Gestural Properties of the Non-Target Velar Stop /k/ and Its Variance with Respect to Supralaryngeal Constrictions in C2

Minjung Son†
Hannam University

ABSTRACT

In this study, we examine overlap-induced gestural properties in C1. Specifically, we examine the spatio-temporal properties of the non-target velar stop /k/ in Korean, as a function of three segmental contexts in C2 (pre-/h/, pre-/p/, and pre-/t/), two-boundary conditions, and two speech-rate conditions. The results show that gestural overlap with different supralaryngeal constrictions in C2 is positively correlated with the constriction duration of /k/, and is negatively correlated with the constriction maxima. Different constrictions in C2 distinctly affect the spatio-temporal properties of C1. The pre-/t/ context leads to distinct patterns in the C1 velar stop, compared to the pre-/h/ and pre-/p/ contexts. The pre-/t/ velar stops observed in most combinatoric phonological contexts have shorter constriction durations, and have greater constriction maxima for /k(#)t/. Greater constriction maxima is also observed for the within-word condition as well as for fast speech rate. The contradictory spatio-temporal properties of /k/ seem to be determined by physiological constraints on a consecutive lingual-lingual sequence, and by a higher jaw position facilitated by a coronal.

Keywords: normalized gestural overlap, gestural reduction, jaw height

1. Introduction

Typologically, if word-initial position allows for consonantal sequences, so does intervocalic position. For English, consonant sequences may contain up to three in the onset (#sC(L)V), while the same sequence can be derived postlexically (s#C(L)V) (e.g.,/leɪ # **sp**ɪn/'lay spin' vs./leɪ**s** **p**ɪn/'lace pin' (Cho et al. 2014)). In Georgian,

* The work was fully supported by NIH Grant DC-00403 conferred upon Catherine T. Best (PI) and Haskins Laboratories. I am grateful to eight speakers for voluntarily participating in EMMA experiments, Sean C. O'Rourke for proofreading of this paper, and two anonymous reviewers for their constructive suggestions and valuable comments. All remaining errors are my own.

† Corresponding author: minjungson@hnu.kr

stop-stop sequences appear word-initially as well as word-internally (e.g.,/**bg**era/ ‘sound’ vs. /a**bg**a/‘saddle bag’ (Chitoran et al. 2002)). With respect to intervocalic C1C2 sequences in general, they can be phonetically realized the same as in the phonological representation. However, they can further undergo phonological processing such as for place assimilation motivated by ease of articulation (Jun 1995, 1996).

Focusing on intervocalic stop-stop sequences, it is possible for an unreleased C1 stop in coda to be perceptually weaker. However, phonetic cues for place of articulation in this context are available during vocalic transitions or release burst (Borden & Harris 1984), which are simultaneously used in speech perception. Phonetic cues of a C1 stop can be relatively more vulnerable to phonological alternation since they can be degraded either by gestural overlap with C2, gestural reduction of C1, or both. With respect to segmental context effects, previous studies have shown that assimilation is attributed to relatively weak perceptual cues for the place of articulation in C1 compared to C2 in VC1C2 sequences (Malécot 1958; Hoseholder 1956; Ohala 1990)—perceptual cues of VC1 formant transitions can be easily masked by the following C2. With regard to relative perceptual saliency of onset, CV has more perceptual cues for place of articulation than VC since the former (CV) has a burst of noise for stop consonants during transition to the nucleus, as well as the formant transitions (Stevens 1989, 1997).

Among three places of articulation (labial, coronal, and dorsal), coronal stops undergo anticipatory place assimilation most frequently according to Jun’s (1995) typological study (e.g., /mit#bɔl/→[mitbɔl] ~ [mi**pb**ɔl], ‘meat ball’ in English (Jun, 1995); /sen#mans/→[se**mm**ans] ‘seven hands’ in Catalan (Mascaró 1978 from Jun 1995); [mitbɹɪŋən]→[mi**pb**ɹɪŋən] ‘to bring along’ in German (Kohler 1990 from Jun 1995); /mitko/→[mitko] ~ [mi**kk**o], ‘believe and’ in Korean (Jun 1995)), followed in turn by labial targets, which, in turn, is followed by velar targets. In his study, velar is least likely to be the target of place assimilation: if velar is such a target within a language, the other two places of articulation (i.e., coronal and labial) also become the target of place assimilation (Jun 1995, 1996).

As a basis for weakened perceptibility of the place of articulation in coda, which is characterized by weak VC1 formant transitions and lack of burst noise, previous articulatory studies attributed this to gestural overlap (Browman & Goldstein 1986, 1989, 1990, 1991, 1992, 1995; Byrd 1992, 1996; Chen 2003; Son 2008; Surprenant & Goldstein 1998). In Chen’s (2003) gestural simulation study based on X-ray microbeam data, dynamic parameters (e.g., activation interval, target, damping, and stiffness) of relevant tract variables (tongue tip constriction degree (TTCD) and lip aperture (LA)) were specified. One set of gestural simulation data was generated by varying degrees of gestural overlap between C1 and C2, and the other set by varying gestural reduction of C1. The gestural scores generated were submitted to the task-dynamics model (Saltzman 1995), from which the resulting time function became the input for the Haskins articulatory synthesizer (Rubin et al. 1981). Using the output of the gestural simulation as the input to the listener recovery algorithm, she furthermore tested how temporally more overlapped final stops are recovered perceptually. In her paired gestural simulation and recovery algorithm studies, C1 coronal stop /d/ was less likely to be fully recovered in in coronal-labial (/d#b/) sequences with greater gestural overlap, which was compatible with the result of gestural reduction. Meanwhile, C1 non-targets in labial-coronal (/b#d/) sequences were more consistently recovered with lip constriction by the listener algorithm irrespective of varying gestural overlap or different degrees of C1 constriction: labial stop /b/ in the non-assimilating context /b#d/ was more reliably recovered as the underlying labial. Putting the results together, Chen (2003) concluded that listeners' weakened perceptibility of coronal coda in VC1C2V admitted to gestural overlap, which in turn influenced speakers towards misproduction of greater overlapped targets as reduced.

However, in Cho and McQueen (2008), the results of human listeners' perceptual decision of C1 targets in assimilating contexts were the opposite of results of perceptual decision making in Chen's (2003) listener recovery algorithm. Cho and McQueen (2008) found that the phonological knowledge of place assimilation helped listeners to perceptually reconstruct the underlying form of C1, since they are better with C1 coronal reconstruction (e.g., /t/-detection in response to hearing [VkkV]), compared to C1 labial (e.g., /p/-detection in response to hearing [VkkV]). This means that Korean listeners were able to reconstruct the underlying form of coronal stop /t/ better since they possess phonological knowledge of place assimilation such that coronal stop was more likely to be gradiently/categorically assimilated to C2 velar (or labial). As soon as they heard [kk] sequences, they might have instantly accessed the phonological representation /tk/, when dealing with possible coronal target of place assimilation. Secondly, in a phoneme monitoring task of non-assimilating /k#t/ sequences, C1 velar stop was detected much faster when it was perceived as [k] and C1 non-target velar stop was detected much slower when it was perceived as [t] (e.g., hearing assimilated sequences that are disallowed), compared to C1 labial stop in non-assimilating /p#t/ sequences. The results were taken to indicate that the phonetic cues for velar stop (e.g., a velar pinch) are strong perceptual cues, which might have in turn exempted a velar segment from diachronic phonological alternation (e.g., place assimilation).

Further evidence for gestural reduction of C1 triggered by gestural overlap was found in some articulatory studies (Browman & Goldstein, 1995; Kühnert & Hoole, 2004; Nolan, 1992; Surprenant & Goldstein, 1998; Son, 2008). Kühnert & Hoole (2004) examined kinematic tongue movement of voiced coronal (/d/) in the pre-/k/ assimilating context and the pre-/h/ non-assimilating context, using a simultaneous electropalatography (EPG) and electromagnetic articulograph (EMA) study. Testing with speakers of British English and German, greater temporal reduction of the tongue tip movement was consistently observed. Browman and Goldstein (1995) also looked into kinematic movement of stop consonants. They analyzed voiceless stops (/p/, /t/, /k/) from the University of Wisconsin (WS) X-ray microbeam (XRMB) study where either one of the three target words ('pop', 'tot', 'caulk') was embedded in a carrier intonational phrase ('my ___ huddles/puddles/tuddles'). The results showed that three segments (/p/, /t/, and /k/) in coda were spatially more reduced than those in onset. In addition, coronal exclusively showed more reduction in the pre-/p/ assimilating context, compared to the pre-/h/ non-assimilating context; labial and velar did not demonstrate this contextual sensitivity.

As a possible attribute for the relatively weak perceptibility, syllable position effects were proposed to account for regressive directionality (Fujimura et al. 1978; Jun 1995, 1996, 2004; Steriade 2000, 2001): coda is the target of assimilation, not onset. For English, reduction has been widely observed in the temporal and spatial magnitude of all oral gestures in coda as compared to onset (e.g., stop (/p/, /t/, and /k/); sonorant (/m/, /l/, /j/, and /w/) (Browman & Goldstein 1995; Gick 2003; Giles & Moll 1975; Krakow 1989, 1999)). In an articulatory study of intervocalic velar stop /k/, Son (2011b) found that syllable-final reduction was observed, providing a phonetic basis for coda reduction of C1 even in intervocalic C1C2 sequences. Furthermore, pre-stop contexts, compared to the pre-/h/ contexts, can induce weakened perceptibility given their effect on release bursts and formant transitions in VC. Note that this comparison can assume *a priori* that the pre-/h/ context in word-medial position should not, at the very least, be an articulatory weakening position. With respect to intervocalic, lexically aspirated stops in Korean, Cho and Keating (2001) showed that word-medial is an articulatory lengthening position, compared to word-initial position, where domain-initial strengthening (e.g., word-initial strengthening) does not prevail. In an acoustic study on underlying lenis-/h/ sequences and lexically aspirated stops, Jang (2006) showed that no durational difference in aspiration in accentual phrase-medial position was detected. Taken together, if assimilation in pre-stop contexts is caused in part by a general tendency towards reduction in those contexts, we should also be able to see reduction in C1C2 contexts that do not trigger place assimilation. If we do not see this, then we may be able to conclude that reduction is caused solely by grammatical knowledge that some environments are assimilating contexts. In particular, we compare Korean velar stop /k/ in non-assimilating contexts as we vary C2 with different degrees of constriction.

1.1. Research questions

Firstly, we explore whether gestural overlap in C1(#)C2 varies with C2 (/p/ and /t/) in Korean non-assimilating contexts. In particular, we examine whether gestural overlap in lingual-lingual (/k(#)t/) sequences is less overlapped as compared to lingual-labial (/k(#)p/) sequences. In Kochetov et al. (2007)'s cross-linguistic study on Korean and Russian, they examined three non-assimilating stop-stop contexts (e.g., /k(#)p/, /k(#)t/, /p(#)t/) in two morphosyntactic conditions (across-word vs. within-word) as well as two speech rate conditions (comfortable vs. fast). For Korean, less overlap was consistently observed in back-to-front (/k(#)p/ and /k(#)t/) sequences, compared to front-to-back (/p(#)k/) sequences; this was attributed to perceptual recoverability constraints. Rate effects were only observed in constriction duration, indicating a shorter temporal interval in the fast speech rate (comfortable>fast). Son's (2011a) articulatory study observed similar degrees of gestural overlap between word-internal /kt/ and /pt/ sequences. However, these previous articulatory studies had limited the scope of study to investigating some temporal properties, thus not including spatial properties or the constriction duration of C1 non-targets as a function of sequence types. In the current study, we start to look into the C1 non-target velar stop /k/ dependent on a following segment in C2 with different supralaryngeal constrictions (the pre-/p/ context with a labial constriction vs. the pre-/t/ context with a lingual constriction) for reduction purposes. We systematically examine spatial as well as temporal aspects of the non-target /k/ in the three non-assimilating contexts. Also included were two morphosyntatic conditions (across-word vs. within-word) and two speech rates (comfortable vs. fast) so that we can provide a balanced analysis within a given language.

Secondly, we are concerned with determining whether, and if so, how spatio-temporal properties of C1 non-target velar stop /k/ in non-assimilating contexts varies with different constriction degrees of C2 (/h/ with no supralaryngeal constriction; /p/ and /t/ with supralaryngeal constrictions) as we examine the kinematic properties of C1. Note that for American English, the spatial magnitude of a C1 non-target velar stop did not differ as a function of different degrees of gestural overlap in non-assimilating sequences (e.g., /k#h/=/k#p/, 'my chaulk huddle' vs. 'my chaulk puddle' in Browman & Goldstein (1995)), while context-dependent spatial reduction was observed in assimilating contexts (/t#h/</t#p/) (e.g., 'my tot huddle' vs. 'my tot puddle'). Given that the pre-/h/ context is not a condition for spatial reduction of C1 non-targets for American English under any circumstances, it is of interest to determine how similar or different spatio-temporal properties of C1 non-target velar stop /k/ in Korean are. In the current study, we also employ three different segmental contexts (pre-/h/ vs. pre-/p/ vs. pre-/t/) while continuing to examine the constriction duration of non-target velar stop /k/, constriction maxima of C1, and vertical jaw maxima spanning across C1C2 sequences so as to discuss the implications of articulatory findings in Korean.

2. Method

2.1. Stimuli and data collection

A total of five subjects participated in the production experiment, but data from only five of the subjects are used for further analysis.[1] In the current study, velar stop /k/ is compared in the pre-/h/ and the pre-stop contexts (pre-/p/ and pre-/t/). A total of ninety-six tokens were acquired from each subject (3 segmental contexts (pre-/h/, pre-/p/, pre-/t/) × 2 speech rates (comfortable vs. fast) × 2 boundaries (across-word vs. within-word) × 8 repetitions). The presumed syllabification in the phonetic form is indicated by a dot and the complete set of stimuli set is shown in (1).

The stimuli in Korean are prepared as short ordinary sentences. In (1.a), the across-word condition is shown for transliterations, phonetic forms, and glosses.

1) The original data collection included two non-assimilating targets, velar stop /k/ and labial stop /p/, with eight speakers. Three subjects, however, mispronounced /...apta.../ as [...ap*ha*ta...] by inserting [ha]. This was not properly monitored during the experiment sessions and we excluded data for these three speakers from further analysis.

Likewise, the within-word condition is shown in (1.b). The target sequences are in bold and broad phonetic transcription is used. The complete list of stimuli appears in Appendix.

(1) Stimuli list

(a) Across-word condition

Transliteration	Phonetic form	Gloss
(i) cɔnha**k** **h**acamaca	[cɔn.ha.**kʰ**a.ɟa.ma.ɟa]~ [cɔn.ha**k̚**.**h**a.ɟa.ma.ɟa]	'as soon as transferred (from the old school to the new one)'
(ii) cɔnha**k** **p**atci	[cɔn.ha**k̚**.**p**at.ci]	'(We do not) accept (any) transferred (students)'
(iii) ɔha**k** **t**aɨmɨlo	[ɔ.ha**k̚**.**t**a.ɨ.mɨ.lo]	'next to the verbal part'

(b) Within-word condition

Transliteration	Phonetic form	Gloss
(i) cɔnha**kh**ataka	[cɔn.ha.**kʰ**a.da.ga] ~ [cɔn.ha**k̚**.ha.da.ga]	'during transfer (from one school to another)'
(ii) cɔnha**kp**anenɨn	[cɔn.ha**k̚**.**p**a.ne.nɨn]	'class for the transferred' – NOM.
(iii) ɔha**kt**apɨn	[ɔ.ha**k̚**.**t**a.bɨn][2]	'answers for the verbal part' – NOM.

We collected kinematic data pertinent to articulatory movement of speech production using a two-dimensional electromagnetic midsagittal articulometer (EMMA in Perkell et al. 1992) at Haskins Laboratories in New Haven, Connecticut. This 2D point-tracking technique records of x and y values from electric transducers (a.k.a. pellets or receiver coils) glued to seven different articulators: one on the the lower incisor, one on the upper lip, one on the lower lip, one on the tongue tip, two on the tongue body, and one on the tongue dorsum (see Löfqvist (1993), for a full description of this point-tracking system). After sampling at 200 Hz, articulatory data were further smoothed by a low-pass filter of 20 Hz in post-processing procedures completed with Matlab software. Using a Sennheiser shotgun microphone, the acoustic data were also acquired while collecting kinematic data. In the current study, the scope of analysis was limited to kinematic characteristics of the tongue dorsum, tongue tip, and lip articulators to examine three C1(#)C2 sequences (/k(#)h/, /k(#)p/, and /k(#)t/). In particular, we used vertical movement of the

2) A subset of data, /kt/ sequences in the within-word condition was used in Son (2011a).

tongue dorsum (TD-y) for /k/, tongue tip constriction degree (TTCD) for /t/ (taking into account both the vertical and horizontal movement of the tongue tip articulator), lip aperture (LA) for /p/ (using the Euclidean distance between the upper and lower lip), and vertical movement of the jaw (Jaw-y) for C1C2.

2.2. Measurements

Seven gestural landmarks (i.e., movement onset, peak velocity of the formation duration, target attainment, constriction maxima/minima, constriction release onset, peak velocity of the opening duration, and movement offset) are semi-automatically marked using the function of *lp_Snapex* in MVIEW (Tiede, 2005). For the purpose of the current study, we take into account the first five gestural landmarks and acquire 480 tokens for further analysis. However, if there is no obvious movement trajectory showing vertical jaw maxima during C1C2 sequences, we alternatively measure its value as aligned with directly relevant time points of non-target velar stop /k/ in the pre-/h/ context, and we do not apply this if C2 is articulated with supralaryngeal constrictions. A total of 459 tokens are available for further analysis.

With respect to overlap measures, although, to the best of our knowledge, there has not been a systematic study to compare different gestural overlap measures, normalized estimation has been used more prevalently than absolute estimation, in more recent articulatory studies (Marin & Pouplier, 2014; Pouplier et al. 2017, among others). We adopt the normalized estimation method used in Pouplier et al. (2017) (e.g., plateau overlap normalized by temporal intervals between the peak velocity of the formation duration and the constriction offset) and modified so as to reflect a better description of C1 from our data (e.g., plateau overlap normalized by temporal intervals between the movement onset and the constriction offset).

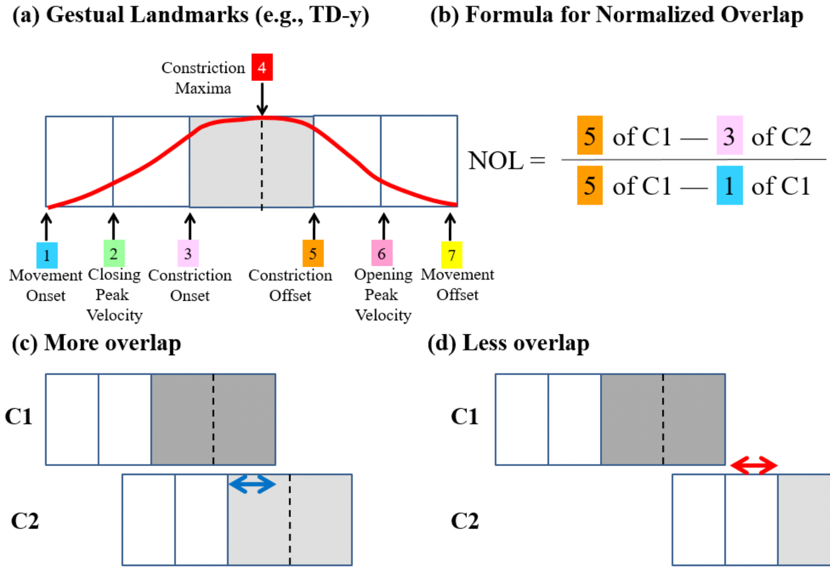

Figure 1. (a) indicates the schematic representation of seven demarcated landmarks with which a simplified vertical tongue dorsum movement trajectory is superimposed. (b) gives a formula with which normalized gestural overlap is calculated—this formula will give positive values in the case of (c) (more overlap or smaller constriction plateau lag) and it will give negative values in the case of (d) (less overlap or greater constriction plateau lag). More overlap is represented by the diagram in (c) and less overlap in (d) based on normalized overlap measurements.

2.3. Statistical analysis

We carry out linear mixed effects models in R, considering individual participant differences (R Development Core Team, 2014). The results from articulatory examination, converted to z-scores, are fitted with the lmer function from the lme4 package (Bates et al. 2011). We use CC sequence types (/kh/, /kp/, /kt/), boundary types (across-word vs. within-word), and speech rate (comfortable vs. fast) as fixed factors, with subjects as a random factor. Post-hoc analysis is done with Tukey HSD tests. The car package is also used: the pairs function is used to generate scatter plots and the cor.test function to estimate Spearman's rank correlation coefficient (rho (ρ)) using a non-parametric measure of rank correlation.

3. Results

3.1. Gestural overlap

3.1.1. Lingual-lingual sequences in comparison with lingual-labial sequences in non-assimilating contexts

There is an interaction between Sequence type and Speech rate ($\chi^2(1)=16.67$, $p<0.0001$): greater overlap in the fast speech rate is observed in /k(#)p/ sequences (comfortable<fast) and greater overlap in the comfortable speech in /k(#)t/sequences (comfortable>fast).

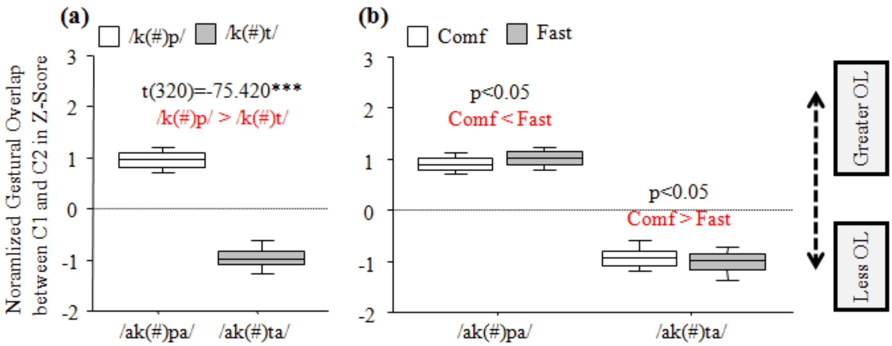

Figure 2. Normalized gestural overlap of C1C2 sequences measured with (a) Sequence type and (b) Sequence type by Speech rate. Greater values represent greater gestural overlap. (The symbol '***' represents $p<0.0001$.)

3.1.2. Correlation between gestural overlap of C1C2 and spatio-temporal properties of non-target velar stop /k/ in C1

In Figure 3, using a correlation coefficient (Spearman's rho for ranked data, also known as a non-parametric test), normalized gestural overlap is positively correlated with the constriction duration of non-target velar stop /k/ in pre-/p/ and pre-/t/ contexts, indicating that greater gestural overlap is related to longer constriction duration of C1 ($r_s=0.557$, $p<0.0001$). Normalized gestural overlap is negatively correlated with the constriction maxima of velar stop /k/ in C1, suggesting greater gestural overlap is associated with greater gestural reduction ($r_s=-0.394$, $p<0.0001$).

However, the formation duration of C1 is independent of normalized gestural overlap of C1C2 sequences ($r_s = -0.075$, $p>0.05$).

Based on the results, we continue to investigate whether, and if so how, constriction duration and constriction maxima vary with three fixed factors (e.g., Sequence type, Boundary type, and Speech rate). In particular, we examine whether the relative articulatory strength of C1 is dependent on the constriction degree of a following segment in non-assimilating contexts (e.g., pre-/h/ (i.e., no constriction), pre-/p/ (i.e., non-lingual constriction), and pre-/t/ (i.e., lingual constriction)).

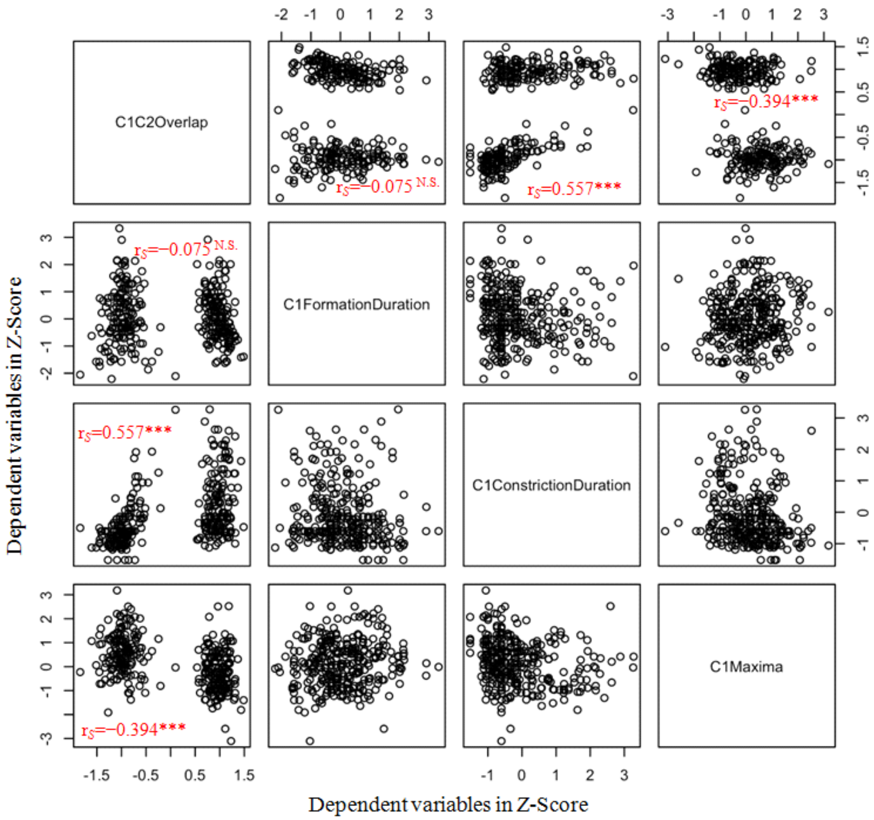

Figure 3. Scatter plot of the normalized gestural overlap of C1C2 (/k(#)p/ and /k(#)t/) sequences with respect to formation duration, constriction duration, and constriction maxima of velar stop in C1.

3.2. Constriction duration of non-target velar stop /k/ in C1

There is a three-way interaction among Sequence type, Boundary type, and Speech rate ($\chi^2(7)=14.34$, $p<0.05$). In Figure 4, the results of post-hoc tests using Tukey HSD show that the constriction duration of velar stop /k/ in C1 is consistently shorter in a consecutive lingual-to-lingual movement in the comfortable speech rate ((/k(#)h/=/k(#)p/)>/k(#)t/). In the fast speech rate, the constriction duration of velar stop /k/ in C1 does not differ between the pre-/h/ context and either the pre-/p/ or pre-/t/ context (/kh/=/kp/; /kh/=/kt/). In contrast, the constriction duration of velar stop /k/ in C1 does not vary with different segmental contexts across the board in the within-word condition of the fast speech rate (/kh/=/kp/=/kt/). Turning to the main effects of different boundary types, the within-word condition exhibits longer constriction duration of non-target velar stop /k/ in C1, compared to the across-word condition, with an increase of 0.22 in z-score (across-word<within-word) ($t(480)=2.68$, $p<0.01$). For different speech rates, the fast speech rate condition provides for a shorter constriction duration, with a decrease of 0.59 in z-score, in comparison to the comfortable speech rate (comfortable>fast) ($t(480)=-7.38$, $p<0.0001$).

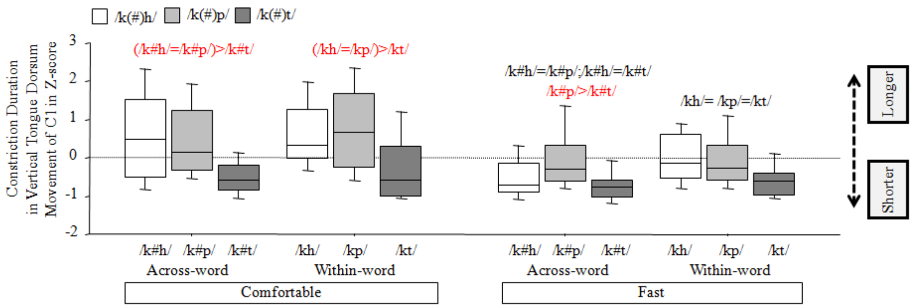

Figure 4. Constriction duration of velar stop /k/ in C1 measured with Sequence type interacting with Boundary type and Speech rate. Larger values in z-scores are for longer constriction duration, with smaller values for shorter constriction duration.

3.3. Constriction maxima of non-target velar stop /k/ in C1

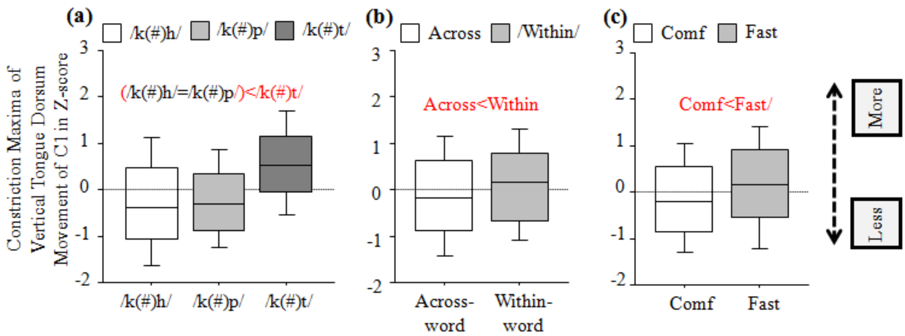

Figure 5. Constriction maxima of non-target velar stop /k/ in C1 measured with (a) Sequence type, (b) Boundary type, and (c) Speech rate. Greater values in z-scores represent more constricted vertical tongue dorsum movement and smaller values, less constriction.

There is no interaction among Sequence type, Boundary type, and Speech rate: a model with a three-way interaction is not significantly better than one without such interaction ($\chi^2(2)=1.05$, $p>0.05$). The interaction of Sequence type and Speech rate does not significantly improve the goodness of fit ($\chi^2(2)=0.04$, $p>0.05$). The same applies to Boundary type interacting with Speech rate ($\chi^2(1)=0.67$, $p>0.05$) and Sequence type interacting with Boundary type ($\chi^2(2)=4.01$, $p>0.05$). As we refit the models, the results indicate that the constriction maxima of the vertical tongue dorsum movement vary with Sequence types ((/k(#)h/=/k(#)p/)</k(#)t/) ($\chi^2(2)=87.49$, $p<0.0001$), Boundary types (Across-word<Within-word) ($\chi^2(1)=8.53$, $p<0.01$), and Speech rate (Comfortable<Fast) ($\chi^2(1)=12.33$, $p<0.001$) as shown in Figure 5.

3.3.1. Correlation between the constriction maxima of non-target velar stop /k/ in C1 and jaw maxima of C1C2 sequences

Using a correlation coefficient (Spearman's rho for ranked data), jaw height is positively correlated with the constriction maxima of non-target velar stop /k/ in C1 ($r_s=0.40$, $p<0.0001$) as shown in Figure 6. The results indicate that greater constriction of the tongue dorsum can be attributed to higher jaw position.

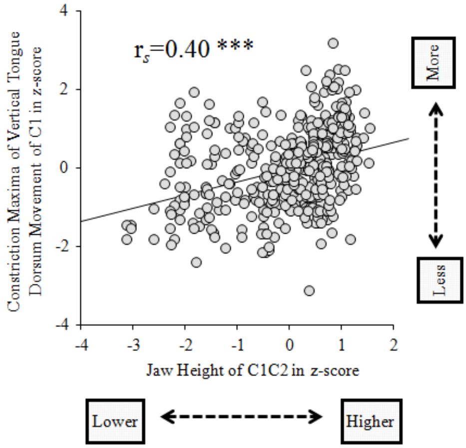

Figure 6. Scatter plot of vertical jaw position in C1C2 sequences relative to constriction maxima of the vertical tongue dorsum gesture in C1. Also shown is the result of Spearman's rank correlation coefficient.

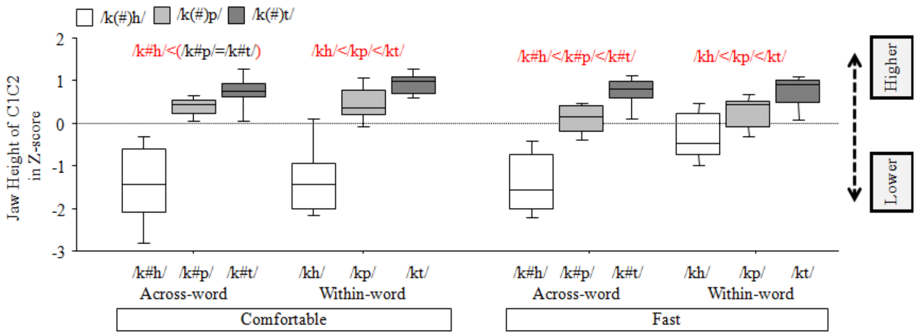

Figure 7. Vertical jaw position in C1C2 measured with Sequence type interacting with Boundary type and Speech rate. Greater values in z-scores represent higher jaw position and smaller values, lower jaw position.

There was a three-way interaction among Sequence type, Boundary type, and Speech rate ($\chi^2(7)=95.26$, $p<0.0001$). In Figure 7, results of the post-hoc tests using Tukey HSD indicate that the jaw height in /k(#)h/ with no constriction in C2 is consistently lower than C1C2 with supralaryngeal constriction across the board. Most of the time, jaw height is also higher in pre-/t/ contexts compared to pre-/p/ contexts, except for in the across-word boundary context at the comfortable speech rate, where a jaw height difference is not detected. Examining main effects, the within-word context is a condition for higher jaw position, with an increase of 0.30 in z-score in comparison with the across-word condition (within-word>across-word) ($t(459)=5.55$, $p<0.0001$). Vertical jaw position does not vary with different Speech rates (comfortable=fast) ($t(459)=1.11$).

4. Summary and Discussion

4.1. Summary

To sum up, firstly, normalized gestural overlap in lingual-labial (/k(#)p/) sequences is greater than that in lingual-lingual (/k(#)t/) sequences. Looking at correlations between normalized gestural overlap and three other variables (i.e., formation duration, constriction duration, and constriction maxima), we learn that there is a positive correlation between gestural overlap and a temporal attribute (e.g., the more overlap, the longer the constriction duration) and a negative correlation between gestural overlap and a spatial attribute (e.g., the more overlap, the more spatial reduction). Secondly, we observe an interaction between Sequence type and Speech rate such that lingual-labial /k(#)p/ sequences showed greater normalized gestural overlap in the fast speech rate while lingual-lingual /k(#)t/ sequences in the comfortable speech rate. Thirdly, spatio-temporal properties of non-target velar stop /k/ show inter-variable and inter-context variability. If temporal reduction occurs, the pre-/t/ position is a context of decreasing tongue dorsum constriction duration in three combinatoric phonological contexts (the across-word/within-word condition of the comfortable rate and across-word condition of the fast rate). Meanwhile, more muscular effort of the dorsal gesture prevails in the pre-/t/ context in terms of constriction maxima across the board ((/k(#)h/=/k(#)p/)</k(#)t/), and the same applies to the within-word condition (across-word<within-word) and the fast speech rate (comfortable<fast). Lastly, the constriction maxima of the dorsal gesture in C1 are positively correlated with the jaw height estimated during the articulation of C1C2: the higher the jaw position, the more constricted the constriction degree of the dorsal gesture. Moreover, jaw position during the articulation of the C1 and C2 reflects more finely tuned positional values in the order /k(#)h/</k(#)p/ </k(#)t/ in most combinatoric phonological contexts, except that a two-way distinction occurs in the across-word condition at the comfortable rate (/k#h/< (/k#p/=/k#t/)).

4.2. Discussion

4.2.1. Cross-linguistic and language-specific spatio-temporal attributes of the dorsal gesture of C1 in Korean non-assimilating C1C2 sequences

In Browman and Goldstein's (1995) X-ray microbeam study, spatial reduction in terms of constriction maxima in coda was more extreme in assimilation contexts than in contexts which do not employ an oral constriction (/t#p/>/t#h/), while this was not observed for non-targets of an assimilation process (e.g., /k/ as C1). Given that gestural reduction as a function of segmental context effects can be understood as gestural overlap, an inducing factor of reduced perception and production (Chen, 2003; Surprenant & Goldstein, 1998), we expect some reduction of C1 to occur in environments having oral constrictions with a stop in C2 regardless of targethood of an assimilation process. However, American English did not demonstrate reduction of non-target stops in environments with a following stop (e.g., /k#p/ in 'my caulk puddles') in comparison with environments without any constrictions (e.g., /k#h/ in 'my caulk huddles'). For Seoul Korean, comparing /k(#)h/ sequences with no constriction in C2 to /k(#)p/ sequences with labial constriction in C2, we found no evidence of spatio-temporal reduction of dorsal non-targets triggered by overlap: this is in congruent with American English.

Lingual-lingual sequences (/k(#)t/) behave differently from /k(#)p/ counterparts: temporal reduction of the dorsal gesture occurred in the pre-/t/ context and spatial strengthening occurred as well. Dorsal data from our current study imply that a physiological constraint (Mooshammer et al. 1995; Kochetov et al. 2007) imposed by a coronal gesture in C2 may have caused durational reduction on the one hand, and more constriction of the tongue dorsum gesture on the other. With regard to earlier temporal completion of the tongue dorsum constriction gesture in /k#t/ sequences (both comfortable and fast speech rates) and /kt/ sequences (comfortable rate only), we conjecture that tongue tip constriction in the alveolar ridge for C2 should influence more advanced displacement of the tongue dorsum due to coarticulation, and an earlier release of the tongue dorsum gesture in the pre-/t/ context can be understood as a consequence of extensive forward movement of tongue dorsum to the extent that the constriction of the tongue dorsum cannot be sustained and therefore not perceived as such by listeners. According to Mooshammer et al.'s (1995) kinematic study on German velar segments, the tongue dorsum constriction gesture against the palate was coupled with forward movement of the tongue dorsum. This forward movement was more enhanced by a high front vowel /i/. In line with this, we assume that the horizontally forward motion could have occurred due to co-production of the following /t/, and this can be further related to narrower constriction. The temporally tighter constriction of /k/ in pre-/t/ contexts might also be accounted for by the forward motion of the tongue dorsum. Note that phonetic cues such as velar place of articulation can be distorted if velar contact is articulated further forward to the extent that forward movement of the tongue dorsum becomes intolerable. In this regard, earlier temporal completion of the tongue dorsum constriction gesture can be taken to be speakers preserving purer velar place of articulation. However, as this is beyond of the scope of the current study, we will leave it for future research.

Next, with respect to greater constriction of the tongue dorsum gesture in C1 before coronals in C2, we learn that there is higher jaw position in this particular context. Looking at a C1 dorsal coda segment in C1C2 sequences, the lingual-lingual (/k(#)t/) sequences avoid gestural reduction of the dorsal gesture, compared to /k(#)h/ sequences with no constriction in C2 and /k(#)p/ sequences with labial constriction in C2. Previous literature has consistently shown that coronal stop (either /t/, /d/, or both) employed higher jaw position (Keating, 1991; Keating et al. 1994; Mooshammer et al. 2007; Son et al. 2011). Due to gestural overlap between the tongue dorsum gesture for /k/ in coda and the tongue tip gesture for /t/ in onset, this can give rise to higher jaw position for the adjacent tongue dorsum gesture in C1, which in turn results in greater constriction degree (r_s=0.40, p<0.0001). Therefore, we conclude that inter-gestural overlap-driven gestural reduction in C1 cannot be considered independently of other factors such as jaw height (e.g., a possible factor to facilitate articulatory strengthening) to counteract gestural reduction (cf., Chen, 2003; Surprenant & Goldstein, 1998).

Another point that we would like to address is that jaw height estimated at time points during C1C2 is controlled more finely than tongue dorsum constriction in C1: the results indicate that the jaw articulator involving both C1 and C2 is more refined to reflect phonological context effects than the main articulator itself. The distribution of the dorsal spatial gesture demonstrates a bi-modal distribution, one subset in the data distribution is for both /k(#)h/ and /k(#)p/ sequences and the other for /k(#)t/ sequences. The vertical jaw gesture indicates a tri-modal distribution with jaw height in the order /k(#)t/>/k(#)p/>/k(#)h/ in the across-/within-word conditions of the fast rate and in the within-word condition of the comfortable rate. Even in the across-word condition of the comfortable rate, vertical jaw position was lower with no supralaryngeal constriction in C2, as compared to C2 sequences with oral constrictions. This may be attributed to the fact that the articulatory freedom of the tongue dorsum gesture is constrained by the adjacent tongue tip gesture, allowing little room for finer articulatory tuning.

4.2.2. Articulatory compensation in order to avoid the loss of prominent articulatory cues in non-target velar stop /k/ in C1

Prominent velar phonetic cues were attributed to velar's resistance to place assimilation; velar stop /k/ was characterized as a perceptually stronger segment, which has in turn made it cope relatively more persistently with diachronic phonological change (Blevins, 2004; Cho & McQueen, 2008). In terms of articulation, we find some supporting evidence that velar stop /k/ predominantly counteracts gestural reduction of C1, indicating i) temporal strengthening (i.e., lengthening) with respect to increased gestural overlap in /k(#)p/ sequences and ii) spatial strengthening (i.e., spatial reinforcement) compensating for shorter constriction duration in conjunction with less gestural overlap in /k(#)t/ sequences. Previous perceptual studies were limited to testing speakers' articulatory maneuverability to rescue phonetic cues being at risk due to more gestural overlap (as shown in /k(#)p/ sequences) or physiological constraints (as observed in /k(#)t/ sequences). The current study adds a piece of articulatory evidence that speakers seem to adjust the articulation of a C1 velar stop in non-assimilating contexts such that C2-dependent spontaneous temporal strengthening (e.g., (/k(#)h/=/k(#)p/)>/k(#)t/) or spatial strengthening (e.g., (/k(#)h/=/k(#)p/) </k(#)t/) are manifested. Cautions need to be made in concluding that temporal reduction is attested in Korean non-assimilating contexts (e.g., /k(#)t/), and we still need to consider various factors such as physiological constraints and the role of the jaw during the articulation of a gesture so as to determine whether it is reduced or strengthened. By doing this, we may come to understand inter-gestural coordination in C1C2 sequences in a more comprehensive way.

References

Bates, D., Maechler, M. and Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version, 1, 1-23. Retrieved from http://CRAN.R-project.org/package =lme4

Borden, G. J. and Harris, K. S. (1984). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. 2nd Ed. Baltimore: Williams & Wilkins.

Browman, C. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook* 3, 219-252.

Browman, C. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology* 6, 201-251.

Browman, C. and Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In Kingston, John and Mary Beckman, eds., *Papers in Laboratory Phonology* I: *Between the Grammar and the Physics of Speech*, 341-376. Cambridge: Cambridge University Press.

Browman, C. and Goldstein, L. (1991). Gestrural structures: Distinctiveness, phonological processes, and historical change. Modularity and the motor theory of speech perception. In Ignatius G. Mattingly and Michael Studdert-kennedy, eds., *Proceedings of a Conference to Honor Alvin M. Liberman*, 313-338. Lawrence Erlbaum Associates Inc., Publishers.

Browman, C. and Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica* 49, 155-180.

Browman, C. and Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti and L. J. Raphael, eds., *Producing Speech: Contemporary Issues* (for K.S. Harris), 19-33. Woodbury, N.Y., AIP Press.

Byrd, D. (1992). Perception of assimilation in consonant clusters: a gestural model. *Phonetica* 49, 1-24.

Byrd. D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24, 209-244.

Chen, L. H. (2003). The origins in overlap of place assimilation. In Gina Garding and Mimu Tsujimura, eds., *Proceedings of the XXIIth West Coast Conference on Formal Linguistics*, 137-150. Somerville, MA: Cascadilla Press.

Chitoran, I., Goldstein, L. and Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven and N. Warner, eds., *Laboratory Phonology* 7, 419-448. Berlin: Mouton de Gruyter.

Cho, T. , Lee, Y. and Kim, S. (2014). Prosodic strengthening on the /s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics* 46, 128-146.

Cho, T. and Keating, P. (2001) Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29, 155-190.

Cho, T. and McQueen, J. (2008). Not all sounds in assimilation environments are perceived equally: Evidence from Korean. *Journal of Phonetics* 36, 239-246.

Fujimura, O., Macchi, M. and Streeter, L. A. (1978). Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech* 21, 337-346.
Gick, B. (2003). Articulatory correlates of ambisyllabicity in English glides and liquids. In John Local, Richard Ogden, and Rosalind Temple, eds., *Phonetic Interpretation: Papers in Laboratory Phonology* VI, 222-236.
Giles, S. B. and Moll, K. L. (1975). Cinefluorographic study of selected allophones of English /l/. *Phonetica* 31 206-227.
Householder, F. W. (1956). *Unreleased /p t k/ in American English. For Roman Jakobson*, M. Halle, ed., 235-244. The Hague: Mouton.
Jang, M. (2006). The acoustic characteristics of aspiration merger in Korean. The Universtiy of Texas at Austin. Handout.
Jun, J. (1995). *Perceptual and Articulatory Factors in Place Assimilation: An Optimality Theoretic Approach*. Ph.D. dissertation. University of California in Los Angeles.
Jun, J. (1996). Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology*, 13 377-407.
Jun, J. (2004). Place assimilation. In Bruce Hayes, Robert Kirchner and Donca Steriade, eds., *Phonetically Based Phonology*, 58-86. Cambridge University Press.
Keating, P. A. (1991). Coronal places of articulation. In C. Paradis and J.-F. Prunet, eds., *The Special Status of Coronals*, 29-48. Academic Press.
Keating, P. A., Lindblom, B., Lubker, J. and Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics* 22, 407-422.
Kochetov, A., Pouplier, M. and Son, M. (2007). Cross-language differenxces in overlap and assimilation patterns in Korean and Russian. *Proceedings of the 16th International Congress International Congress of Phonetic Sciences*, 1361-1364.
Krakow, R. (1989). *The articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures*. Ph.D. dissertation. Yale University.
Krakow, R. (1999). Physiological organization of syllables: a review. *Journal of Phonetics* 27, 23-54.
Kühnert, B. and Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinical Linguistics and Phonetics* 18, 559-575.
Löfqvist, A. (1993). Electromagnetic transduction techniques in the study of speech motor control. *PHONUM* 2, 87-106.
Malécot, A. (1958). The role of releases in the identification of released final stops. *Language* 34, 370-380.
Marin, S. and Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics* 42, 24-36.
Mooshammer, C., Hoole, P. and Geumann, A. (2007). Jaw and order. *Language and Speech* 50, 145-176.
Mooshammer, C., Hoole, P. and Kühnert, B. (1995). On loops. *Journal of Phonetics* 23, 3-21.
Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In Gerry

Docherty and D. Robert Ladd, eds., *Paper in Laboratory Phonology 2: Segment, Gesture, and Tone*, 261-280. Cambridge: Cambridge University Press.
Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In Kingston, John and Mary E. Beckman, eds., *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, 258-275. Cambridge University Press.
Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. and Jackson, M. (1992). Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America* 92, 3078-3096.
Pouplier, M., Marin, S., Hoole, P. and Kochetov, A. (2017). Speech rate effects in Russian onset clusters are modulated by frequency, but not auditory cue robustness. *Journal of Phonetics* 64, 108-126.
R Development Core Team. (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.rproject.org
Rubin, P., Baer, T. and Mermelstein, P. (1981). An articulatory synthesizer for perceptual research. *Journal of the Acoustical Society of America* 70, 321-328.
Saltzman, E. (1995). Intergestural timing in speech production: data and modeling, *Proceedings of the 13th International Congress International Congress of Phonetic Sciences*, 84-91.
Son, M. (2008). Gestural overlap as a function of assimilation contrast. *Korean Journal of Linguistics* 33, 665-691.
Son, M. (2011). A language-specific motor constraint in Korean non-assimilating consonant sequences. *Speech Sound and Speech Sciences* (말소리와 음성과학) 3, 27-33.
Son, M. (2011). The projection of syllable structure: A case study of intervocalic /k/ in Korean. *Korean Journal of Linguistics* (언어) 36, 395-414.
Son, M., Kim, S. and Cho, T. (2011). Supralaryngeal articulatory characteristics of coronal consonants /n, t, t^h, t*/ in Korean. *Phonetics and Speech Sciences (말소리와 음성과학)* 3, 33-43.
Steriade, D. (2000). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms., UCLA, Los Angeles, CA.
Steriade, D. (2001). Directional asymmetries in place assimilation: a perceptual account. In Elizabeth V. Hume and Keith Johnson, eds., *The Role of Speech Perception in Phonology*, 219-250. Academic Press.
Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonology* 17, 3-45.
Stevens, K. N. (1997). Articulatory-acoustic-auditory relationships. In W. J. Hardcastle and J. Laver, eds., *Handbook of Phonetic Sciences*, 462-506. Cambridge, MA: Blackwell.
Surprenant, A. M. and Goldstein, L. (1998). The perception of speech gestures. *Journal of the Acoustical Society of America* 104, 518-529.

Minjung Son
Associate Professor
Department of English Language & Literature (#218)
Hannam University
70 Hannamro, Daedeok-Gu, Daejeon 34430, Korea
E-mail: minjungson@hnu.kr

Received: June 30, 2019
Revised version received: August 12, 2019
Accepted: August 14, 2019

Appendix

1. cɔnhak hacamaca iltɨŋhessɔjo (전학 하자마자 일등 했어요.)
   ‘(I) was top (in all subjects) as soon as (I was) transferred.’
2. cɔnhakpatci anha (전학 받지 않아.)
   ‘(We) do not accept (any) transferred (students).’
3. ɔhak taɨmɨlo suhaki munceja (어학 다음으로 수학이 문제야.)
   ‘Mathematics matters next to the verbal (part).’
4. cɔnha**kh**ataka koseŋhessɔjo (전학하다가 고생했어요.)
   ‘(I) had trouble during my transfer (from one school to another).’
5. cɔnhakpanenɨn cɔnhak on hakseŋman issɔ (전학반에는 전학 온 학생만 있어.)
   ‘The class for the transferred only has students transferred students (from other schools).’
6. ɔhaktapɨn nonsullo hejatwey (어학답은 논술로 해야돼.)
   ‘Answers for the verbal part ought to be given in the essay.’