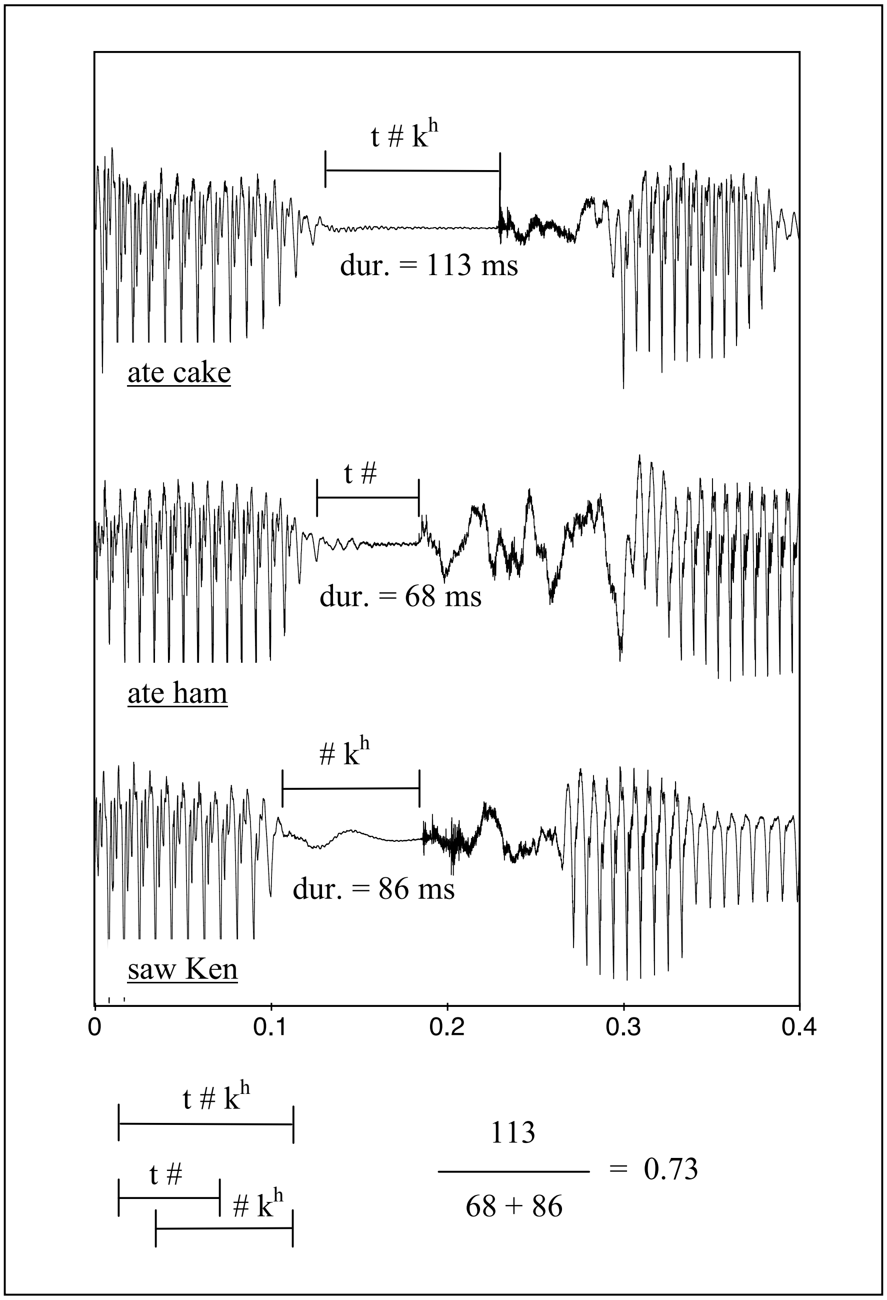


Figure 18. Illustration of duration ratio for a cluster with a close transition (gestural overlap). Example phrases from speaker K9. The duration of the cluster is less than the duration of its parts.

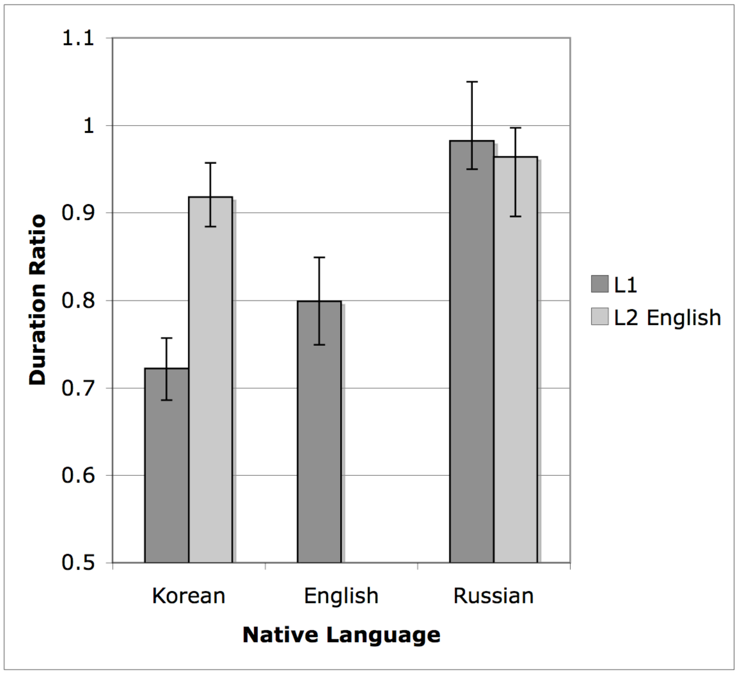


Figure 19. Mean duration ratio in obstruent#obstruent clusters, for 5 L1 and L2 combinations. Error bars = 95% CI.

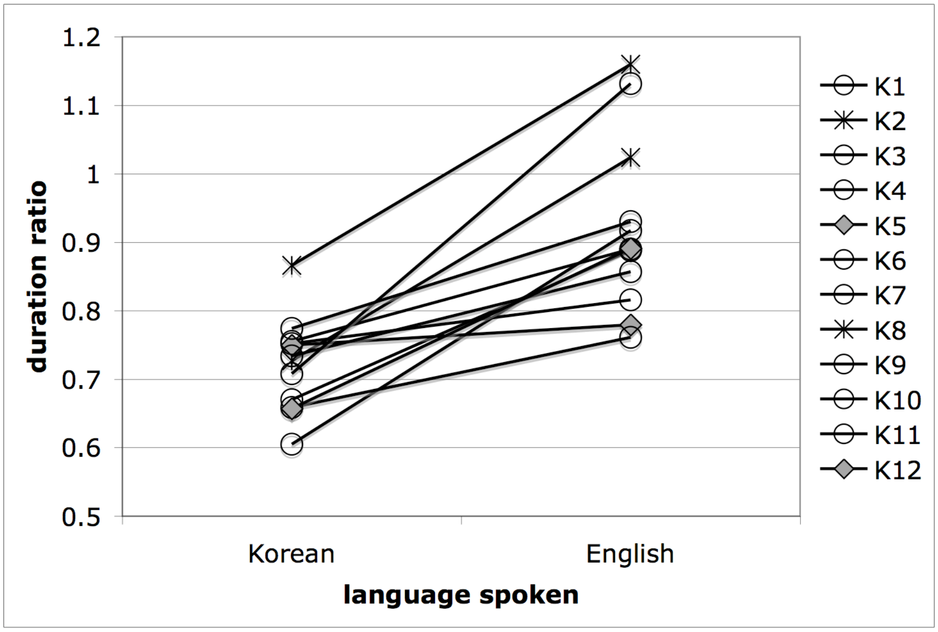
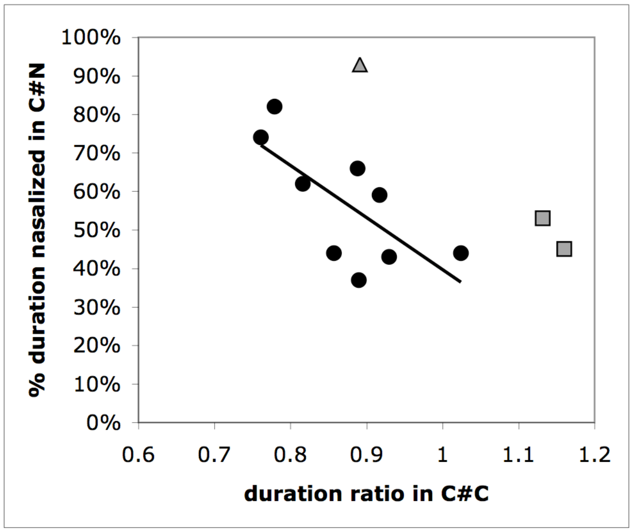


Figure 20. Duration ratios in obsturent#obstruent clusters in native Korean and L2 English, by speaker. Symbols indicate preferred realization of obstruent#nasal sequences in English: Diamond = nasalization; asterisk = audible release; circle = unreleased, with variable voicing and nasality.



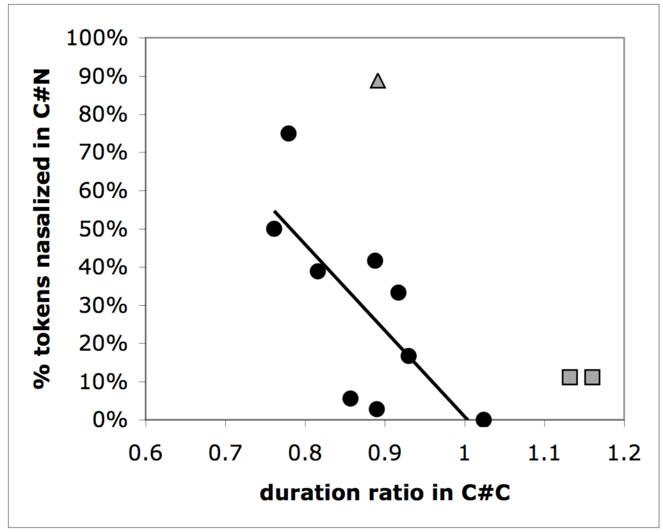


Figure 21: Correlation between nasal assimilation (Experiment I) and overlap (Experiment II), all tokens included. A. Percent of cluster duration nasalized plotted against duration ratio for each Korean English speaker. B. Percent of tokens transcribed as nasal plotted against duration ratio for each Korean English speaker. Triangle = K12, characterized by categorical nasalization. Squares = K8 and K10, characterized by word integrity.

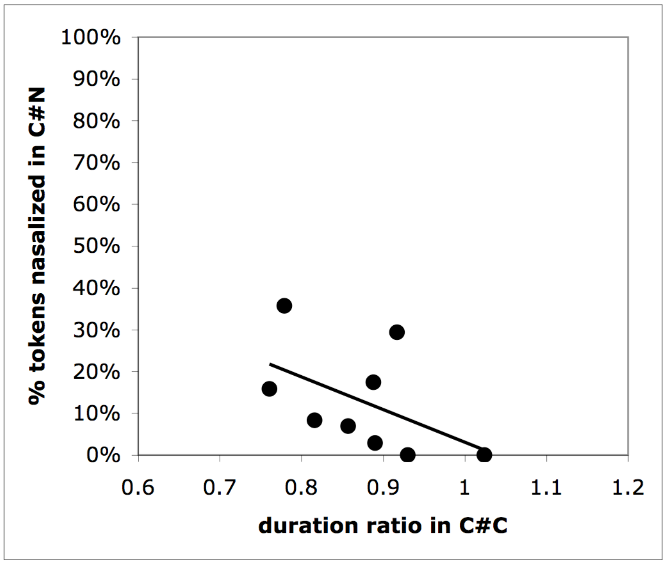


Figure 22: Correlation between gradient nasal assimilation (Experiment I) and overlap (Experiment II). Full nasals are excluded. Percent of tokens transcribed as nasal (that is, pre-oralized tokens) plotted against duration ratio for each Korean English speaker. Speakers K8, K10, and K12 are excluded.

Notes

---

[i] There were no clear instances of epenthesis to break up the sequences of voiceless stop plus nasal in this dataset. When a release was audible, no voicing or formant structure indicative of a vowel was apparent. Previous research on Korean English (Park 2005, Kim 2000, Zsiga & Kim 2005) found that epenthesis was more likely following voiced than voiceless final consonants.

[ii] As a reviewer points out, in order to explicitly test for deletion vs. assimilation, it would be necessary to compare the duration of assimilated clusters with that of singleton nasals. Data on singletons was not collected in the present experiment, however. While the durations are certainly more consistent with assimilation, a deletion analysis can not be definitively ruled out.

[iii] It might be argued that the gaps between 85% and 100% in Figure 10 above are in the eye of the labeler: if the oral portion is short enough, it might be missed by the phonetician's eye as well as her ear, and the token classed as 100% nasal by mistake. The average total duration in the Korean English sequences was 188 ms, 10% of which is 19 ms. For the average male voice (180 Hz), this is space enough for three to four glottal pulses, long enough to detect the presence or absence of the complexity due to nasal resonance. Pulses will be even shorter for female voices. See Figure 8, where the oral portion, with a duration of 17 ms, is clearly evident. For some of the smaller gaps, it may be the case that some tokens in the full nasal category were in fact examples of more extreme overlap rather than categorical reorganization. However, gaps the size of those found for most of the speakers in Figure 10 would not go unnoticed. Further,

even if velum opening began right at consonant closure, a changing waveform amplitude would distinguish the pre-oralized tokens from the full nasals. All of the pre-oralized tokens showed strong voicing throughout the closure, with no breathiness, so the evidence points to the difference in amplitude being the result of variably-delayed velum opening rather than variation in the laryngeal gestures.

[iv] The perceptions referred to here are those of the author and the research assistants who worked on labeling the data. Controlled perceptual experiments with naive participants were not undertaken as part of this study, but are planned for future research.

[v] It is not clear if categorical assimilation *without* close transition would ever be attested. If, as has been hypothesized, the presence of a coupling between oral closing gestures corresponds to membership in the same prosodic domain, that configuration would correspond to assimilation that applied *only* across phrase boundaries. While such a pattern could in principle be represented in a coupling graph (or autosegmental representation, for that matter), extra-grammatical constraints on the processes by which phonological alternations arise and are transmitted might rule out such a disparity. See section 6.2, which provides further discussion of predicted and actual coupling relations. As is noted there, further modeling studies with more extensive data from different languages are needed in order to address these questions.

[vi] The difference between Korean and English was borderline: a Tukey test found that English and Korean were within a homogenous subset, but a Student-Neuman-Keuls test found that they were not. (All tests agree that Russian is significantly different from both.) If the difference

between native Korean and native English is real, it may be due to a greater reduction of coda consonants in Korean than in English, or possible categorical place assimilation in certain clusters (Jun 1995, Kochetov & Pouplier 2008, Son et al. 2007). For the purposes of this analysis, however, the important finding is that native Korean has a pattern of coordination at least as overlapped as native English, if not more.