Journal of Phonetics 39 (2011) 344–361

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Communicatively driven versus prosodically driven hyper-articulation in Korean

Taehong Cho[a,*], Yoonjeong Lee[a], Sahyang Kim[b]

[a] Hanyang Phonetics and Psycholinguistics Lab, Department of English Language and Literature, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul, Republic of Korea
[b] Department of English Education, Hongik University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 27 June 2010
Received in revised form
7 February 2011
Accepted 10 February 2011
Available online 2 May 2011

ABSTRACT

This study investigated how three different kinds of hyper-articulation, one communicatively driven (in clear speech), and two prosodically driven (with boundary and prominence/focus), are acoustic-phonetically realized in Korean. Several important points emerged from the results obtained from an acoustic study with eight speakers of Seoul Korean. First, clear speech gave rise to global modification of the temporal and prosodic structures over the course of the utterance, showing slowing down of the utterance and more prosodic phrases. Second, although the three kinds of hyper-articulation were similar in some aspects, they also differed in many aspects, suggesting that different sources of hyper-articulation are encoded separately in speech production. Third, the three kinds of hyper-articulation interacted with each other; the communicatively driven hyper-articulation was prosodically modulated, such that in a clear speech mode not every segment was hyper-articulated to the same degree, but prosodically important landmarks (e.g., in IP-initial and/or focused conditions) were weighted more. Finally, Korean, a language without lexical stress and pitch accent, showed different hyper-articulation patterns compared to other, Indo-European languages such as English—i.e., it showed more robust domain-initial strengthening effects (extended beyond the first initial segment), focus effects (extended to V1 and V2 of the entire bisyllabic test word) and no use of global F0 features in clear speech. Overall, the present study suggests that the communicatively driven and the prosodically driven hyper-articulations are intricately intertwined in ways that reflect not only interactions of principles of gestural economy and contrast enhancement, but also language-specific prosodic systems, which further modulate how the three kinds of hyper-articulations are phonetically expressed.

1. Introduction

Speech production is by nature variable, so that no single utterance can be repeated with exactly the same physical properties. Some aspects of speech variability may be inevitably physiological or biomechanical, as no two speakers can have exactly the same vocal tract configurations nor can even the same speaker assume the exact same articulatory posture twice. However, speech variability may also arise in linguistically or communicatively relevant ways. One source of linguistically relevant speech variation can be found with communicative conditions, as described by the H&H theory (Lindblom, 1990). For example, when speakers face communicatively adverse situations (e.g., speaking in noise, or talking to hearing-impaired or non-native listeners), they are likely to make efforts to produce the utterance more clearly (or hyper-articulate it) to enhance its intelligibility. On the other hand, when the communicative situation is optimal (e.g., speaking in a quiet environment or talking casually to a friend about mutually understood topics), speakers are likely to employ a low-cost form of motor behavior to produce the utterance (or hypo-articulate it). As a result of such communicatively driven adjustment of speech production, an utterance may be produced with a wide range of acoustic-phonetic variation along the hypo- to hyper-articulated speech continuum (see Smiljanić & Bradlow, 2009 for a comprehensive review).

Another source of linguistically relevant speech variation is prosodic structure. Spoken utterances are prosodically modified, depending on the prosodic structure with which the utterance is produced (see Cho, in press; Shattuck-Hufnagel & Turk, 1996 for reviews). The prosodic structure of an utterance is known to be expressed not only by suprasegmental features such as F0, duration, and amplitude, but also by segmental or articulatory features. For example, speech production can vary with degree of prominence, such that segments are articulated strongly in lexically stressed syllables, and the lexically stressed segments are articulated even more strongly when they receive higher-order prominence such as

* Corresponding author. Tel.: +82 2 2220 0746; fax: +82 2 2220 0741.
E-mail addresses: tcho@hanyang.ac.kr (T. Cho),
yj.cynthia.lee@gmail.com (Y. Lee), sahyang@hongik.ac.kr (S. Kim).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.02.005

accentuation or focus (e.g., in English nuclear pitch accent) (e.g., Cho, 2004, 2005; Cho & Keating, 2009; de Jong, 1995, 2004; Fowler, 1995). Prosodically driven articulatory variation is also found with domain-initial strengthening: segments that occur in the initial position of larger prosodic units (after a larger prosodic boundary, e.g., the Intonational Phrase, IP, boundary) are produced more strongly than those in the initial position of smaller prosodic units (after a smaller prosodic boundary, e.g., the Word boundary) (e.g., in English, Cho, 2005; Cho & Keating, 2009; Fougeron & Keating, 1997; in French, Fougeron, 2001; in Tamil, Byrd, Kaun, Narayanan, & Saltzman, 2000; in Dutch, Cho & McQueen, 2005; in German, Kuzla, Cho, & Ernestus, 2007; in Japanese, Onaka, 2003; Onaka, Watson, Palethorpe, & Harrington, 2003; in Korean, Cho & Keating, 2001; Jun, 1995; Kim, 2003). The degree of domain-initial strengthening has been thought to perform the function of marking prosodic boundaries, which contributes to signaling prosodic structure (cf. Cho, in press).

The sources of speech variation discussed so far can largely be divided into three different types, all of which are thought to give rise to some kind of hyper-articulation. When Lindblom (1990) first introduced the H&H theory, hyper-articulation was expected to occur globally in the utterance, which may be termed *communicatively driven (global) hyper-articulation*. The term 'communicatively driven' could be used in a broader sense, given that the ultimate goal of speech production, in most cases, is communication. However, in the present study, we use the term narrowly defined as driven by adverse communicative situations which would call upon speakers to employ a clear speech mode for the improvement of the overall intelligibility of the utterance, as used in Lindblom (1990). de Jong (1995) distinguished this from prominence-induced hyper-articulation by characterizing the latter as *localized hyper-articulation* as it is local to a lexically stressed syllable with nuclear pitch accent. The boundary-induced articulatory strengthening (e.g., domain-initial strengthening) can also be thought of as another kind of localized hyper-articulation as its effect occurs in the vicinity of prosodic juncture (cf. Byrd & Saltzman, 2003; Cho & Keating, 2009). The latter two kinds of localized hyper-articulation can therefore be called *prosodically driven (local) hyper-articulation*.

The primary purpose of the present study is to investigate how the different sources of hyper-articulation are acoustic-phonetically realized in Korean by examining the above three conditions (i.e., clear speech, prominence, and boundary conditions) simultaneously. One of the important questions is about the extent to which they are similar and different in their acoustic-phonetic characteristics. The three kinds of hyper-articulation are all characterized by some kind of extreme articulation, and they, whether communicatively or prosodically driven, converge on one common goal—i.e., the successful delivery of linguistic messages to the listener as intended by the speaker, which is perhaps best described by the Jakobson, Fant, and Halle's (1965) classic statement, "We speak to be heard in order to be understood." One can therefore expect that there will be at least some acoustic-phonetic patterns that all three types of hyper-articulation have in common. In principle, however, the sources of hyper-articulation are different in terms of communicative and/or linguistic functions—i.e., enhancing the global intelligibility of the utterance (in a clear speech mode), demarcating the continuous speech into prosodic groups (marking prosodic boundaries) or signaling information locus (marking prominence). The assumed different functions allude to separate encoding of different types of hyper-articulation in speech planning, which makes it plausible that they are phonetically expressed in a distinct way—at least in some phonetic dimensions. Previous studies have indeed provided partial support for this prediction with English. Speakers differentiate prominence-induced from boundary-induced hyper-articulation, supporting the view that prosodic strengthening serves dual functions, prominence marking and boundary marking, and they are encoded separately in speech planning (Cho & Keating, 2009; Keating & Shattuck-Hufnagel, 2002; see Xu & Wang, 2009, for a related claim regarding boundary marking in Mandarin). However, no studies have systematically compared all three types of hyper-articulation in any given language, addressing the issue of multiple functions of hyper-articulation. In the present study, we therefore explore this issue with Korean by examining how the prosodically driven hyper-articulation is distinguished from the communicatively driven hyper-articulation, and how prominence- versus boundary-induced hyper-articulations are further differentiated from each other.

The second goal of the present study is to explore whether and how the different sources of hyper-articulation interact with each other. Although the communicatively driven and prosodically driven hyper-articulation can be independently motivated (e.g., prosodic strengthening may occur in both the casual and the clear speech modes), they appear not to be mutually exclusive. Smiljanić and Bradlow (2008b), for example, showed that more prosodic phrases are formed in clear speech than in casual speech in English, demonstrating that modification of prosodic structure can be communicatively driven. As now understood, prosodically driven hyper-articulation has functions to mark the information loci or linguistically important positions (cf. Keating & Shattuck-Hufnagel, 2002) and the resulting local phonetic clarities are assumed to facilitate speech comprehension (Cho, McQueen, & Cox, 2007; Cutler & Butterfield, 1992; Gow, Melvold, & Manuel, 1996). It is then reasonable to assume that, all else being equal, the goal of the communicatively driven hyper-articulation is more effectively achieved if speakers heighten the phonetic clarity of segments in prosodically important locations (i.e., accented syllables or domain-initial positions) more than those in prosodically weak locations (cf. Smiljanić & Bradlow, 2008a). We will test this possibility by examining the interaction effects between the three types of hyper-articulation in Korean.

The third goal of the present study is to expand our knowledge on hyper-articulation to a language other than Indo-European languages. We have chosen Korean as a test language for a number of reasons. There have been increasingly a large number of studies reported on clear speech effects primarily with English and a few other Indo-European languages, such as Spanish and Croatian (e.g., Bradlow, 2002; Bradlow & Bent, 2002; Helfer, 1998; Picheny, Durlach, & Braida, 1986; Schum, 1996; Smiljanić & Bradlow, 2005, 2008a, b; Uchanski, 2005). Previous studies with English have definitely improved our understanding of clear speech effects under various adverse communication situations with different listener populations (e.g., listeners with hearing impairments, elderly adults, non-native listeners, and children with learning impairments). Our knowledge on clear speech effects in non-European languages such as Korean, however, has been extremely limited (cf. Smiljanić & Bradlow, 2009). Only a few studies (e.g., Kang, 2010; Kang & Guion, 2008) have examined clear speech aspects of Korean with limited scopes (e.g., focusing only on how the three-way stop contrast in Korean is enhanced in a clear speech mode), and as a result there remains much room for further investigation. Our knowledge on prosodic strengthening in Korean has been quite limited as well. Although domain-initial strengthening effects have been investigated with Korean (Cho & Keating, 2001; Jun, 1993; Kim, 2003), no studies have examined interactions between boundary-induced strengthening and prominence (focus)-induced strengthening in Korean, let alone interactions between prosodically driven and communicatively driven hyper-articulation.

Expanding our knowledge of both clear speech effects and prosodic strengthening effects in Korean will therefore provide us a better and more balanced insight into the different types of hyper-articulation in general, and it will serve as a basis for understanding cross-linguistic similarities and differences of hyper-articulation in particular, by allowing us to compare Korean data with already-existing data in English and other languages. In what follows, we will discuss specific questions that bear on language-specificity and cross-linguistic similarities that are to be addressed in the present study.

The first specific question is concerned with the relationship between the scope of domain-initial strengthening and the prosodic system of a given language. Quite a few studies (Cho & Keating, 2001, 2009; Cho & McQueen, 2005; Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003) have demonstrated that boundary effects are mainly local to the consonant in domain-initial CV, and the effects on the following vowel have been found to be rather limited (cf. Byrd, 2000; Byrd, Krivokapić, & Lee, 2006; Cho, 2006, 2008; Cho & Keating, 2009; Krivokapić, 2007). For example, in English, although phonological vowel features are known to be enhanced in a way to enhance phonological contrasts between vowels in a clear speech mode (Smiljanić & Bradlow, 2005, 2008a) as well as in accented (focused) condition (Cho, 2005; de Jong, 1995, 2004), boundary-induced hyper-articulation effects have not been clearly observed in terms of enhancement of vowel features (Cho, 2005; Cho & Keating, 2009). Barnes (2002) attributed the lack of initial strengthening effects on the vocalic articulation in CV syllables to the specific role of the vowel in English that is arguably reserved for acoustic manifestation of lexical stress. By examining domain-initial strengthening patterns in four different languages, including Korean and English, Keating et al. (2003) also suggested that the more robust domain-initial strengthening effects found in Korean than in English may be due to the fact that Korean marks prosodic structure primarily by phrasing while both phrasing and prominence markings are employed in English.

It is therefore possible that, unlike English, languages without lexical stress and pitch accent such as Korean would show more robust domain-initial strengthening effects, such that the effects may spread well into the vowel in a domain-initial CV syllable as its domain of influence is not restricted by the lexical prominence system. In the present study, we will test this prediction by examining whether the three peripheral vowels /i,a,u/ in Korean are hyper-articulated in domain-initial CV position, and if so, how they are realized in connection with enhancement of the phonological contrast between the vowels. We will also examine the extent to which boundary-induced hyper-articulation (domain-initial strengthening) spreads into the test word even beyond the first syllable to see how far the effect can be extended beyond the segments immediately adjacent to the boundary (cf. Byrd et al., 2006; Cho, 2008; Cho & Keating, 2009; Krivokapić, 2007). Any observable boundary-induced hyper-articulation effects will then be compared with prominence (focus)-induced and clear speech-induced hyper-articulation effects to see whether and how the three kinds of hyper-articulation are differentiated in the acoustic realization of the vowels.

Our next specific question is concerned with the relationship between the vowel inventory size and the vowel space expansion in hyper-articulation environments. Smiljanić and Bradlow (2008a) explored this question comparing English and Croatian in a clear speech mode, and showed a cross-linguistically comparable degree of vowel space expansion, despite the fact that the two languages differ drastically in the number of contrastive vowels (14 versus 5 vowels for English and Croatian, respectively). Based on the results, they concluded that the vowel contrast enhancement can be considered as universally applicable in clear speech production, irrespective of vowel inventory size. Korean also has a relatively small vowel inventory size with 7 contrastive vowels (Shin & Cha, 2003). Crucially, the three peripheral vowels /i,a,u/ are each positioned in a section of the vowel space with no adjacent neighboring vowels in that section (which would otherwise jeopardize their distinctiveness). Hence, there would be no compelling force to expand the vowel space in the clear speech mode. In order to address this issue, we will examine how the vowel (F1–F2) space formed by the three peripheral vowels in Korean is expanded by all three kinds of hyper-articulation conditions. We will then compare the results with English and Croatian data, and discuss implications for the principles of contrast enhancement and effort minimization (Liljencrants & Lindblom, 1972; Lindblom, 1990; see also Flemming, 1995, 2001).

Finally, the Korean data will allow us to consider a question about the universality versus language-specificity of pitch range expansion associated with hyper-articulation. Previous studies on both English and Croatian (Smiljanić & Bradlow, 2005) showed a global expansion of pitch range in clear speech, and in a review paper, Smiljanić and Bradlow (2009) implied that pitch range expansion may be a universally applicable feature of clear speech. It is of course a viable assumption that the perceptibility of the intonational structure may be enhanced with more pitch excursion, which may in turn contribute to the intelligibility enhancement in clear speech, as found with other languages (e.g., English and Croatian). However, English and Croatian both have lexical stress generally expressed by high (rising) pitch (as well as increase in duration and amplitude, cf. Lehiste, 1970), and the pitch excursion is expected to be heightened with higher F0 when it receives phrasal accent (e.g., focus) serving as the locus of prominence (i.e., the head of a prosodic phrase; cf. Beckman & Edwards, 1994). It is then plausible that the pitch range expansion in clear speech may be a feature of prosodic systems with lexical stress and pitch accent, while languages without lexical stress such as Korean may not employ pitch excursion as much.

In sum, the present study investigates systematic acoustic-phonetic variation in Korean as a function of different hyper-articulation conditions, in order to understand how communicatively driven (in a clear speech mode) versus prosodically driven (domain-initial or focused) hyper-articulation effects are similar and different, how they interact, and how they reflect the language-specific phonological and prosodic systems of Korean.

2. Method

2.1. Subjects and recording

Eight male native speakers of Seoul Korean participated in this experiment. All speakers were students at Hanyang University in Seoul and were paid for the participation in the study. The participants were not aware of the purpose of the present study. The acoustic data from five speakers were collected in a sound attenuated booth along with the articulatory data acquisition using Electromagnetic Midsagittal Articulography (Carstens Articulograph AG 200) with an AGK C420 head-mounted microphone at a sampling rate of 44 kHz. The articulatory data, which we plan to report in the future, are currently being analyzed. The remaining three speakers' acoustic data were collected, without acquisition of EMA data, at a sampling rate of 44 kHz using a SHURE KSN44 dynamic microphone and a Tascam HD-P2 digital recorder in a sound attenuated booth at the Hanyang Phonetics and Psycholinguistics Lab.

2.2. Test sentences and procedure

The test syllables were /pʰi, pʰa, pʰu/. The bilabial aspirated stop /pʰ/ was chosen as it is known to show a robust prosodic strengthening (Cho & Jun, 2000; Jun, 1993, 1995). The three peripheral vowels /i,a,u/ were included to see how the acoustic vowel space formed by them varies as a function of three hyper-articulation-inducing factors. They were embedded in disyllabic words in Korean (i.e., /p̲ʰa̲tʃʌn/ 'Korean pancake', /p̲ʰi̲pu/ 'skin', and /p̲ʰu̲tiŋ/ 'pudding'). These words were then included in sentences with which the three critical factors were manipulated: the

Table 1
Experimental sentences with the test syllable /p^{h}a/ in /**p^{h}atʃʌn**/ in IP-initial focused (a), IP-initial unfocused (b), IP-medial focused (c), and IP-medial unfocused (d) conditions. The test word is underlined, and focused items are in bold.

(a)	IP-initial /p^{h}a/: Focused
Q:	*hjoli-nɨn ʌntʃena, [IP tʃ^{h}ik^{h}in-hako sul-ɨl mʌk-ni]?*
	Hyori-Top. always chicken-and wine-Acc. eat-Q
	"Does Hyori always eat chicken and wine?"
A:	*ani, hjoli-nɨn ʌntʃena, [IP **p^{h}atʃʌn**-hako sul-ɨl mʌkʌ].* (Test Sentence)
	No, Hyori-Top. always pancake-and wine-Acc. eat
	"No, Hyori always eats PAJUN (pancake) and wine."
(b)	IP-initial /p^{h}a/: Unfocused
Q:	*hjoli-nɨn ʌntʃena, [IP p^{h}atʃʌn-hako k^{h}olla-lɨl mʌk-ni]?*
	Hyori-Top. always pancake-and coke-Acc. eat-Q
	"Does Hyori always eat pajun and coke?"
A:	*ani, hjoli-nɨn ʌntʃena, [IP p^{h}atʃʌn-hako **sul**-ɨl mʌkʌ].* (Test Sentence)
	No, Hyori-Top. always pancake-and wine-Acc. eat
	"No, Hyori always eats pajun (pancake) and WINE."
(c)	IP-medial /p^{h}a/: Focused
Q:	*hjoli-nɨn ʌntʃena, [IP kokuma tonk*asɨ joli-laŋ sotfu-lɨl mʌk-ʌs*-ni]?*
	Hyori-Top. always sweet potato pork cutlet dish-and soju-Acc. eat-Past-Q
	"Did Hyori always have the sweet potato pork cutlet with soju (liquor)?"
A:	*ani, hjoli-nɨn ʌntʃena, [IP kokuma **p^{h}atʃʌn** joli-laŋ sotfu-lɨl mʌk-ʌs*ʌ].* (Test Sentence)
	No, Hyori-Top. always sweet potato pancake dish-and soju-Acc. eat-Past
	"No, Hyori always had the sweet potato PAJUN (pancake) dish with soju (liquor)."
(d)	IP-medial /p^{h}a/: Unfocused
Q:	*hjoli-nɨn ʌntʃena, [IP kamtʃa p^{h}atʃʌn joli-laŋ sotfu-lɨl mʌk-ʌs*-ni]?*
	Hyori-Top. always potato pancake dish-and soju-Acc. eat-Past-Q
	"Did Hyori have the potato pajun (pancake) dish with soju (liquor)?"
A:	*ani, hjoli-nɨn ʌntʃena, [IP **kokuma** p^{h}atʃʌn joli-laŋ sotfu-lɨl mʌk-ʌs*ʌ].* (Test Sentence)
	No, Hyori-Top. always sweet potato pancake dish-and soju-Acc. eat-Past
	"No, Hyori always had the SWEET POTATO pajun (pancake) dish with soju (liquor)."

communicatively driven factor was Speaking style (Clear versus Casual conditions), and the prosodically driven factors were Boundary (IP-initial versus IP-medial conditions) and Prominence (Focus versus Unfocused conditions). Test sentences for /p^{h}a/ are given in Table 1 (see the Appendix for the other test sentences). The various conditions associated with these factors were obtained with the following procedure.

In order to induce variation with the factors as naturally as possible in a limited laboratory setting, and to avoid orthographic influences, a mini discourse situation was created in such a way that the subjects were shown a contextual picture on a computer screen and were asked to answer questions according to the context given in the pictures without written scripts. For example, the IP-initial /p^{h}a/ in the focused condition (Table 1a) was created as follows. The subject was first given the contextual picture with a famous Korean celebrity 'Hyori' eating *pajun* /p^{h}atʃʌn/ (a test word, 'Korean style pancake') with wine. The experimenter then asked the question, "Does Hyori always eat CHICKEN and wine?" by intentionally using "chicken" instead of the target word *pajun* in the picture. The subject was asked to answer the question in a complete sentence according to the contextual information given in the picture by correcting the statement as in, "No, Hyori always eats PAJUN and wine," which induced a corrective narrow focus on the target word *pajun*.[1] Speakers had to go through a practice session before actual recording because they often produced sentences with unintended renditions. The practice session was especially necessary in order to reduce the time taken for the EMA recording as speakers' endurance is generally limited in EMA experiments. During the practice session, we also asked the speaker to use same declarative markers at the end of the sentence for the sake of consistency across speakers. While this practice session may reduce spontaneity of the speech, we were still confident that with this elicitation strategy, we could elicit more natural spoken utterances than those obtained in read speech.

With respect to the IP-initial condition, by putting the frequency adverb /ʌntʃena/ ('always') just before the target word, an IP boundary was induced between them. (Without any specific instruction, subjects tended to put the IP boundary after the adverb.) For the unfocused condition with the same IP-initial /p^{h}a/, the information locus was made somewhere else other than the target word (Table 1b), so that the target word was unfocused.

For the IP-medial condition (=the Word boundary condition) (Table 1c and d), the same focused versus unfocused conditions were created in a similar way, but this time, an IP-medial prosodic boundary was induced by putting the target word as the second member in a compound (as in /kokuma **p^{h}atʃʌn**/ 'sweet potato pancake'). Note that the preceding word /kokuma/ ('sweet potato') ends with /a/ to be matched with the final vowel in /ʌntʃena/ ('always') in the IP-initial condition, to control the segmental context before the target word.

To induce different speaking styles (Clear versus Casual), the speakers were first asked to speak casually as if they were talking to their close friends. After the speakers' casual style answer in each mini discourse situation, they were then asked to repeat the answer more clearly as if it were directed to a non-native speaker of Korean who had just started to learn Korean. In order to facilitate the clear speech mode, the speakers were shown a picture of a person who looked like a non-Korean and the experimenter pretended to be the person. In every trial, the experimenter, with disguised voice mimicking a non-native speaker of Korean, told the speaker that he could not understand what had been said, and asked the speaker to repeat the sentence.

Finally, in the experimental sentences, the test syllables /p^{h}i, p^{h}a, p^{h}u/ were always preceded by an /a/-final word (e.g., *ʌntʃen**a*** 'always' in …*ʌntʃen**a** p^{h}atʃʌn*… in IP-initial condition and *kokum**a*** 'sweet potato' in …*kokum**a** p^{h}atʃʌn*… in IP-medial condition). This allowed us to examine how the domain-final vowel /a/ before the test syllable is acoustically realized under different speech styles and prosodic conditions.

The entire corpus was repeated five times in a pseudo-randomized order as follows. Speakers produced sentences with three test words (/p^{h}atʃʌn/, /p^{h}ipu/, and /p^{h}utiŋ/) in a block, which was repeated five times with different orders between the test words within the block. The order between IP-initial and IP-medial and that between focused and unfocused conditions were also alternated. In total, 960 sentences (2 styles × 2 boundaries × 2 focus conditions × 3 vowels × 5 repetitions × 8 speakers) were collected and analyzed in the present study.

2.3. Measurements

2.3.1. Global measurements for clear versus casual speech

Speaking rate: Speaking rate was included as it has been known to be heavily influenced by the speech style, which modifies the temporal structure of the sentence (e.g., Bradlow, Kraus, & Hayes, 2003; Picheny et al., 1986; Smiljanić & Bradlow, 2005). It was defined as the number of syllables per second, which was calculated based on the entire number of syllables over the entire utterance duration with silent pauses excluded.

Global F0 measures (F0 peak, F0 minimum, and pitch (F0) range): Another clear speech effect has been found on the modification of

[1] Note that we have employed 'lexical' focus on the target word, in that the target word was lexically different from the word to be corrected with no phonological resemblance between them. The focus manipulated in the present study was therefore different from phonological or segmental focus which highlights the contrast in terms of a particular phonological feature or a segment while the rest of the word remains constant (e.g., as employed in de Jong, 2004 and van Heuven, 1994).

pitch movements, primarily reflected in the extended pitch range in utterances (e.g., Bradlow et al., 2003; Picheny et al., 1986; Smiljanić & Bradlow, 2005). To examine the global pitch modification more comprehensively, we included three global F0 measures: F0 peak, F0 minimum, and the pitch (F0) range over the course of the entire utterance. F0 peak and F0 minimum values were taken from the vowel portions of the utterance—i.e., F0 peak was taken from the vowel which has the highest F0 peak of all the vowels and F0 minimum was taken from the vowel which has the lowest F0 minimum of all the vowels, using a Praat script. To ensure that local pitch perturbation was not included, each value obtained was cross-checked and corrected by inspecting the F0 contour of each utterance. Pitch (F0) range was then calculated as a difference between F0 peak and F0 minimum of the utterance.

The number of IPs (phrasing): To examine how prosodic phrasing is modified as a function of speech style, the number of IPs formed in each utterance was counted by the two of the K-ToBI trained authors. The overall agreement rate between the two transcribers was 92%. The remaining 8% of disagreed tokens were checked by a third trained K-ToBI transcriber, so that the number of IPs for these tokens was finalized when at least two out of three transcribers agreed on it.

2.3.2. Local measurements on post-boundary test words with /pʰi, pʰa, pʰu/

Voice onset time (VOT): VOT for the aspirated stop /pʰ/ was measured from the time point of the stop release burst to the voice onset of the following vowel in domain-initial position. VOT for stops has been known to be an important acoustic parameter that reflects both prosodically driven (e.g., Cho & Jun, 2000; Cho & Keating, 2001; Jun, 1993, 1995) and communicatively driven speech variation (e.g., Smiljanić & Bradlow, 2008a).

V1 and C2V2 durations: In order to examine how the temporal structure of the test word is modified by the speaking style and prosodic factors, we measured the acoustic duration of the vowel of the first syllable (V1) (i.e., the interval from the onset of the voicing of the vowel to the onset of the following consonant) and the duration of the second syllable (C2V2) of the postboundary test word. Here, we did not separate C2 and V2 of the second syllable because the test words contained different consonants. Note that VOT is often inversely correlated with the acoustic duration of the following vowel as it takes over some portion of the supralarynageal vocalic opening articulation (cf. Cho, 1996). So where there is a robust lengthening effect on VOT, there is likely a reduced or a null lengthening effect on the following vowel as some of the vocalic lengthening effect is saturated with VOT.

F0 peaks of V1 and V2: F0 maximum values of V1 and V2 (i.e., #C1<u>V1</u>C2<u>V2</u>) were measured. Each F0 peak value was taken from the acoustic vowel portion by inspecting each token visually to avoid local microscopic pitch perturbations that might occur immediately after the onset consonant. According to the prosodic model of Seoul Korean (Jun, 1993, 1995, 2000), the basic intonational structure of Korean Accentual Phrase (AP) is THLH where 'T' (tone) is phonologically determined—i.e., H is assigned when the onset consonant is either aspirated or tensed (see Kim & Cho, 2009 for further discussion on the high frequency of the basic THLH). As an AP is assumed to be embedded inside an IP (i.e., in a strictly layered prosodic structure), the first syllable of the test words in IP-initial position is assigned an H due to the presence of an initial aspirated stop while the second syllable receives an H by default. Therefore, our bisyllabic test word is expected to be associated with an HH tonal pattern (with both the first and the second syllables associated with H) as the onset is the aspirated stop (/pʰ/). The variation of F0 peak in the first and the second vowels was examined to see how the phonologized tonal patterns in the domain-initial test word are conditioned by speaking style and focus. Here it should be noted that we did not examine effects of boundary on F0 peaks of the test words because IP-initial and IP-medial tonal patterns were different in our corpus. The test word in IP-initial position started with HH while the test word in IP-medial position was the second member of the two-word compound, which formed a single AP. Therefore, IP-medial test words were not produced with a comparable HH tonal pattern, which made it difficult to compare IP-initial versus IP-medial F0 patterns directly.

V1 and V2 intensity peaks: For the intensity variation of the test word, intensity peak values were taken from the acoustic intensity profile (dB SPL) of the first and the second vowels (V1, V2). Each intensity peak value obtained from the acoustic intensity profile was visually inspected along with the display of the waveform to ensure that the value was taken from the acoustic vowel portion.

Vowel formant measures (F1, F2, and the Euclidean area): In order to investigate how the vowel formant structure of the vowels in the test words changes according to different speaking styles and prosodic factors, F1 and F2 of the test vowels (/i,a,u/) were taken from the steady-state region from the spectrographic displays. The values were taken at the hand-marked steady-state points by an LPC formant tracking function in Praat, along with visual inspection of the spectrogram for each token, which necessitated some corrections of values obtained by the formant tracking function. Based on the obtained F1 and F2, the Euclidean area of the triangle in the acoustic vowel space, formed with three coordinates based on (F1, F2) of the three vowels, was made to quantify the modification of vowel space under different test conditions.

2.3.3. Local measurements at the preboundary position

The preboundary (domain-final position) has been known as a prosodically important location as it demarcates the end of a prosodic domain. The present study therefore included some acoustic measures taken from the preboundary (domain-final) syllable (/ma/ or /na/) to see how the prosodically important domain-final articulation is acoustically modified under different speech styles and prosodic conditions. In particular, it will be interesting to see whether the preboundary articulation shows increased spatial magnitude (as to be reflected in F1 and F2) in Korean, given that boundary strengthening often induces lengthening of the domain-final segments without increased spatial magnitude in English (e.g., Beckman, Edwards, & Fletcher, 1992).

Domain-final (preboundary) V duration: The final vowel duration (which was always /a/ as in /kokum<u>a</u>/ and /ʌntʃen<u>a</u>/) was measured to see how domain-final vowels would be temporally modified by the test conditions.

Domain-final (preboundary) V intensity peak: As was the case with V1 and V2 of the postboundary test words, this measure was taken as the peak value of the acoustic intensity (dB SPL) during the final vowel /a/ to examine how it is influenced by the test conditions.

F1 and F2 of domain-final /a/: To see how the final /a/'s formant structure is modified by the test conditions, F1 and F2 were taken in the same way used for measuring the F1 and F2 of the vowels in the postboundary (domain-initial) test words.

2.4. Statistical analyses

We conducted a series of repeated measures analyses of variance (RM ANOVAs) for statistical evaluation of the influence of the communicatively driven speaking style and the prosodically driven prominence and boundary on the various acoustic measures. The within-subject factors considered were Speaking style (Clear versus Casual), Boundary (IP-initial or IP-final versus IP-medial), and Prominence (Focused versus Unfocused). RM ANOVAs with data pooled across eight speakers (with each speaker contributing one averaged score per condition) would return significance only if most speakers contributed consistently to any observed variations. In reporting the results, we will focus on the general patterns that are statistically consistent across speakers. When there were interaction effects between factors, we conducted posthoc pairwise comparisons with Bonferroni/Dunn corrections. Effect size was estimated by conducting η^2 analyses, which provide a measure of how much the observed variability can be ascribed to a given factor, and therefore show how large the observed effect might be (Sheskin, 2000: pp. 553–556). In all ANOVAs, *p*-values less than 0.05 were considered significant.

3. Results

In this section, we will first report results with respect to global characteristics of clear versus casual speech over the entire utterance and then results on local effects of Speaking style (Clear vs. Casual) on the test words along with effects of Boundary (IPi or IPf vs. IPm) and Prominence (Focused vs. Unfocused).

3.1. Global characteristics of clear versus casual speech

One-way repeated measures ANOVAs showed that there was a significant main effect of Speaking style on speaking rate ($F[1,7]=115.85$, $p<0.001$, $\eta^2=0.94$), showing that utterances in clear speech were produced more slowly than those in casual speech—i.e., speakers produced fewer syllables per second as compared to casual speech (Fig. 1a). There was also a significant main effect of Speaking style on prosodic phrasing (the number of IPs, $F[1,7]=42.83$, $p<0.001$, $\eta^2=0.86$), with more IPs in clear speech than in casual speech (Fig. 1b). As shown in Fig. 1c–e, however, there was no main effect of Speaking style on any of the three pitch measures (F0 peak, F0 minimum, Pitch (F0) range), showing that clear speech is not associated with either higher pitch or expanded pitch range.

3.2. Local effects on the test words with /p^{h}i, p^{h}a, p^{h}u/

VOT: VOT for the test consonant /p^h/ showed main effects of all three factors. VOT was longer in clear speech than in casual speech ($F[1,7]=16.35$, $p=0.005$, $\eta^2=0.7$), longer in IP-initial position than in IP-medial position ($F[1,7]=11.18$, $p<0.05$, $\eta^2=0.62$), and longer when the test word was focused than when unfocused ($F[1,7]=18.03$, $p<0.005$, $\eta^2=0.72$), as shown in Fig. 2a. A significant Speaking style × Prominence interaction was found ($F[1,7]=17.53$, $p<0.005$, $\eta^2=0.48$), which was in part due to a more robust clear speech effect in the focused condition (mean diff. 12 ms, $t(7)=3.9$, $p<0.01$, $\eta^2=0.69$) than in the unfocused condition (mean diff. 6 ms, $t(7)=3.51$, $p<0.05$, $\eta^2=0.64$) (Fig. 2b). There was also a significant Boundary × Prominence interaction ($F[1,7]=17.53$, $p<0.005$, $\eta^2=0.72$), which stemmed from two interrelated facts—i.e., the focus-induced VOT lengthening was reliable in IP-medial position (mean diff. 22 ms, $t(7)=5.24$, $p=0.001$, $\eta^2=0.8$), but not in IP-initial position (mean diff. 6 ms, $t(7)=1.73$, $p>0.1$), and the boundary-induced VOT lengthening was reliable only in the unfocused condition (mean diff. 18 ms, $t(7)=3.97$, $p=0.005$, $\eta^2=0.69$) (Fig. 2c).

V1 duration: For the vowel (V1) duration, a significant main effect was found only with Speaking style ($F[1,7]=17.28$, $p<0.001$, $\eta^2=0.87$), showing longer V1 duration in clear speech than in casual speech (Fig. 3a). Boundary and Prominence yielded no significant main effects. The Speaking style effect, however, interacted with Boundary and Prominence ($F[1,7]=22.34$, $p<0.005$, $\eta^2=0.76$, $F[1,7]=6.46$, $p<0.05$, $\eta^2=0.48$, respectively). Posthoc comparisons showed that the interaction was due to a more robust clear speech effect found in prosodically strong positions: The effect was more robust in IP-initial CV position (mean diff. 19 ms, $t(7)=7.33$, $p<0.001$, $\eta^2=0.89$) than in IP-medial position (mean diff. 10 ms, $t(7)=4.95$, $p<0.005$, $\eta^2=0.78$) (Fig. 3b), and in the focused condition (mean diff. 16 ms, $t(7)=6.92$, $p<0.001$, $\eta^2=0.87$) than in the unfocused condition (mean diff. 12 ms, $t(7)=5.86$, $p=0.001$, $\eta^2=0.83$) (Fig. 3c). As shown in Fig. 3d, there was also a significant Boundary × Prominence interaction ($F[1,7]=6.21$, $p<0.05$, $\eta^2=0.47$), which was attributable to a significant boundary effect in the focused condition (mean diff. 7 ms, $t(7)=3.37$, $p<0.05$, $\eta^2=0.62$), but not in the unfocused condition (mean diff. 2 ms, $t(7)=-0.67$, $p>0.1$, $\eta^2=0.06$).

The second syllable (C2V2) duration: All three factors showed significant main effects on the second syllable (C2V2) duration (Speaking style, $F[1,7]=30.25$, $p=0.001$, $\eta^2=0.81$; Boundary, $F[1,7]=70.03$, $p<0.001$, $\eta^2=0.91$; Prominence, $F[1,7]=41.49$, $p<0.001$, $\eta^2=0.86$), with a longer duration in clear (versus casual), focused (versus unfocused), and IP-initial (versus IP-medial) conditions (Fig. 4a). Again, Speaking style interacted

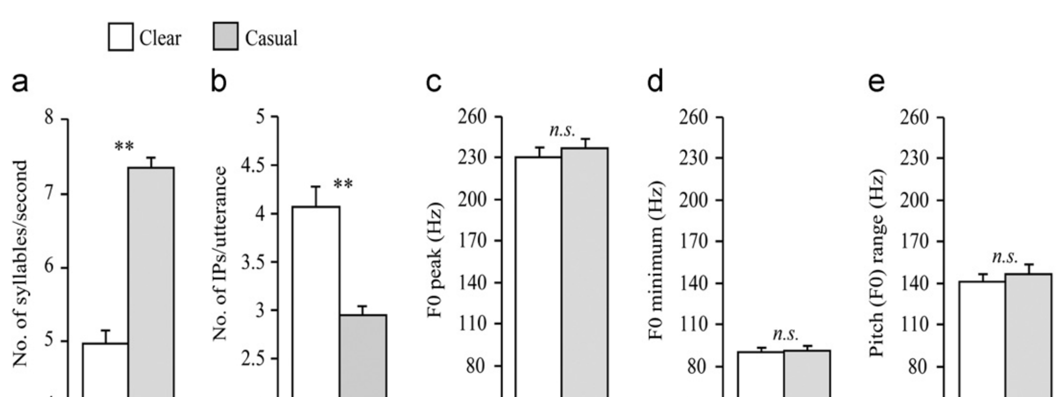

Fig. 1. Effects of Speaking style (clear versus casual) on global acoustic measures: Speaking rate (a), Prosodic phrasing (b), F0 peak (c), F0 minimum (d), and Pitch (F0) range (e) (*$p<0.05$ and **$p<0.01$).

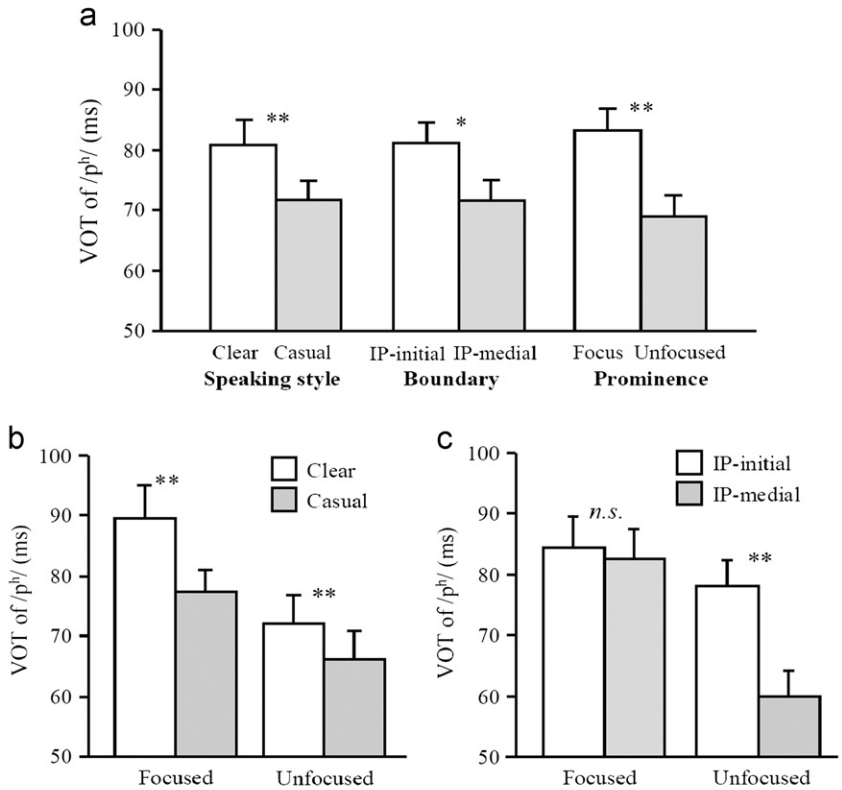

Fig. 2. Main effects of Speaking style, Boundary and Prominence on VOT of /p^h/ (a), Speaking style × Prominence interaction (b) and Boundary × Prominence interaction (c) ($^*p<0.05$ and $^{**}p<0.01$).

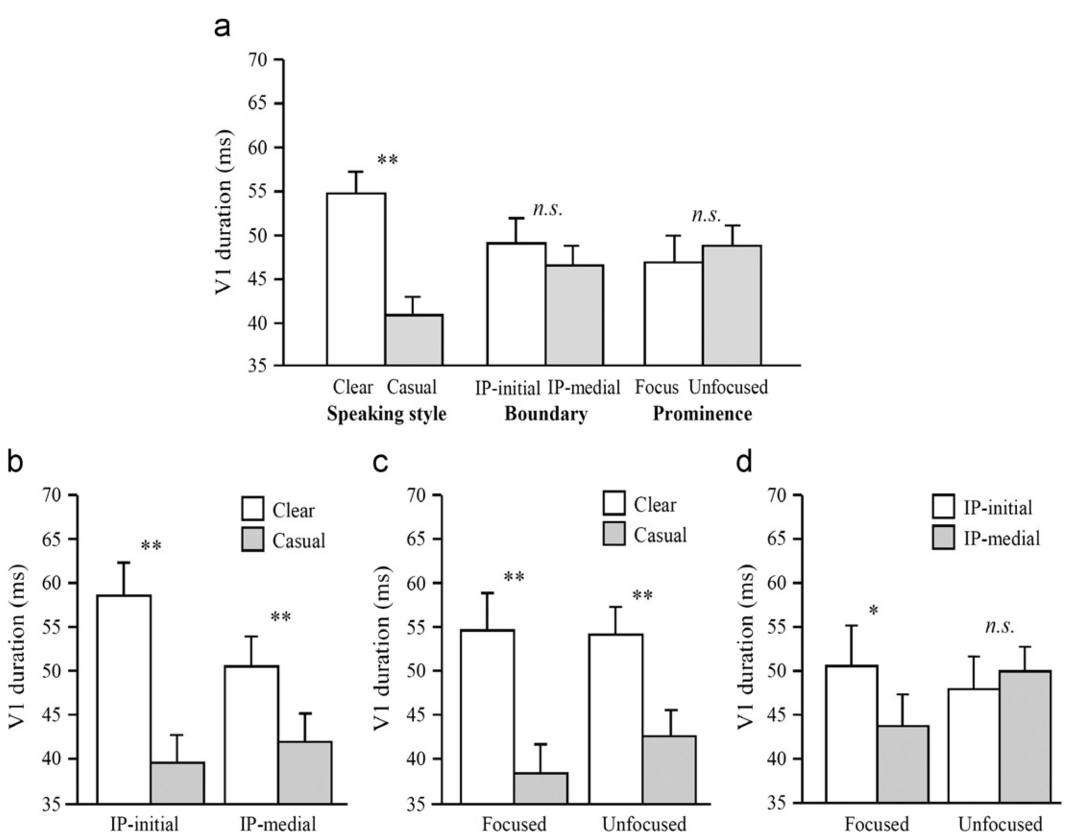

Fig. 3. Main effects of Speaking style, Boundary and Prominence on V1 duration (a), Speaking style × Boundary interaction (b), Speaking style × Prominence interaction (c), and Boundary × Prominence interaction (d) ($^*p<0.05$ and $^{**}p<0.01$).

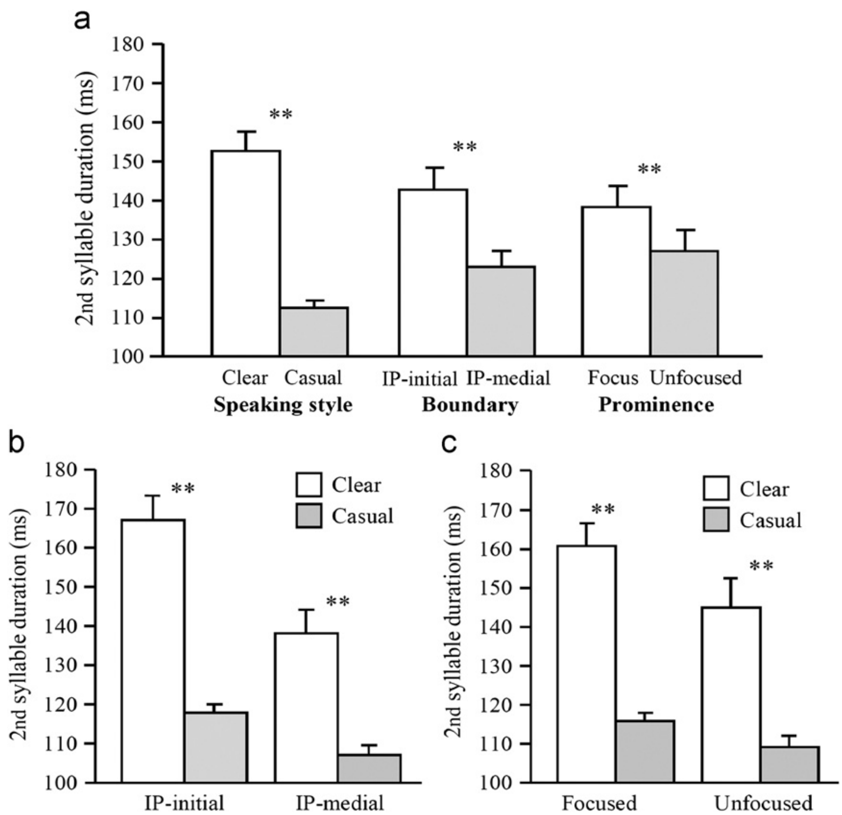

Fig. 4. Main effects of Speaking style, Boundary and Prominence on the second syllable (C2V2) duration (a), Speaking style × Boundary interaction (b) and Speaking style × Prominence interaction (c) ($^{*}p<0.05$ and $^{**}p<0.01$).

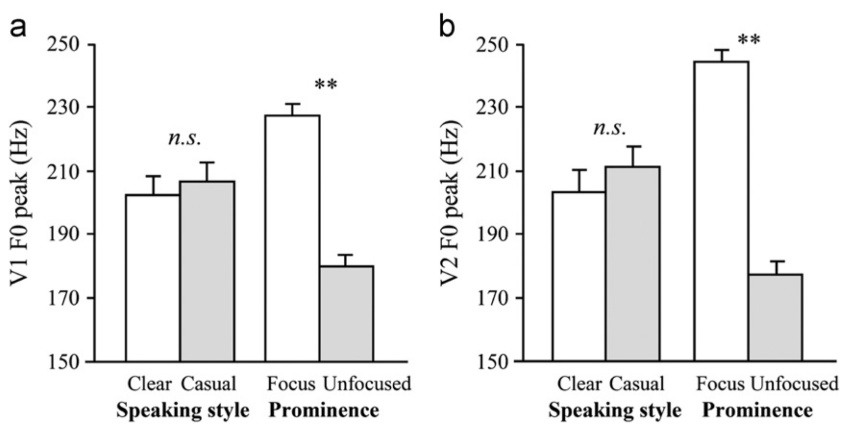

Fig. 5. Main effects of Speaking style and Prominence on V1 F0 peaks (a) and V2 F0 peaks (b). ($^{*}p<0.05$ and $^{**}p<0.01$.).

with Boundary ($F[1,7]=12.97$, $p<0.01$, $\eta^2=0.65$) and Prominence ($F[1,7]=6.61$, $p<0.05$, $\eta^2=0.49$). As was the case with V1 duration, the interaction was mainly due to a more robust clear speech effect in prosodically strong positions: a more robust clear speech effect was found when the test word was in IP-initial position (mean diff. 50 ms, $t(7)=5.79$, $p=0.001$, $\eta^2=0.83$) than in IP-medial position (mean diff. 32 ms, $t(7)=4.64$, $p<0.005$, $\eta^2=0.76$), and in the focused condition (mean diff. 46 ms, $t(7)=7.01$, $p<0.001$, $\eta^2=0.88$) than in the unfocused condition (mean diff. 36 ms, $t(7)=4.07$, $p=0.005$, $\eta^2=0.7$) (Fig. 4b and c).

V1 and V2 F0 peaks: Both V1 and V2 F0 peaks showed no main effect of Speaking style ($F[1,7]<1$ for both V1 and V2), but a significant main effect of Prominence (V1: $F[1,7]=119.36$, $p<0.001$, $\eta^2=0.95$; V2: $F[1,7]=174.88$, $p<0.001$, $\eta^2=0.96$)—i.e., it was higher when the test word was in the focused condition than in the unfocused condition (Fig. 5a and b). Interestingly, however, there was no Speaking style effect, nor was there any interaction between factors, showing no pitch modification due to Speaking style in the test word.

V1 and V2 intensity peaks: For V1 intensity peak, only the Prominence factor showed a significant main effect ($F[1,7]=39.86$, $p<0.001$, $\eta^2=0.85$): V1 intensity peak was greater in the focused condition than in the unfocused condition (Fig. 6a). However, there was a significant Speaking style × Boundary interaction ($F[1,7]=9.44$,

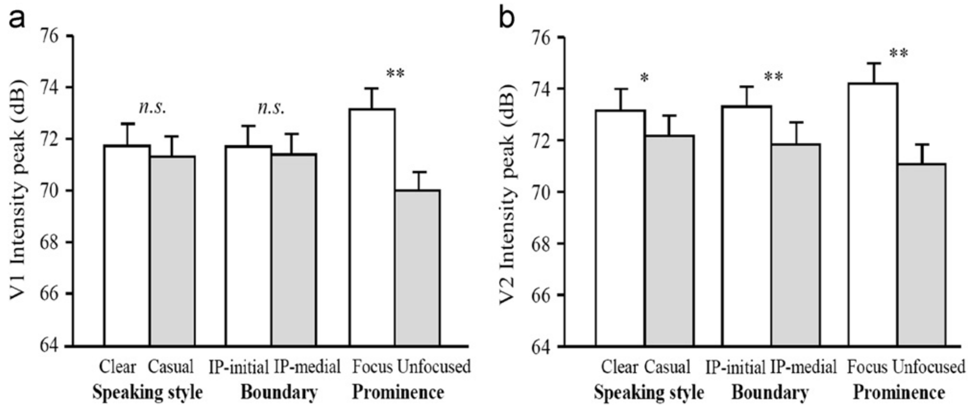

Fig. 6. Main effects of Speaking style, Boundary and Prominence on V1 Intensity peak (a) and V2 Intensity peak (b) (* $p<0.05$ and ** $p<0.01$).

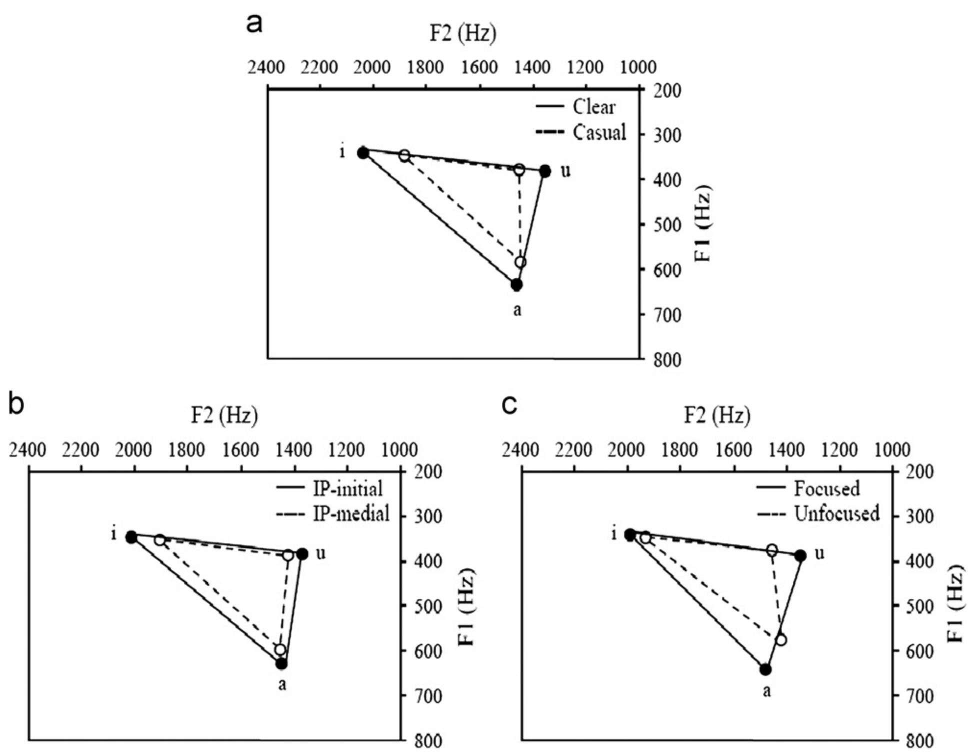

Fig. 7. /i/-/a/-/u/ Euclidean area in Clear versus Casual speech (Speaking style) (a), IP-initial versus IP-medial position (Boundary) (b), and Focused versus Unfocused position (Prominence) (c).

$p<0.05$, $\eta^2=0.57$), which was again due to the fact that the clear speech effect did arise but only in a prosodically strong, IP-initial position (mean diff. 1 dB, $t(7)=2.32$, $p=0.05$, $\eta^2=0.44$), not in IP-medial position (mean diff. 0.1 dB, $t(7)=-0.17$, $p>0.1$, $\eta^2=0.01$) (the figure is not given). There was also a significant Boundary × Prominence interaction ($F[1,7]=8.53$, $p<0.05$, $\eta^2=0.55$), showing that a boundary effect existed only when V1 was focused—i.e., the greater V1 intensity for IP-initial test words than IP-medial ones in the focused condition (mean diff. 1.36 dB, $t(7)=3.37$, $p<0.05$, $\eta^2=0.62$).

As for V2 intensity peak, all three factors showed a significant main effect (Speaking style, $F[1,7]=6.83$, $p<0.05$, $\eta^2=0.49$; Boundary, $F[1,7]=19.22$, $p<0.005$, $\eta^2=0.73$; Prominence, $F[1,7]=35.79$, $p=0.001$, $\eta^2=0.84$). V2 intensity peak was higher in clear speech than in casual speech, higher when the test word was in IP-initial position than in IP-medial position, and higher in the focused condition than in the unfocused condition (Fig. 6b). No interactions between factors were found.

F1 and F2 of V1: /a/ in /pʰa/: All three factors showed significant main effects on F1 of /a/ (Speaking style, $F[1,7]=10.95$, $p<0.05$, $\eta^2=0.61$; Boundary, $F[1,7]=33.47$, $p=0.001$, $\eta^2=0.83$; Prominence, $F[1,7]=43.31$, $p<0.001$, $\eta^2=0.86$). F1 was higher (thus positioning /a/ lower in the acoustic vowel space) in clear than in casual speech; higher in IP-initial CV position than IP-medial CV position, and higher in the focused condition than in the unfocused condition. There was a Speaking style × Boundary interaction ($F[1,7]=5.37$, $p=0.05$, $\eta^2=0.43$), which was again in part due to an asymmetrical clear speech effect between IP-initial and IP-medial position: a significantly higher F1 (positioning /a/ lower in the vowel space) in clear speech was found in IP-initial CV position (mean diff. 73 Hz, $t(7)=4.57$, $p<0.005$, $\eta^2=0.75$), but no clear speech effect was observed in IP-medial position (mean diff. 23 Hz, $t(7)=0.8$, $p>0.1$, $\eta^2=0.08$).

F2 of /a/ showed a significant main effect only with Prominence ($F[1,7]=9.68$, $p<0.05$, $\eta^2=0.58$)—i.e., F2 was higher (positioning /a/ more advanced in the acoustic vowel space) in the focused condition than in the unfocused condition. There was a significant Boundary × Prominence interaction ($F[1,7]=6.18$, $p<0.05$, $\eta^2=0.47$) due to the fact that the focus effect was significant only when the vowel was positioned IP-medially (mean diff. 83 Hz, $t(7)=3.86$, $p<0.01$, $\eta^2=0.68$).

F1 and F2 of V1: /i/ in /p^h^i/: No factor showed a significant effect on F1 of /i/, but there was a significant interaction between Boundary and Prominence ($F[1,7]=5.71$, $p<0.05$, $\eta^2=0.45$), which was due to a significant boundary effect only in the focused condition: showing boundary-induced lower F1 (positioning /i/ higher in the vowel space) only in the focused condition (focused: mean diff. 25 Hz, $t(7)=-2.43$, $p<0.05$, $\eta^2=0.46$; unfocused: mean diff. 4 Hz, $t(7)=0.53$, $p=0.61$, $\eta^2=0.04$). Unlike F1, however, F2 showed significant main effects of all three factors (Speaking style, $F[1,7]=32.94$, $p=0.001$, $\eta^2=0.83$; Boundary, $F[1,7]=17.51$, $p<0.005$, $\eta^2=0.71$; Prominence, $F[1,7]=11.59$, $p<0.05$, $\eta^2=0.62$), showing that F2 was higher (positioning /i/ more advanced in the acoustic vowel space) in clear than in casual speech, in IP-initial than in IP-medial position, and in the focused condition than in the unfocused condition. No between-factor interactions were found.

F1 and F2 of V1: /u/ in /p^h^u/: As was the case with /i/, no factor showed a significant main effect on F1 of /u/. There was no between-factor interaction, either. On the other hand, F2 of /u/ showed significant main effects of Speaking style ($F[1,7]=5.8$, $p=0.05$, $\eta^2=0.49$) and Prominence ($F[1,7]=14.56$, $p<0.01$, $\eta^2=0.71$) with no Boundary effect ($F[1,7]<1$): F2 was lower (showing a more retracted vowel quality in the vowel space) in clear than casual speech, and it was also lower in the focused condition than in the unfocused condition. No between-factor interactions were found with F2.

/i/-/a/-/u/ Euclidean Area: All three factors showed significant main effects on /i-a-u/ Euclidean area (Speaking style, $F[1,7]=29.47$, $p<0.005$, $\eta^2=0.83$; Boundary, $F[1,7]=26.73$, $p<0.005$, $\eta^2=0.82$; Prominence, $F[1,7]=34.71$, $p=0.001$, $\eta^2=0.85$). As shown in Fig. 7, the Euclidean area was larger in clear than in casual speech (Fig. 7a), larger in IP-initial than in IP-medial position (Fig. 7b), and larger in the focused condition than in the unfocused condition (Fig. 7c). (Note also that the main effects on F1 and F2 for all three test vowels can be inferred from Fig. 7.) There was a significant interaction effect of Speaking style and Boundary ($F[1,7]=14.16$, $p<0.01$, $\eta^2=0.7$), which was in part due to a more robust clear speech effect in IP-initial (mean diff. 52655 Hz^2, $t(7)=5.51$, $p=0.001$, $\eta^2=0.81$) than in IP-medial position (mean diff. 21813 Hz^2, $t(7)=3.31$, $p<0.05$, $\eta^2=0.61$) (the figure is not given).

3.3. Local effects on preboundary /a/

Final vowel (/a/) duration in preboundary position: Results of repeated measures ANOVAs showed significant main effects of Speaking style and Boundary on the final vowel duration ($F[1,7]=19.56$, $p<0.01$, $\eta^2=0.69$; $F[1,7]=114.43$, $p<0.001$, $\eta^2=0.94$, respectively), showing that the final vowel was longer in clear speech than in casual speech and it was longer in IP-final than in IP-medial position (Fig. 8a). There was, however, a significant Speaking style x Boundary interaction ($F[1,7]=9.75$, $p<0.05$, $\eta^2=0.58$), due to the fact that the clear speech effect size (i.e., mean difference) was larger in IP-final position than in IP-medial position (IP-finally: mean diff. 137 ms, $t(7)=3.57$, $p<0.01$, $\eta^2=0.65$; IP-medially: mean diff. 24 ms, $t(7)=6.9$, $p<0.001$, $\eta^2=0.87$) (Fig. 9).

Final vowel (/a/) intensity: The intensity peak of the final vowel (/a/) showed a significant Boundary effect ($F[1,7]=23$, $p<0.005$,

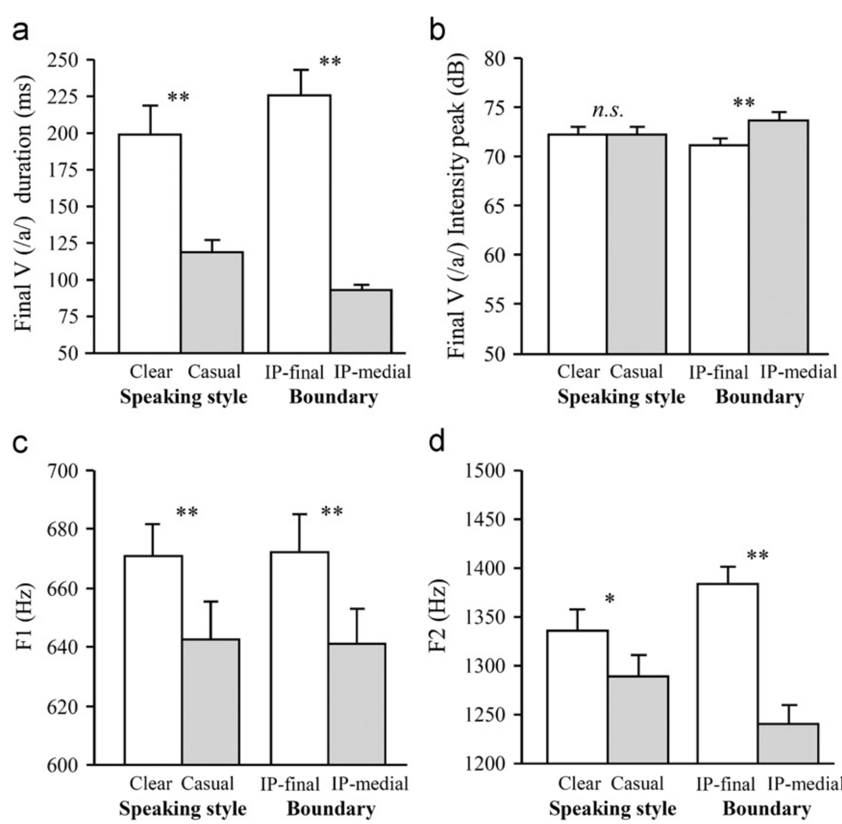

Fig. 8. Main effects of Speaking style and Boundary on the final V (/a/) duration (a), Intensity (b), F1 (c), and F2 (d) (*$p<0.05$ and **$p<0.01$).

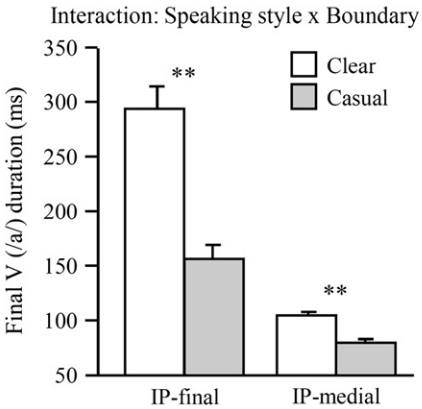

Fig. 9. Speaking style × Boundary interaction on Final V (/a/) duration (*$p<0.05$ and **$p<0.01$).

$\eta^2=0.94$)—i.e., it was *smaller* in IP-final position than in IP-medial position (Fig. 8b). There was no effect of Speaking style, nor was there a significant Speaking style × Boundary interaction.

F1 and F2 of final vowel (/a/): There were significant main effects of Speaking style and Boundary on F1, which was higher (positioning the final /a/ lower in the vowel space) in clear speech than in casual speech (670 Hz vs. 641 Hz; $F[1,7]=14.67$, $p<0.01$, $\eta^2=0.68$) and higher in IP-final than IP-medial position (671 Hz vs. 641 Hz; $F[1,7]=15.03$, $p<0.01$, $\eta^2=0.68$) (Fig. 8c). F2 also showed significant effects of Speaking style (1336 Hz vs. 1290 Hz; $F[1,7]=10.32$, $p<0.05$, $\eta^2=0.6$) and Boundary (1383 Hz vs. 1242 Hz; $F[1,7]=69.9$, $p<0.001$, $\eta^2=0.91$). F2 was significantly higher (positioning /a/ more advanced in the acoustic vowel space) in clear than in casual speech, and it was higher in IP-final than in IP-medial position (Fig. 8d). No interactions between factors were found.

4. Summary and discussion

In this section, we will first summarize the main findings about communicatively driven hyper-articulation (clear speech effects) versus prosodically driven hyper-articulation (boundary and prominence effects) in Korean. We will then discuss their implications in terms of specific questions that were raised at the outset of the paper.

4.1. Communicatively driven hyper-articulation: global and local clear speech effects

4.1.1. Global clear speech effects

The results showed global clear speech effects (i.e., effects over the course of the entire utterance) in Korean most clearly in temporal dimension and prosodic phrasing. The clear speech mode induced a decrease in overall speaking rate (i.e., fewer syllables per second) and an increase in the number of prosodic phrases (i.e., more IPs per sentence). This builds on and extends previous findings of clear speech effects in Indo-European languages such as English, Spanish, and Croatian (cf. Bradlow, 2002; Picheny et al., 1986; Smiljanić & Bradlow, 2005, 2008a, 2008b; Uchanski, 2005) to a new language, Korean, which is typologically and prosodically different. The results thus confirmed the cross-linguistic tendency that the communicatively driven global hyper-articulation gives rise to modification of the temporal and prosodic structure of the utterance to accommodate listeners in difficult communicative situations (in this case those with limited experience in the test language). The global slowing down of the utterance would allow more time for the listeners to process the speech signal, and placing a greater number of major prosodic boundaries is likely to enhance lexical segmentation (e.g., Christophe, Peperkamp, Pallier, Block, & Mehler, 2004; Kim & Cho, 2009), which, taken together, function to enhance the overall intelligibility of the utterance. The present study, however, showed an interesting pattern which runs counter to previously observed clear speech effects in English and Croatian (e.g., Smiljanić & Bradlow, 2005, 2008a)—i.e., the clear speech mode in Korean had no effect on the pitch range over the course of the sentence (see Section 4.5.3 for further discussion on this point).

4.1.2. Local clear speech effects

Along with the global clear speech effects, we have also examined the extent to which syllables locally positioned in the vicinity of prosodic juncture are acoustically modified as a function of Speaking style. Results showed that clear speech triggered local temporal expansion, as reflected in lengthened preboundary vowel /a/ and postboundary syllables (both the first and the second syllables) as well as lengthened VOT of the aspirated stop /p^h/. In addition, we have also found evidence for clear speech-induced spatial expansion—i.e., clear speech was associated with more extreme F1 and/or F2 values of each of the three peripheral test vowels /i,a,u/ along with expansion of the F1-F2 vowel space formed by the vowels. Notably, /a/ was produced with higher F1, suggesting a more lowered vowel quality (interpretable as enhancement of [+low] for /a/); /i/ was produced with higher F2, showing a more advanced vowel quality (interpretable as enhancement of [−back] for /i/); and /u/ showed lowered F2, showing a retracted vowel quality (interpretable as enhancement of [+back] for /u/). These individual clear speech patterns converged on the expansion of the /i-a-u/ Euclidean (triangular) area, showing the expansion of the acoustic vowel space in a clear speech mode in line with Smiljanić and Bradlow (2005). Finally, as was the case with the global clear speech effects, there was no notable clear speech effect on pitch in the local test syllables (i.e., the domain-final and domain-initial syllables).

4.2. Prosodically driven hyper-articulation: effects of boundary and prominence (focus)

4.2.1. Boundary effects

The results showed substantial final lengthening for /a/ in preboundary position—i.e., the final /a/ (in /na/ or /ma/) was longer in IP-final position than in IP-medial position, which is in line with the general cross-linguistic phrase-final lengthening pattern (e.g., Cho & Keating, 2001; Edwards, Beckman, & Fletcher, 1991; Klatt, 1975; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). In addition to the final lengthening effect, we have also found lowering of domain-final /a/ (with higher F1) in the acoustic vowel space. These results suggest that domain-final articulation undergoes some sort of hyper-articulation in both spatial and temporal dimensions (see Section 4.5.2 for related discussion).

The present study also found robust domain-initial strengthening effects in both temporal and spatial dimensions. The test consonant /p^h/ was produced with longer VOT IP-initially than IP-medially, which can be interpreted as having resulted from strengthening of glottal abduction articulation (cf. Cho & Keating, 2001; Pierrehumbert & Talkin, 1992). Regarding the vowels after the consonant, we have found mixed domain-initial strengthening effects. On the one hand, we have not observed boundary-induced lengthening of the vowel after the initial consonant, which is consistent with Cho and Keating (2001), who tentatively concluded that the domain-initial strengthening was limited to the initial consonant in Korean as with other languages such as English and French (Fougeron, 2001; Fougeron & Keating, 1997). On the other hand, we have found some evidence that domain-initial strengthening is indeed extended beyond the initial consonant in Korean. The test vowels /i/ and /a/ were produced with spatial expansion when they were produced in IP-initial CV syllables as reflected in their F1 and F2 values. Notably, F2 of /i/ was higher in IP-initial CV syllables than IP-medial CV syllables, positioning /i/ more advanced, and F1 of /a/ was higher, positioning /a/ lower in the F1-F2 vowel space. The boundary-induced spatial expansion in initial CV syllables was further evident with the expansion of the vowel space as reflected in the enlarged Euclidean area (see Section 4.5.1 for further discussion on this point).

In addition to initial strengthening effects on the vowel of the first syllable, the results also showed that the domain-initial effect can spread even into the second syllable of the test words. The V2 intensity peak and the second syllable duration increased when the test word was positioned IP-initially than IP-medially.

Before we move on, it is worth considering another interesting finding regarding domain-initial effects—i.e., there was no significant main effect of Boundary on the initial vowel's (V1) duration even though the duration of the second syllable was significantly longer in IP-initial position. At first glance, this lengthening effect on the second syllable may look inconsistent with the lack of V1 lengthening. It looks as if the rightward temporal influence bypassed the first vowel and affected the second syllable. Furthermore, Cho and Keating (2001) suggested that domain-initial strengthening is closely related with lengthening, such that in prosodically weak position, reduced time may induce articulatory target undershoot, but in prosodically strong position with enough time allowed, the target would be fully reached (cf. Lindblom, 1963). One might therefore expect that the vowel space expansion that was found with V1 would go hand in hand with temporal expansion of the vowel, which has not been observed in the present study.

These seeming paradoxes, however, can be understood when we consider the sum of VOT and the duration of the following vowel, which is significantly longer IP-initially than IP-medially. VOT for the aspirated stop was on average by far longer than the following vowel duration (76.4 ms vs. 47.9 ms). Given this durational asymmetry between VOT of the aspirated stop and the following vowel, Cho (1996) characterized VOT as the 'voiceless' vowel as it takes over a substantial part of the interval of the supralaryngeal articulation due to a long voicing lag. This can be further understood from the kinematic point of view. It is well known that in CV articulation the vocalic gesture starts well before the consonant release (cf. Browman & Goldstein, 1990, 1992), and the interval of VOT indeed overlaps fully with the vocalic movement duration of the gesture for the vowel target. It is therefore reasonable to assume that much of the acoustic temporal effect on the first syllable is saturated with VOT, and hence no temporal expansion is observed in the vowel. Nonetheless, the articulatory movement duration of the domain-initial vocalic gesture has been found to be lengthened in Korean (Cho, Yoon, & Kim, in preparation), which implies that the acoustic vowel expansion may indeed be closely correlated with the kinematic duration of the vocalic gesture, which is masked acoustically by VOT.

4.2.2. Prominence (focus) effects

As with initial strengthening effects, the VOT of /p^h/ and the second syllable of the postboundary test words were lengthened in the focused condition, which suggests that the duration of the entire word is influenced by prominence. Similar to the findings with boundary effects, the V1 of the test words did not show prominence-induced temporal expansion. (Again the lack of the focus effect on V1 duration could be accounted for by the inverse relationship between VOT and the following vowel duration, as discussed in the preceding section.) As was with initial strengthening, spatial expansion was found in the focused condition: F1 and F2 of /a/ were higher (indicating a more lowered and advanced vowel quality in the acoustic vowel space); F2 of /i/ was higher (indicating a more advanced vowel quality); F2 of /u/ was lower (indicating a more retracted vowel quality) in the focused condition than in the unfocused condition. The /i-a-u/ Euclidean area was significantly greater in the focused condition than in the unfocused condition as well, showing prominence-induced vowel space expansion. There was also an increase in F0 peaks for V1 and V2 in the focused condition. Finally, unlike the clear speech and boundary strengthening effects, the focused condition was associated with an increase in both V1 and V2 intensity peaks.

4.3. Interaction effects

4.3.1. Interactions between communicatively driven and prosodically driven factors

One of the questions that the present study has endeavored to answer was whether communicatively driven hyper-articulation (induced by clear speech) interacts with, or is conditioned by, prosodically driven hyper-articulation associated with boundary and prominence, and if so, how. The results indeed showed that the former interacts significantly with the latter in several acoustic measures. First, there were significant interaction effects between Speaking style and Boundary on temporal measures such as V1 duration and duration of the second syllable, and non-temporal measures such as V1 intensity and the overall /i-a-u/ Euclidean area. These interactions stemmed from a consistent pattern that the communicatively driven hyper-articulation was more robust when segments occurred in IP-initial CV syllables than IP-medial CV syllables, except for V1 intensity which showed a significant clear speech effect only when the vowel was in IP-initial CV syllables. A similar Speaking style $\times$ Boundary interaction was found with the duration of the preboundary (domain-final) /a/, again showing a more robust clear speech effect IP-finally than IP-medially. Second, Speaking style also interacted with Prominence on VOT, V1 duration, duration of the second syllable, and V1 intensity, which also showed a more robust clear speech effect in the focused condition than in the unfocused condition, except for V1 intensity which showed a significant clear speech effect only when the test word was in the focused condition. These interaction effects suggest that the clear speech effects are conditioned by prosodic factors in a way that segments positioned in prosodically strong locations are weighted more (see Section 4.4 for further discussion on this point).

4.3.2. Interactions between the two prosodic factors, boundary and Prominence

The results also showed some interaction effects between two prosodic factors, Boundary and Prominence, with three acoustic measures: V1 duration, V1 intensity, and VOT. Among these measures, V1 duration and V1 intensity showed no main effect of Boundary, but significant boundary effects were found only when the test word was in the focused condition. The interaction effect on VOT, however, revealed the opposite pattern. The boundary-induced VOT difference (longer IP-initially than IP-medially) turned out to be significant, this time in the prosodically 'weak' condition, that is, only when the test word was 'unfocused.' The lack of boundary effect on VOT in the focused condition can be interpreted as coming from a ceiling effect associated with the focused condition. In fact, among the three factors (Speaking style, Boundary, and Prominence), the Prominence factor induced the largest VOT difference (86 ms in the focused condition versus 69 ms in the unfocused condition, with 17 ms mean diff.). It is then plausible that there is no room left in the focused condition for additional lengthening of VOT coming from boundary.[2] This ceiling effect account is not new, but is consistent with Cho (2006) who reported that /ɑ/ in domain-initial CV in English showed spatial expansion only when unaccented while no such effect was found in accented condition, which was interpreted as coming from a ceiling effect.[3]

F1 and F2 measures for the three test vowels also showed some Boundary by Prominence interactions, but the patterns were inconsistent. For /i/, only F1 showed a significant Boundary by Prominence interaction due to the fact that boundary-induced lower F1 (positioning /i/ higher in the vowel space) was significant only in the prosodically strong, focused, condition, thus being consistent with the general Boundary by Prominence interaction pattern found with V1 duration and V1 intensity. For /a/, however, F2 showed a significant Boundary by Prominence interaction, but this time it was due to a significant focus effect only in the prosodically weak, IP-medial position. Other than these two interaction effects, no other formant measures showed comparable interaction effects.

Considering all these Boundary × Prominence interaction effects, although it is hard to make an across-the-board generalization about the direction of the interactions, the results do support the view that the two prosodically driven factors are not entirely independent, but that they do influence each other, as was observed with English (cf. de Jong, 2004; Cho & Keating, 2009). In particular, as far as the two general acoustic measures, V1 duration and V1 intensity, and one /i/-specific F1 measure are concerned, it appears that boundary effects on these measures are not fully realized when the test words occur in the prosodically weak, unfocused, condition, but they are reinforced in the prosodically strong, focused, condition.

4.4. Similarities and differences between the communicatively driven and the prosodically driven hyper-articulation

Another important question that the present study aimed to answer was whether and how the speakers would differentiate between the communicatively driven versus the prosodically driven hyper-articulation, and between the boundary-induced versus the prominence-induced hyper-articulation.

Let us first consider the similarities. Clear speech effects were found to be similar to prosodic strengthening effects (due to boundary and prominence) in various aspects in both temporal and spatial dimensions. First, all three factors gave rise to temporal expansion in strong environments (i.e., in clear speech, focused and IP-initial conditions), specifically with longer VOT and longer duration of the second syllable of the test words. Second, all three factors showed some degree of increase in loudness of the test words, especially on V2 of the test words whose intensity peak was augmented in clear speech, focused and domain-initial (IPi) conditions. Third, all three factors induced expansion of the acoustic vowel space, as reflected in some F1 and F2 values of the three test vowels which converged on the enlarged /i-a-u/ Euclidean area.

Although the three kinds of hyper-articulation with different sources have common acoustic properties, we have also seen that they differ in some other acoustic aspects. Most notably, the communicatively driven hyper-articulation was distinct from prosodically driven hyper-articulation in that while the F0 peaks for the first and the second vowels of the test word (C1V1C2V2) were higher in IP-initial (than IP-medial) and in focused (than unfocused) conditions, no increase in F0 peak and pitch range was found in clear speech condition. Furthermore, only the communicatively driven (clear speech-induced) hyper-articulation induced an increase in V1 duration across the board without further interactions with other factors. Finally, although all three kinds of hyper-articulation were associated with the vowel space expansion, the sources of the vowel expansion were different to some degree. As for /a/, all three factors induced higher F1 of /a/ (indicating a more lowered vowel quality, and therefore the greater mouth opening), but F2 of /a/ was influenced only by the Prominence factor, being higher in the focused condition than in the unfocused condition. As for /i/, the Boundary × Prominence interaction showed that F1 of /i/ was lower (therefore raised in the vowel space) in IP-initial CV condition than IP-medial CV condition, but again only when it occurred in the focused condition. As for /u/, F2 differentiated the boundary effect from the clear speech and the prominence effects, in that /u/ was associated with lower F2 (indicating a more retracted vowel quality) in clear speech and focused conditions, but not in IP-initial CV condition.

Regarding boundary-induced versus prominence-induced hyper-articulation patterns, there are a few acoustic parameters that still make them distinct. One clear difference was in the presence or absence of the main effect on V1 intensity—i.e., only the focused condition gave rise to an increase in V1 intensity. Another dimension in which they differed was F2 for /u/ and /a/, which was influenced by Prominence but not by Boundary: /u/ was associated with lower F2 (thus more retracted in the vowel space) in the focused condition, and /a/ was produced with higher F2 (thus more advanced in the acoustic vowel space) in the focused condition.[4]

The similarities and differences between three kinds of hyper-articulation and their interactions have implications for how the hypo- and hyper-articulation continuum is employed by the speaker in order to achieve communicatively driven and prosodically driven articulatory goals.

First, the similarities suggest that although the sources of hyper-articulation are in principle different, the three kinds of hyper-articulation can be characterized by some common hyper-articulation patterns, converging on one common goal—i.e., heightening of phonetic clarity, whether communicatively driven in the sense of the H&H theory or prosodically driven in the sense of prosodic strengthening, in order to get the intended linguistic message across to the listener successfully. The phonetic clarity

[2] Silva (2006) reported that mean VOT values of aspirated stops produced by young Koreans (born after 1980) were centered around 70 ms when stops were produced in a frame sentence. Given that our speakers were the same generation as those in Silva (2006), it is reasonable to assume independently that the mean VOT of 86 ms observed in the focused condition in the present study is already exceptionally long, leaving no much room for further expansion.

[3] Another similar ceiling effect could be found in Smiljanić and Bradlow (2008a). They found that although the durational contrast between the tense and the lax vowels in English was larger in clear speech in domain-final position compared to domain-medial position, the durational difference of the vowel as a function of the following consonant's voicing was not larger in clear speech. They suggested that the lack of clear speech effects in the latter case was due to a limit to the amount of combined lengthening effects of domain-final position and clear speech.

[4] In English, it has been suggested that the low back vowel /ɑ/, which may be specified with [+low, +back], is produced with *lower* F2 in prosodically strong environments, suggesting that the feature [+back] is enhanced (Cho, 2005). Unlike English /ɑ/, however, Korean /a/ is rather central (somewhat fronted) (Yang, 1996). We did not, therefore, have a priori prediction on how the place feature of /a/ in the backness dimension would be enhanced. Given that the F2 value in the focused condition was 1484 Hz, positing /a/ more fronted as opposed to 1430 Hz in the unfocused condition, and given that focus can be used as a diagnostic for what phonetic content is used to mark contrastiveness of the phoneme in a given language (e.g., de Jong & Zawaydeh, 2002; de Jong, 2004), the pattern suggests that the backness is not an important feature for /a/ in Korean. The more fronted quality of /a/ under focus may then be interpreted as reinforcing its centrality when focused.

obtained in a clear speech mode can be taken to enhance the intelligibility of the utterance for the successful delivery of the global linguistic message; the phonetic clarity of segments at edges of prosodic domains can be understood as marking prosodic boundaries for the successful delivery of prosodic phrasing information; and the phonetic clarity of focused words can be considered to highlight the semantic message embedded in the informational structure. After all, all three types of hyper-articulation can be taken together to eventually facilitate speech comprehension by virtue of similar patterns of articulatory strengthening. In this sense, the same hypo- to hyper-articulation continuum, which was originally thought to be effectively used for communicative purposes in the H&H theory (Lindblom, 1990), is also employed for prosodically driven 'localized' hyper-articulation (see relevant discussion in the Introduction section).

Second, despite the remarkable similarities arising with communicatively driven versus prosodically driven hyper-articulation, differences between them suggest that the different sources of hyper-articulation are indeed differentially manifested in phonetic output, supporting the view that they are phonetically encoded separately in speech planning (cf. Cho, in press; Cho & Keating, 2009; Keating, 2006; Keating & Shattuck-Hufnagel, 2002). However, the between-factor interactions imply that the separate encoding of different sources of hyper-articulation is not entirely independent, but that they are closely intertwined. Most notably, we have seen evidence that the communicatively driven hyper-articulation is prosodically modulated: When speakers are in need of enhancing their intelligibility in adverse speaking condition, for example, they do not simply hyper-articulate every segment in the utterance to an equal degree, but rather they appear to make more efforts to heighten the phonetic clarity of segments that are positioned in prosodically important landmarks—i.e., in domain-initial position or in focused words. This is indeed comparable to the conclusion made with regard to the interaction between clear speech and stress in English and Croatian in Smiljanić and Bradlow (2008a)—i.e., "a major feature of clear speech production (and a source of its increased intelligibility) is at the level of prosodic structure" (p. 108), which was based on their finding that clear speech effects were more robustly manifested in stressed syllables in both English and Croatian. This is also similar to the interaction effects between stress and focus reported in de Jong (2004), which showed that focus effects are mediated by stress in English, such that the vowel duration difference due to the following voicing contrast is enhanced under focus, but the effect is localized in stressed syllables. The heightened phonetic clarity of segments that arises at the prosodic boundary and with prominence is likely to help listeners with prosodic parsing, which in turn facilitates speech comprehension (Cho et al., 2007; Christophe et al., 2004, Cutler & Butterfield, 1992; Gow et al., 1996; Pitt & Samuel, 1990).

In short, the communicatively driven hyper-articulation can be taken to be prosodically modulated, in such a way that speakers hyper-articulate segments in selective and economical ways, taking into account the role of positions of segments in a given prosodic structure, which in turn increases effectiveness of speech comprehension on the listener's part. Note that this also has a broader implication for the H&H theory (Lindblom, 1990). In the original version of the H&H theory, the economy of articulatory effort was discussed as a speaker-oriented driving force for hypo-articulation in a casual communicative situation. We propose that the general principle of gestural economy can also apply to the communicative situation, which requires hyper-articulation (in a clear speech model), so that all else being equal, speakers make more efforts to strengthen prosodically important segments from which listeners would benefit more, while avoiding undue expenditure of effort by not putting too much effort to hyper-articulate segments in prosodically less important locations whose strengthening would not give rise to the same degree of communicative effectiveness for a given amount of effort.[5]

4.5. Issues on the role of language-specific prosodic/phonological structure on hyper-articulation

The present study also aimed to explore how language-specific prosodic and phonological system would influence the way that the hyper-articulation effect, whether communicatively or prosodically driven, is phonetically realized. In this section, we discuss the results that bear directly and indirectly on this issue.

4.5.1. Relationship between the size of the phonological inventory and the degree of vowel space expansion

One of the questions discussed at the outset of the paper was about the relationship between the size of the phonological vowel inventory and the degree of vowel space expansion in hyper-articulation environments. Given that Korean has a relatively small vowel inventory size with 7 contrastive vowels, and also given that the three test vowels /i,a,u/ are positioned separately in each section in the vowel space (high front, low central, and high back sections, respectively for /i,a,u/) with no adjacent vowel in that section, the principle of effort minimization (gestural economy) (cf. Liljencrants & Lindblom, 1972; Lindblom, 1986, 1990; Bradlow, 1995, 1996) would lead to no substantial vowel space expansion even in a situation where hyper-articulation is called for in Korean. We found, however, substantial expansion of the vowel space in all three types of hyper-articulation conditions. Although no decisive conclusion could be made until Korean data can be directly compared with data of other languages which have more crowded vowel spaces, the result, especially regarding the clear speech-induced vowel space expansion, is in line with the cases with English and Croatian (Smiljanić & Bradlow, 2005). It appears that speakers strive to maximize perceptual contrasts between vowels as much as possible in adverse communicative situations, overriding the general principle of effort minimization, which would otherwise function in not-so-adverse communicative situations. The vowel space expansion that arises especially with clear speech can therefore be thought of as one of the common strategies employed by speakers across languages to enhance their intelligibility, independently of vowel inventory size.

4.5.2. The scope of boundary-induced strengthening

Another question that is related to language-specificity was how far the boundary-induced strengthening effect can be extended beyond the segments immediately adjacent to the boundary, especially to the right of the prosodic boundary in Korean. Studies documented in the literature have so far shown mixed results on this (Barnes, 2002; Byrd, 2000; Byrd et al., 2006; Cho, 2006, 2008; Cho & Keating, 2001, 2009; Cho & McQueen, 2005; Fougeron & Keating, 1997; Keating et al., 2003; Krivokapić, 2007), but with a general consensus that domain-initial strengthening is not robust after the initial consonant. The present study

[5] de Jong and his colleagues (de Jong, 1991, 1995, 2004; de Jong & Zawaydeh, 2002; Silbert & de Jong, 2008) discuss how the communicative effectiveness can be achieved by hyper-articulation. For example, based on the finding that the durational difference of the vowel before a voiced versus a voiceless consonant is enhanced in English under stress, but not in Arabic, they proposed that the durational difference as a function of voicing of the following consonant is "linguistically specified" in English but not in Arabic, and that the specified phonetic contents of phonological contrast are most likely to be subject to hyper-articulation (de Jong & Zawaydeh, 2002; de Jong 2004). That is, what phonetic content is enhanced under hyper-articulation conditions is modulated by the linguistic roles of the phonetic content in a given language.

however showed that domain-initial strengthening in Korean is indeed extended into the vowel after the initial consonant as reflected in extreme F1 and F2 values and its accompanying vowel space expansion, and even beyond the first syllable as reflected in the increased V2 intensity and the longer second syllable duration. This suggests that as far as Korean is concerned, the scope of domain-initial strengthening is at least as large as the initial syllable, possibly including the second syllable (i.e., the entire bisyllabic test word). This extension of the domain-initial strengthening effect in Korean is clearly different from what has been found with English (e.g., Cho & Keating, 2009; Fougeron & Keating, 1997), which showed no robust evidence on the domain-initial strengthening effect beyond the initial consonant.

The apparent cross-linguistic difference between Korean and English prompts the question: What factors would influence determination of its domain of influence? Although these questions cannot be answered directly by the experimental findings of the present study, a possible determining factor that can be thought of is language-specific prosodic systems. Given the possibility that the lack of initial strengthening effects on the postconsonantal vowel in English is attributable to the role of the nucleus vowel which, in English, is arguably reserved for the stress-induced prominence realization (Barnes, 2002), our results suggest that languages without lexical stress and pitch accent such as Korean are associated with more robust domain-initial strengthening effects as its domain of influence is not restricted by the lexical prominence system (cf. Keating et al., 2003).

These results also have implications for the prosodic boundary gesture model of Byrd and Saltzman (Byrd, 2000, 2006; Byrd et al., 2000, 2006; Byrd & Saltzman, 2003; Saltzman, 1995). They proposed that the boundary influence on articulation can be understood as a result of the influence of so-called 'π-gesture' that is governed by prosodic constituency in the task dynamics model (e.g., Saltzman & Munhall, 1989). It is an abstract 'prosodic' gesture which determines articulatory movement speed by modulating the rate of the clock, which in turn controls articulatory activation of constriction gestures. The rate of the clock is assumed to decrease in proportion to the boundary strength, and as a result the articulatory movement at the juncture becomes slower, and possibly spatially larger. In the temporal domain, the π-gesture is anchored at a prosodic boundary, such that its clock-slowing effect is stronger at the juncture, dwindling farther from the edge. Although it appears extremely complicated to understand the complex results of the present study in terms of the local clock-slowing mechanism currently proposed in the theory of π-gesture, to the extent that the theory works, the results of the present study suggest that the scope of the influence of the π- gesture may vary from language to language, reflecting the language-specific prosodic system. For example, some languages such as Korean may operate on a larger scope when their prosodic structure is signaled primarily by boundary marking without restrictions associated with lexical stress and pitch accent, while languages such as English may have a reduced scope when the boundary marking is restricted by the system for prominence marking (cf. Barnes, 2002; Cho & Keating, 2009), limiting initial strengthening effects to the edges of prosodic domains. Much more work appears to be needed to determine how the theory of π-gesture explains the complicated patterns that arise with interactions of boundary and prominence effects and how its scope is modulated by language-specific prosodic systems.

The cross-linguistic differences in prosodic systems also prompt another important question about whether the boundary effect to the left (on the domain-final vowel) in Korean gives rise to more robust strengthening effects as compared with English. Beckman et al. (1992) found that in English /ɑ/ in "pop, opposing" (domain-final position) was more lengthened compared with /ɑ/ in "poppa, posing" (domain-medial position), but a spatial effect (e.g., jaw lowering) did not come with the final lengthening, which arguably makes the domain-final effect distinct from prominence(stress)-induced strengthening, which is characterized by both temporal and spatial expansions.

Again, however, given that the boundary effect is not restricted by stress in Korean, one would expect more robust strengthening effects in the final position as well, which is indeed what we found in the present study. Specifically, the higher F1 along with lengthening for /a/ in final position indicates that /a/ is produced with greater mouth opening, providing an additional case in which domain-final articulation is indeed subject to modifications in the spatial dimension.[6] This domain-final effect in Korean is indeed similar to stress effects in English reported in de Jong (2004) and in Arabic (de Jong & Zawaydeh, 2002). In exploring the prominence (stress) effects, de Jong (2004), showed that stress in English increases both F1 and duration of low vowels (e.g., /æ/) in a correlated fashion (similar to stress effects on /a/ in Arabic), which alluded to the possibility that the longer duration of the vowel would prevent the vocalic opening gesture from being truncated by the following consonantal closing gesture (i.e., stress eliminates gestural rephasing or target undershoot of the vocalic gesture due to an earlier initiation of the subsequent closing gesture). The spatio-temporal effect found with the domain-final /a/ in Korean could therefore be accounted for by the dynamical mechanisms similar to those of stress effects in English. Note, however, that Cho (2005) demonstrated that English vowels /i,ɑ/ in an open syllable are indeed produced with extreme articulations in domain-final position (e.g., a higher tongue position for domain-final /i/; a lowered and retracted tongue position for domain-final /ɑ/) relative to domain-medial position. Therefore, whether the robust final strengthening effect found in Korean is due to language-specific prosodic system or not is subject to further corroboration.

4.5.3. The cross-linguistic versus language-specific use of F0 in hyper-articulation contexts

Focus is often phonetically realized by raising F0 of the words that receive focus, and in languages with lexical stress such as English, its phonetic realization of focus occurs in a lexically stressed syllable (e.g., Xu & Xu, 2005). Smiljanic (2006) also showed that, although Serbian and Croatian differ in terms of pitch-peak alignment pattern, depending on whether or not the language uses a lexical contrast between 'rising' and 'falling' accents, they both exhibited an effect of raising F0 in focused conditions, which was realized in stressed syllables as in English. Korean with no lexical stress showed similar focus effects, showing that F0 of vowels was raised in the focused condition. Combined, these results suggest that raising F0 is one of the cross-linguistically applicable strategies to enhance prominence as it may increase perceptual saliency of the focused item (cf. Warner, Otake, & Arai, 2010). However, Korean differs from the languages with stress in that the focus effect on F0 was not localized to a specific syllable, but it was found on both the first and the second vowels (V1, V2) of the bisyllabic test word. That is, with no mediation of focus

[6] Given the greater mouth opening associated with domain-final /a/ as reflected in lower F1, one might expect that /a/ would be louder domain-finally than domain-medially, as is predicted by sonority expansion (cf. Beckman et al., 1992; de Jong, 1995). Interestingly, however, we found the opposite—i.e., the peak intensity of the final vowel /a/ was weaker in IP-final position than in IP-medial position. This, however, can be accounted for in terms of respiratory declination. In general, F0 and acoustic amplitude tend to decline gradually over the course of utterances as a result of declination of the subglottal (pulmonary) air pressure (Gelfer, 1987; Gelfer, Harris, & Baer, 1987; 't Hart, Collier, & Cohen, 1990; Ladd, 1984). Given that the declination of subglottal and laryngeal articulation is generally independent from the supralaryngeal declination (Krakow, Bell-Berti, & Wang, 1994), the domain-final vowel may well be strengthened in terms of supralarygneal articulation (as reflected in higher F1 for /a/), but weakened in its loudness due to the general respiratory declination.

realization by stress, Korean appears to apply the F0 raising strategy across the entire word which receives focus.

Another language-specific aspect of using F0 could be thought of in terms whether speakers would expand pitch range in a clear speech mode. As introduced at the outset of the paper, previous studies on both English and Croatian showed a global expansion of pitch range (Smiljanić & Bradlow, 2005, 2008a), which may be taken to be cross-linguistically applicable (Smiljanić & Bradlow, 2009). Given that the pitch range expansion was found with Indo-European languages with lexical stress and pitch accent, we asked whether the same would hold for a typologically and prosodically different language such as Korean which does not employ lexical stress and pitch accent. Our results with Korean indeed revealed no clear speech effect on pitch range, which was calculated globally over the course of the entire utterance. F0 maxima during the utterances were not significantly higher in the clear speech condition, either.[7] Higher pitch that usually contributes to the pitch range expansion may of course be perceptually salient, so that listeners may be more attentive to higher than lower pitched speech. The results, however, suggest that the expansion of the overall pitch range is not a universally applicable characteristic of clear speech, but rather is determined language-specifically. It is therefore plausible that speakers with different linguistic backgrounds implement different strategies conditioned by language-specific prosodic systems: Speakers of languages (e.g., English) which use pitch systematically (e.g., marking both lexical and higher-order prosodic structures) are more likely to make use of the pitch range expansion to enhance the overall intelligibility of the utterance, whereas speakers of languages which do not employ lexically determined stress and pitch accent in its prosodic system appear not to utilize the pitch range expansion as much. After all, F0 in the latter type of language carries less lexical information than the former type of language.

5. Conclusion

Speech production is by nature variable, so that segments constituting an utterance are produced in variable forms depending on various factors. Some of them cause random variation such as speakers' anatomical differences (gender and age differences inclusive) from which linguistic aspects are hardly inferable. But some other factors function systematically so that its fine phonetic detail reflects some underlying principles of speech production. Variable communicative situations induce speech variation, for example, but it is realized along the hypo- and hyper-articulation continuum in a linguistically relevant way, reflecting the tug-of-war between output-oriented and production-oriented constraints, as Lindblom (1990) puts it. Prosodically driven speech variation is another example, whose fine-grained phonetic detail is thought to signal higher-order linguistic structures such as prominence and prosodic phrasing in the prosodic system of the language.

In the present acoustic study, we have provided concrete evidence in Korean for how such speech variation coming from communicative needs and prosodic structure are indeed systematically manifested in fine-grained phonetic details. The hyper-articulation patterns coming from different sources (clear speech, prosodic boundary, and prominence or focus) converge on heightening phonetic clarity of the utterance, guided by the principle of contrast maximization. They, in turn, would lead to facilitation of speech comprehension by virtue of increasing perceptual saliency, signaling prosodic boundaries, and marking information locus. They are nevertheless phonetically distinct in some aspects, supporting the view that different sources of hyper-articulation are encoded separately in speech planning. However, the communicatively driven factor (i.e., Speaking style) and the prosodically driven factors (i.e., Boundary and Prominence) do not function independently, but they interact with each other. Most importantly, the communicatively driven hyper-articulation is prosodically modulated, such that speakers appear to enhance the intelligibility of the utterance in a clear speech mode not simply by hyper-articulating every single segment to the same degree, but by highlighting phonetic hallmarks that reflect important prosodic structural information. The present study also provided further evidence that hyper-articulation is conditioned by phonological/prosodic systems of a given language. Most notably, our results imply that languages such as Korean, which do not employ lexical stress and pitch accent in their prosodic systems, may utilize boundary-induced strengthening more (as reflected in the extension of domain-initial strengthening beyond the initial segment and the robust spatio-temporal expansion both domain-initially and domain-finally) than other languages such as English with lexical stress and pitch accent. Limited use of pitch found in a clear speech mode in Korean also supports this assumption.

All in all, the present study implies that the hypo- to hyper-articulation continuum is employed in speech production not only for communicatively driven speech modification but also for prosodically driven variation in a linguistically relevant systematic way, and that the systematic variation is modulated by language-specific phonological and prosodic systems.

Acknowledgments

We thank the speakers who participated as subjects in our experiment, and James H. Clitheroe for proofreading earlier versions of the paper. We are also grateful for the Editor, Ken de Jong, and other reviewers for constructive comments on an earlier version of this paper.

Appendix. Experimental materials for /p^{h}i/ and /p^{h}u/

Experimental sentences with the test syllable /p^{h}i/ in **/p^{h}ipu/** in IP-initial focused (a), IP-initial unfocused (b), IP-medial focused (c), IP-medial unfocused (d) conditions, and /p^{h}u/ in **/p^{h}utiŋ/** in IP-initial focused (e), IP-initial unfocused (f), IP-medial focused (g), and IP-medial unfocused (h) conditions. The test word is underlined, and focused items are in bold.

(a) IP-initial /p^{h}i/: Focused
Q: *t^{h}ehi-nɨn ʌntʃena [IP mʌli-hako sont^{h}op-ɨl kwanliha-ni]?*

---

[7] As pointed out by a reviewer, the null effect of clear speech on F0 measures such as pitch range and F0 maxima may be interpretable alternatively as coming from the laboratory context with EMA and/or the elicitation technique. While caution is needed in interpreting the data, we believe that the F0 results still reflect how Koreans express clear speech for the following reasons. First, all three of us as trained Korean ToBI transcribers had the same impression that Koreans do not manipulate pitch in a clear speech mode. Second, all three F0 measures (F0 peak, F0 minimum, and Pitch range) independently yielded the same null effect of clear speech, while there were robust and meaningful clear speech effects in both global measures and local measures on test words. There is no a priori reason to reject only the F0 results as artifacts of the experimental procedures while other measures are viably interpretable. Third, recall that our data were collected from 5 speakers with EMA and 3 speakers without EMA. Inspection of individual speakers' behaviors (by conducting *t*-test for each speaker) revealed that all three speakers without EMA showed no pitch expansion in clear speech like those with EMA, confirming that the null clear speech effects did not come from potential constraints with EMA. Fourth, previous studies on clear speech effects in English and Croatian (Smiljanić & Bradlow, 2005, 2008a) used read speech in a laboratory setting. While these studies still have limitations coming from elicitation techniques, we made effort to eliminate possible read speech effects by eliciting utterances without scripts.

Taehee-Top. always hair-and nail care-Acc. get-Q
"Does Taehee always get hair and nail care?"
A: *ani, tʰehi-nɨn ʌntʃena [IP* ***pʰipu****-hako sontʰop-ɨl kwanlihe]* (Test Sentence)
No, Taehee-Top. always skin-and nail care-Acc. get
"No, Taehee always gets SKIN and nail care."

(b) IP-initial /pʰi/: Unfocused
Q: *tʰehi-nɨn ʌntʃena [IP pʰipu-hako mʌli-lɨl kwanliha-ni]?*
Taehee-Top. always skin-and hair care-Acc. get-Q
"Does Taehee always get skin and hair care?"
A: *ani, tʰehi-nɨn ʌntʃena [IP pʰipu-hako* ***sontʰop****-ɨl kwanlihe].* (Test Sentence)
No, Taehee-Top. always skin-and nail care-Acc. get
"No, Taehee always gets skin and NAIL care."

(c) IP-medial /pʰi/: Focused
Q: *tʰehi-nɨn ʌntʃena, [IP aloma tʃʰehiʌŋkwanlipʌp-ilaŋ joka-lil pews*ni]?*
Taehee-Top. always aroma body care-and yoga-Acc. study-Past-Q
"Did Taehee always study aroma body care and yoga?"
A: *ani, tʰehi-nɨn ʌntʃena, [IP aloma* ***pʰipu*** *kwanlipʌp-ilaŋ joka-lɨl pew-ʌs*ʌ]* (Test Sentence)
No, Taehee-Top. always aroma skin care-and yoga-Acc. study-Past
"No, Taehee always studied aroma SKIN care and yoga."

(d) IP-medial /pʰi/: Unfocused
Q: *tʰehi-nɨn ʌntʃena, [IP hwaŋtʰo pʰipu kwanlipʌp-ilaŋ joka-lɨl pew-ʌs*-ni]?*
Taehee-Top. always yellow mud body care-and yoga-Acc. study-Past-Q
"Did Taehee always study yellow mud body care and yoga?"
A: *ani, tʰehi-nɨn ʌntʃena, [IP* ***aloma*** *pʰipu kwanlipʌp-ilaŋ joka-lɨl pew-ʌs*ʌ]* (Test Sentence)
No, Taehee-Top. always aroma skin care-and yoga-Acc. study-Past
"No, Taehee always studied AROMA skin care and yoga."

(e) IP-initial /pʰu/: Focused
Q: *hekjo-nɨn ʌntʃena, [IP kwail-hako p*aŋ-ɨl mʌk-ni]?*
Haegyo-Top. always fruit-and bread-Acc. eat-Q
"Does Haegyo always eat fruits and bread?"
A: *ani, hekjo-nɨn ʌntʃena, [IP* ***pʰutiŋ****-hako p*aŋ-ɨl mʌkʌ].* (Test Sentence)
No, Haegyo-Top. always pudding-and bread-Acc. eat
"No, Haegyo always eats PUDDING and bread."

(f) IP-initial /pʰu/:Unfocused
Q: *hekjo-nɨn ʌntʃena, [IP pʰutiŋ-hako kimpap-ɨl mʌk-ni]?*
Haegyo-Top. always pudding-and kimpap-Acc. eat-Q
"Does Haegyo always eat pudding and kimpap (A Korean style California roll)?"
A: *ani, hekjo-nɨn ʌntʃena, [IP pʰutiŋ-hako* ***p*aŋ****-ɨl mʌkʌ].* (Test Sentence)
No, Haegyo-Top. always pudding-and bread-Acc. eat
"No, Haegyo always eats pudding and BREAD."

(g) IP-medial /pʰu/:Focused
Q: *hekjo-nɨn ʌntʃena, [IP panana pʰai setʰɨ-laŋ sakwa-lɨl mʌk-ʌs*-ni]?*
Haegyo-Top. always banana pie set-and apple-Acc. have-Past-Q
"Does Haegyo always eat banana pie set with an apple?"
A: *ani, hekjo-nɨn ʌntʃena, [IP panana* ***pʰutiŋ*** *setʰɨ-laŋ sakwa-lɨl mʌk-ʌs*ʌ].* (Test Sentence)
No, Haegyo-Top. always banana pudding set-and apple-Acc. have-Past
"No, Haegyo always had the banana PUDDING set with an apple."

(h) IP-medial /pʰu/: Unfocused
Q: *hekjo-nɨn ʌntʃena, [IP maŋko pʰutiŋ setʰɨ-laŋ sakwa-lɨl mʌk-ʌs*-ni]?*
Haegyo-Top. always mango pudding set-and apple-Acc. have-Past-Q
"Did Haegyo always have the mango pudding set with an apple?"
A: *ani, hekjo-nɨn ʌntʃena, [IP* ***panana*** *pʰutiŋ setʰɨ-laŋ sakwa-lɨl mʌk-ʌs*ʌ]* (Test Sentence)
No, Haegyo-Top. always banana pudding set-and apple-Acc. have-Past
"No, Haegyo always had the BANANA pudding set with an apple."

References

Barnes, J. A. (2002). *Positional Neutralization: A phonologization approach to typological patterns*. Unpublished PhD Dissertation. University of California, Berkeley.

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Laboratory phonology, Vol. III: Phonological structure and phonetic form* (pp. 7–33). Cambridge: Cambridge University Press.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. *Laboratory phonology Vol. II: Gesture, segment, prosody*, pp. 68–86.

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, *97*(3), 1916–1924.

Bradlow, A. R. (1996). A perceptual comparison of the /i/-/e/ and /u/-/o/ contrasts in English and Spanish: Universal and language-specific aspects. *Phonetica*, *53*, 55–85.

Bradlow, A. R. (2002). Confluent talker- and listener-related forces in clear speech production. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology*, 7 (pp. 241–273). Berlin/New York: Mouton de Gruyter.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, *112*, 272–284.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language and Hearing Research*, *46*(1), 80–97.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M. E. Beckman (Eds.), *Laboratory phonology Vol. I: between the grammar and the physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, *49*, 155–180.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, *57*(1), 3–16.

Byrd, D. (2006). Relating prosody and dynamic events: Commentary on the papers by Cho, Navas, and Smiljanić. In L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory phonology 8: Varieties of phonological competence* (pp. 549–561). Berlin/NY: Mouton de Gruyter.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Laboratory phonology*, Vol. V (pp. 70–87). Cambridge: Cambridge University Press.

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*(2), 149–180.

Cho, T. (1996). *Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents*. MA Thesis. University of Texas at Arlington.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*(2), 141–176.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ,i/ in English. *Journal of the Acoustical Society of America*, *33*, 121–157.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in laboratory phonology,*, Vol. 8 (pp. 519–548). Berlin: Mouton de Gruyter.

Cho, T. (2008). Prosodic strengthening in transboundary V-to-V lingual movement in American English. *Phonetica*, *65*, 45–61.

Cho, T. (in press). The phonetics-prosody interface in laboratory phonology. In: N. C. Kula, B. Botma, & K. Nasukawa (Eds.), The continuum companion to phonology. London/New York: Continuum.

Cho, T., & Jun, S.-A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, *36*, 31–44 (Also appeared in *UCLA Working Papers in Phonetics*, *99*, 57–70.).

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155–190.

Cho, T., & Keating, P. A. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*, 466–485.

Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*, 121–157.

Cho, T., McQueen, J., & Cox, E. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243.

Cho, T., Yoon, Y., & Kim, S. (in preparation). A kinematic study of prosodic boundary effects on the vowel /i/ articulation in Korean.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*, *51*, 523–547.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*, 218–236.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, *97*(1), 491–504.

de Jong, K. J. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*, 493–516.

de Jong, K. J., & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, *30*, 53–75.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, *89*(1), 369–382.

Flemming, E. (1995). *Auditory representations in phonology*. Los Angeles: University of California.

Flemming., E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, *18*(1), 7–44.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109–135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*(6), 3728–3740.

Fowler, C. A. (1995). Acoustic and kinematic correlates of contrastive stress accent in spoken English. In F. Bell-Berti, & J. J. Raphael (Eds.), *Producing speech: Contemporary issues: For Katherine Safford Harris* (pp. 355–373). Melville, Long Island, New York: AIP Publishing.

Gelfer, C. E. (1987). *A simultaneous physiological and acoustic study of fundamental frequency declination*. City University of New York.

Gelfer, C. E., Harris, K. S., & Baer, T. (1987). Controlled variables in sentence intonation. In T. Baer, C. Sasaki, & K. Harris (Eds.), *Laryngeal function in phonation and respiration* (pp. 422–435). Boston: College-Hill.

Gow, D. W., Melvold, J., & Manuel, S. (1996). How word onsets drive lexical access and segmentation: Evidence from acoustics, phonology and processing. In H. T. Bunnell, & W. Idsardi (Eds.), *Proceedings of ICSLP 96 (CD-ROM)*. Philadelphia: University of Delaware & Alfred I. duPont Institute.

't Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.

Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology*, *9*(3), 234.

Jakobson, R., Fant, G., & Halle, M. (1965). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Ph.D. Dissertation. Ohio State University.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell, & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Laboratory phonology*, Vol. IV (pp. 235–253). Cambridge, UK: Cambridge University Press.

Jun, S.-A. (2000). K-ToBI (Korean ToBI) Labelling Conventions (Version 3.1). 〈http://www.linguistics.ucla.edu/people/jun/ktobi/K-tobi.html〉.

Kang, K.-H. (2010). *Clear speech production and perception of Korean stops and the sound change in Korean stops*. Ph.D. Dissertation. University of Oregon.

Kang, K.-H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America*, *124*(6), 3909–3917.

Keating, P. A. (2006). Phonetic encoding of prosodic structure. In J. Harrington, & M. Tabain (Eds.), *Speech Production: Models, Phonetic Processes, and Techniques* (pp. 167–186). New York and Hove: Psychology Press.

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory phonology, Vol. IV: Phonetic interpretation* (pp. 145–163). Cambridge, UK: Cambridge University Press.

Keating, P. A., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA working papers in phonetics*, *101*, 112–156.

Kim, S. (2003). Domain-initial strengthening of Korean fricatives. In S. Kuno (Ed.), *Harvard studies in Korean linguistics*, *Vol. IX*.

Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America*, *125*(5), 3373–3386.

Klatt, D. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics*, *3*, 129–140.

Krakow, R. A., Bell-Berti, F., & Wang, Q. E. (1994). Supralaryngeal declination: Evidence from the velum. In F. Bell-Berti, & L. J. Raphael (Eds.), *Producing speech: Contemporary issues (a Festschrift for Kathering Safford Harris)* (pp. 333–354). Woodbury, New York: AIP Press.

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*, 162–179.

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, *35*, 301–320.

Ladd, D. R. (1984). Declination: A review and some hypotheses. *Phonology Yearbook*, *1*, 53–74.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language*, *48*, 839–862.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*, 1773–1781.

Lindblom, B. (1986). Phonetic universals in vowel systems. *Experimental phonology*, 13–44.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech Production and Speech Modeling*, *55*, 403–439.

Onaka, A. (2003). Domain-initial strengthening in Japanese: An acoustic and articulatory study. In: Proceedings of the 15th international congress of phonetic sciences. (pp. 2091–2094). Barcelona, Spain.

Onaka, A., Watson, C., Palethorpe, S., & Harrington, J. (2003). An acoustic analysis of domain-initial strengthening effect in Japanese. In S. Palethorpe & M. Tabain (Eds.), *Proceedings of the 6th international seminar on speech production.* (pp. 201–206). Sydney, Australia.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, *29*(4), 434–446.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty, & D. R. Ladd (Eds.), *Laboratory phonology, Vol. II: Gesture, segment, prosody* (pp. 90–117). Cambridge, UK: Cambridge University Press.

Pitt, M. A., & Samuel, A. G. (1990). Attentional allocation during speech perception: How fine is the focus?. *Journal of Memory and Language*, *29*, 611–632.

Saltzman, E. (1995). Dynamics and coordinate systems in skilled sensorimotor activity. In R. Port, & T. van Gelder (Eds.), *Mind as motion: Dynamic, behavior, and cognition* (pp. 149–173). Cambridge, MA: MIT Press.

Saltzman, E., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.

Schum, D. J. (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, *7*(3), 212.

Shattuck-Hufnagel, S., & Turk, A. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*, 193–247.

Sheskin, D. J. (2000). *Handbook of parametric and nonparametric statistical procedures* ((2nd ed.). Boca Raton: Chapman & Hall/CRC.

Shin, J., & Cha, J.-E. (2003). *The sound system of the Korean language* (In Korean). Hankookmoonhwasa, Seoul. isbn:89-5726-014-593710.

Silbert, N. H., & de Jong, K. J. (2008). Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production. *Journal of the Acoustical Society of America*, *123*, 2769–2779.

Smiljanic, R. (2006). Early vs. late focus: Pitch-peak alignment in two dialects of Serbian and Croatian. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in laboratory phonology*,, Vol. 8 (pp. 495–518). Berlin: Mouton de Gruyter.

Smiljanic´, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America*, *118*(3), 1677–1688.

Smiljanić, R., & Bradlow, A. R. (2008a). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics*, *36*(1), 91–113.

Smiljanić, R., & Bradlow, A. R. (2008b). Temporal organization of English clear and plain speech. *Journal of the Acoustical Society of America*, *124*(5), 3171–3182.

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Linguistics and Language Compass*, *3*(1), 236–264.

Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni, & R. Remez (Eds.), *The handbook of speech perception* (pp. 207–235). Malden/Oxford: Blackwell.

Van Heuven, V. J. (1994). What is the smallest prosodic domain?. In P. A. Keating (Ed.), *Phonological structure and phonetic form: Papers in laboratory phonology*, Vol. III (pp. 76–97). Cambridge: Cambridge University Press.

Warner, N., Otake, T., & Arai, T. (2010). Intonational structure as a word boundary cue in Japanese. *Language and Speech*, *53*, 107–131.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*, 1707–1717.

Xu, Y., & Wang, M. (2009). Organizing syllables into groups—Evidence from F0 and duration patterns in Mandarin. *Journal of Phonetics*, *37*, 502–520.

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, *33*, 159–197.

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