



Kim, Dong Jin & Kim, Okgi & Park, Hanyong. 2023. Prosodic realization of identificational and contrastive focus in Korean multiple accusative constructions. *Glossa: a journal of general linguistics* 8(1), pp. 1–35. DOI: <https://doi.org/10.16995/glossa.8242>



Open Library of Humanities

## Prosodic realization of identificational and contrastive focus in Korean multiple accusative constructions

**Dong Jin Kim**, University of Wisconsin-Milwaukee, US, [kim252@uwm.edu](mailto:kim252@uwm.edu)

**Okgi Kim**, Kyung Hee University, KR, [okgikim@khu.ac.kr](mailto:okgikim@khu.ac.kr)

**Hanyong Park**, University of Wisconsin-Milwaukee, US, [park27@uwm.edu](mailto:park27@uwm.edu)

---

This study investigated how two different kinds of prominence-induced prosodic strengthening, identificational focus and contrastive focus, are phonetically realized in Seoul Korean. In the experiment, 16 native speakers produced target sentences presented in three different prosodic conditions: default reading, NP1-contrast, and NP2-contrast. The production data presented three crucial characteristics in the phonetic manifestations of prominence-induced prosodic strengthening in Seoul Korean. First, identificational focus gave rise to longer duration and higher pitch over the constructions under the focus. Second, contrastive focus brought about modification of temporal, pitch, and prosodic structures as well as pre- and post-focus compression. Finally, the two different types of prominence marking interacted with each other; the prosodic realization of contrastive focus overrode the prosodic realization of identificational focus, which was in line with the formal (syntactic) distinction of the two types of focus. To conclude, the current study demonstrates the interplay of information structure and discourse contexts in prosodic strengthening effects induced by prominence, especially in Seoul Korean, which lacks lexical stress and pitch accent.

---

*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by the Open Library of Humanities. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

**OPEN ACCESS**



## 1 Introduction

The information structure of a specific sentence interacts with phonetic realizations of speech sounds. Information structure is considered part of the dimension of common ground management (Chafe 1976; Féry & Krifka 2008), a method of representing the mutually shared information between interlocutors in communication (Karttunen 1974; Stalnaker 1974; Lewis 1979). One manifestation of a given sentence's information structure is focus because information structure, often referred to as information packaging, explains prosodic prominence through the marking of activation states (Chafe 1976; 1987).

Diverse types of focus have been widely discussed in semantics/syntax-oriented and phonetics/phonology-oriented research (see Shattuck-Hufnagel & Turk 1996; Dik 1997; Zimmermann & Onea 2011 for an overview). Such research has revealed that the phonetic realization of focus is mediated by language-specific prosodic systems. For example, in head-prominence languages such as German and English, prominence-induced prosodic strengthening (i.e., focus marking) is realized through pitch accent (Beckman & Pierrehumbert 1986). In these languages, word-level prominence, which is expressed as stress, is assigned to prosodic heads. The stressed syllables are possible locations for the assignment of phrasal-level prominence, pitch accent. On the other hand, the prominence-induced prosodic strengthening in edge-prominence languages (e.g., Seoul Korean and accentless dialects of Japanese such as Kobayashi, Koriyama, and Yamagata) is related to prosodic phrasing locating the targets of focus to the edge of a word or phrase (Jun 2014). These languages lack word-level prominence in their prosodic systems. Even though a considerable amount of research has focused on the realizations of focus on speech sounds (see Cho 2015; 2016 for an overview of previous research), a deeper understanding of how the interplay of various types of focus is prosodically manifested has not been achieved. Moreover, the number of edge-prominence language studies has never been closer to the number of studies on head-prominence languages such as English and German. Therefore, a study of how the interplay of information structure is prosodically expressed in Seoul Korean would enhance our understanding of the interaction between higher-order linguistic representations and their realizations in edge-prominence languages in which prominence is marked at the edge of a phrase.

The current study explores the production of identificational focus and contrastive focus<sup>1</sup> in Seoul Korean. The different focus types were tested within a structure of so-called multiple accusative constructions (MAC) in Korean. The structures of MACs allowed us to examine the effects of information structures derived from the two consecutive Noun Phrases (NPs) marked

---

<sup>1</sup> We used the term “contrastive focus” here, adopting Steube's (2001) view that correction is realized by using contrastively focused expressions.

by accusative case. Furthermore, the discourse context induced by corrective questions allowed us to observe the interplay of such information structure and focus types.

The general goals of the current study are as follows. First, the present study explores the realization of identificational focus driven by the semantic weight factor. Semantic weight has been discussed as one of the factors influencing the prosodic manifestations of Accentual Phrase (AP) in Korean (Jun 1993). Words with a generic meaning are semantically light or ‘empty’ compared to words with compositional meaning features, which are semantically richer. For example, a ‘guy’ or a ‘man’ is semantically lighter than a ‘friend’ or a ‘policeman’ (Bolinger 1972). According to Jun (1993), the accentual phrasing in Korean can be affected by the informativeness or semantic weight of the words; a semantically richer head noun (e.g., a ‘president’) of a relative clause tends to constitute an AP on its own, whereas a semantically light head noun (e.g., a ‘person’) does not unless it is focused. Because there has not been a significant amount of production research testing the effects of identificational focus driven by semantic properties of words, a study examining the effects of such factors on prosodic realization would contribute to our understanding of the interaction between semantic factors and information structure represented in prosody.

Second, the present study expands our knowledge of prosodic strengthening driven by contrastive focus, especially in edge-prominence languages. A larger volume of the research on phonetic characteristics of prominence-induced strengthening focuses on Indo-European languages; much less research has addressed the same characteristics in non-European languages (e.g., Korean). Furthermore, the results summarized in the previous literature demonstrated confounding patterns, motivating us to conduct further production studies on prosodic realizations of focus in Korean. On the one hand, some studies reported significant focus effects on phonetic realizations and phrasing in Seoul Korean. A growing body of research found remarkable focus marking employing longer duration, higher fundamental frequency (F0), and increasing intensity (Oh 2001; Lee & Xu 2010; Cho et al. 2011; Choi et al. 2018; Choi et al. 2020). In addition, Jeon & Nolan (2017) reported that focused constituents often initiate higher prosodic boundaries; a particular constituent in a default condition tends to be located in an AP-medial position as a prosodic word, whereas the corresponding constituent in a focus condition is more likely to initiate a new AP or Intonation Phrase (IP). This result was in accordance with the previous studies demonstrating that the minimal boundary condition for the phrase with focused constituents is an AP (Jun & Lee 1998) or IP (Jun & Kim 2007). On the other hand, several studies found comparatively less prominent focus effects in Korean regarding their phonetic realizations and phrasing. (Lee 2017; Lee & Cho 2020). For example, Lee & Cho (2020) pointed out that focus effects in Seoul Korean are neither produced nor perceived saliently. Their results further revealed that the focus effects on phrasal initial low tone were more confusing than those on phrasal initial high tone in production and perception. Choi et al. (2020) also showed that focused constituents

are able to be realized in AP-medial positions, questioning the dependence of focus marking on boundary marking in Korean. That is, a focused constituent in Korean is signaled not necessarily by a phrasing (i.e., an initiation of a new AP or IP) but by a direct regulation of fine-grained phonetic details (i.e., higher F0, longer duration, and higher intensity) (Cho 2022).

Nonetheless, the need persists to consider the different methodologies of the experiments to understand the mixed patterns of focus effects across various research studies. On the one hand, Cho et al. (2011) and Choi et al. (2018; 2020) examined focus effects by comparing contrastive-focused constituents with defocused constituents. On the other hand, Jun & Kim (2007), Jun & Lee (1998), Lee (2017), Lee & Cho (2020), Lee & Xu (2010), and Oh (2001) compared the focused constituents with the constituents produced in a default reading condition. Such comparisons only in a default reading condition, however, may not always be optimal to investigate the effects unless every possible source of focus has been considered. For example, if the target items are located at VP-initial positions (as in Lee 2017; Lee & Cho 2020), they are more likely to demonstrate prosodic strengthening even in a default reading condition (Jun et al. 2006; Jun & Kim 2007). This may obscure the true effect of contrastive focus due to the confounding factors coming from another type of focus realizations. Therefore, it is reasonable to examine the effects of prosodic strengthening driven by contrastive focus, for example, phrasing and phonetic manifestations, considering the reference prosodic conditions and the positions of the focused constituents within a VP.

Third, another important goal of the present study is to explore the interplay of identificational focus and contrastive focus. Although identificational focus and contrastive focus are often regarded as subtypes of a whole range of focus types, there is a tendency to treat contrastive focus as being motivated by different semantic representations (Halliday 1967; Rochemont 1986; Lambrecht 1994; Kratzer & Selkirk 2007). Specifically, identificational focus is a comparatively weaker version of focus that represents a subset of the set given by the discourse context, whereas contrastive focus excludes alternatives from discourse in which multiple possibilities exist (Kiss 1998; Kratzer & Selkirk 2007; Katz & Selkirk 2011). Thus, the contrastive focus requires structural licensing such as morphological, syntactic, or prosodic markings (Vallduví & Vilkkuna 1998; Molnár 2001; Gussenhoven 2004). Because the focus types exhibit semantic and pragmatic differences, it is reasonable to assume that contrastive focus is an independent notion from other types of focus (Vallduví & Vilkkuna 1998; Kenesei 2006; Neeleman & van de Koot 2008). In French, an edge-prominence language, contrastive focus is prosodically marked (Jun & Fougeron 2000; Dohen & Loevenbruck 2004; Klok et al. 2018). We will explore the issue by examining how identificational focus is marked differently from contrastive focus in Korean, another edge-prominence language.

Overall, the goal of the current study is to explore the use of prominence-induced prosodic strengthening by speakers of Seoul Korean. In particular, we examine prosodic realizations of two different types of focus: identificational focus and contrastive focus. This will provide us a

deeper understanding of prosodic strengthening driven by focus. In the following, we will discuss specific research questions and corresponding predictions based on previous research.

The first question is whether identificational focus driven by semantic features is prosodically realized on MACs in Seoul Korean. Identificational focus is usually derived when a subset of the set already given to the discourse is introduced (Kiss 1998). Identificational focus can be located in any part of a sentence and marked by pitch accent in head-prominence languages such as English (Bolinger 1972; Kiss 1998). In Korean, Jun (1993) proposed a similar phenomenon, in which the effect of inherent semantic weight was found in the prosodic realization. She suggested that when new information is introduced to the discourse in Korean, the word is focused and marked by prosodic phrasing. Jun & Kim (2007) explored the phonetic realizations of identificational focus (cf. VP focus in Jun & Kim 2007) induced by answering wh-questions, and their results showed a preference for a higher level of prosodic phrasing (i.e., having IPs rather than APs) on the VP-initial arguments along with a comparatively lower frequency of dephrasing on the post-focus string. Consequently, the peak of the F0 was higher for the constituents with identificational focus located in VP-initial positions (cf. note that their verbs possessed double objects since the verbs were ditransitive verbs).<sup>2</sup>

Therefore, it is possible to hypothesize that the identificational focus driven by information structure is phonetically marked. That is, the constituents under identificational focus are phonetically realized with higher F0 and a tendency to initiate a higher prosodic boundary in Seoul Korean. In this vein, the current study examines the phonetic manifestations of identificational focus on MACs. Among different types of MACs, we will adopt Type-subtype and Subtype-type MACs (Park 2013; Yoon 2015). For example, the Type-subtype MAC in (1) involves the hypernym *koki* ‘meat’ and the hyponym *koni* ‘duck.’ The hypernym designates a general category of the object, and the hyponym designates a specific kind of meat.

- (1) Mimi-ka koki-lul koni-lul p\*alli malk-as\*-a.  
 Mimi-NOM meat-ACC duck-ACC quickly eat-PST-DECL  
 ‘Mimi quickly ate koki (meat), koni (duck).’

Since there exists a taxonomic hierarchy between the two objects (Park 2013; Yoon 2015), the information in the hypernym (i.e., *koki*) is regarded as broad information, and the information in the hyponym (i.e., *koni*) is considered as specific information. The semantic relation between the

---

<sup>2</sup> The target sentences presented in Jun & Kim (2007: 1278) were as below:

- (i) Set 1: Subject + Indirect Object + Direct Object + Verb  
 Subject + Direct Object + Indirect Object + Verb
- (ii) Set 2: Subject + Locative + Direct Object + Verb  
 Subject + Direct Object + Locative + Verb

two objects remains constant in Subtype-type MACs where the word order is reversed between the two objects. Consider (2), where *koni*, specific information precedes *koki*, broad information.

- (2) Mimi-ka koni-lul koki-lul p<sup>\*</sup>alli mAk-As<sup>\*</sup>-A.  
 Mimi-NOM duck-ACC meat-ACC quickly eat-PST-DECL  
 ‘Mimi quickly ate koni (duck), koki (meat).’

Therefore, we hypothesize that the locus of identificational focus would always be in *koni*, which is more specific information, in both Type-subtype and Subtype-type MACs.

In terms of phonetic manifestations of the prosodic marking of identificational focus, we expect an overall higher F0 with a higher peak and longer duration as in Jun & Kim (2007). If the information structure between the first NP (NP1) and the second NP (NP2) is not manifested in speech production, prosodic strengthening will always exist in the VP-initial position (i.e., the left edge of NP1). However, if the information structure between two NPs is realized, strengthening is observed from the exact locus of identificational focus, which is NP2 for Type-subtype MACs and NP1 for Subtype-type MACs. Regarding the boundary condition, the constituents with identificational focus are more likely to be realized at IP-initial positions based on previous research results (Jun & Kim 2007). Nonetheless, previous research also demonstrated that the focused constituents do not necessarily have to initiate a new IP (Jun & Lee 1998; Choi et al. 2020). Therefore, the current study will revisit and examine the boundary realization of the identificational-focused constituents.

The production of identificational focus will be tested in a default reading following the Implicit Prosody Hypothesis (Fodor 1998; 2002) and Jun’s (2003) methodology. Fodor argues that “a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution” (2002: 113) in default reading (also known as silent reading). If all other factors are equal, the speakers will prefer the syntactic analysis consistent with the most natural prosodic contour. Therefore, we hypothesize that native speakers’ preferences for focus and boundary realization are expected to be observed in a default reading, which reflects the information structure of NPs in MACs.

The second question is concerned with the prominence marking induced by contrastive focus in Seoul Korean. Given that Korean is categorized as an edge-prominence language in which prominence marking is associated with the prosodic phrasing (Jun 2006; 2014), it is expected that focused constituents will tend to be located at higher domain initial positions as in Jeon & Nolan (2017). The minimal boundary condition for the focused constraints is expected to be either the AP-initial position (Jun & Lee 1998) or the IP-initial position (Jun & Kim 2007). The focused constituents are expected to be realized with an overall higher F0 and elongated duration, as previous research reported (Jun & Kim 2007; Lee & Xu 2010; Cho et al. 2011; Choi et al. 2020). Furthermore, the presence of post-focus compression, such as reducing pitch range

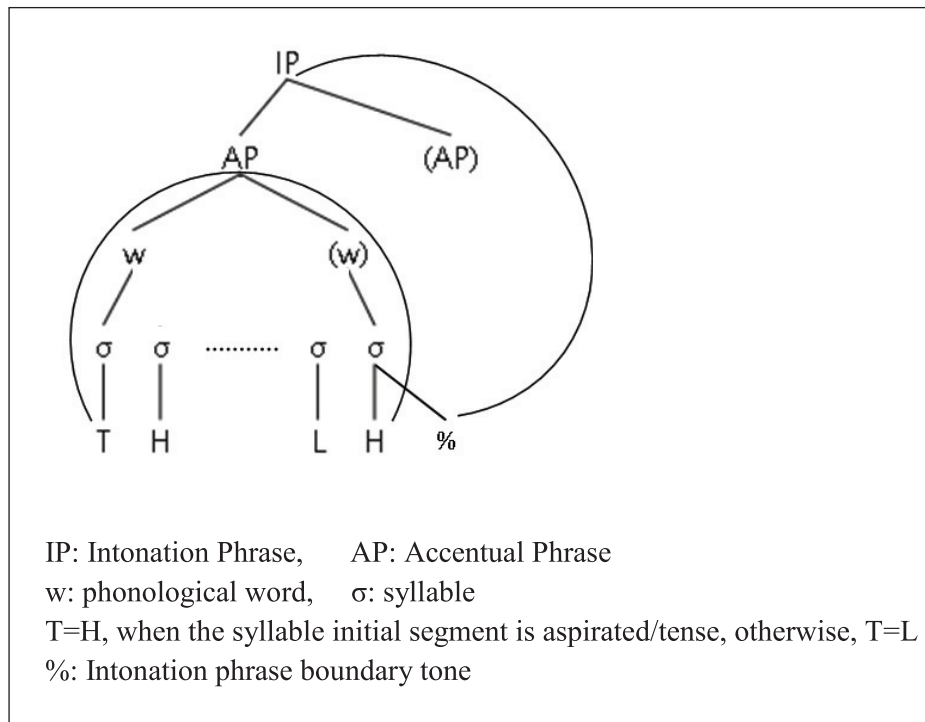
and amplitude, is expected based on the results of previous studies (Jun & Kim 2007; Lee & Xu 2010). To address this issue, we examine how NP1s and NP2s in MACs are phonetically realized in relation to the locus of contrastive focus. We then compare the present study results with the data of Seoul Korean from previous literature and discuss implications for prominence-induced prosodic strengthening in edge-prominence languages.

The third question is whether and how the different sources of focus interact with one another. A substantial amount of research in semantics and syntax has proposed that identificational focus and contrastive focus are independent notions (Vallduví & Vilkkuna 1998; Kenesei 2006; Neeleman & van de Koot 2008). Even though both identificational and contrastive focus are signaled by identical sets of pitch accents (e.g., H\*L) in intonational languages like English, the realizations of such pitch accents in actual speech are more prominent on constituents with contrastive focus (Bolinger 1961; Katz & Selkirk 2011). It is then plausible that the prosodic realization of contrastive focus is more prominent than that of identificational focus because of the facultative formal distinction between the two focus types (Zimmermann & Onea 2011). Therefore, we hypothesize that more prominent focus effects of contrastive focus (i.e., elongated duration and heightened pitch) will be observed in Korean.

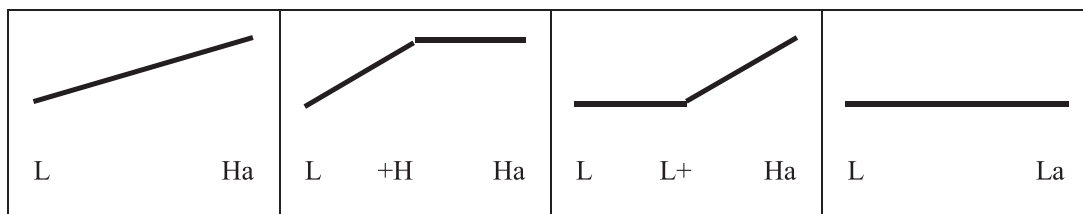
In sum, the current study investigates the phonetic variation of prominence-driven prosodic strengthening as a function of different types of focus (identificational and contrastive focus) by means of two accusative NPs in Korean MACs.

## 2 Intonational structure of Seoul Korean

Before we move on to the current study design, it is worth discussing the intonational structure of Seoul Korean for additional background. The current version of K-ToBI (Korean Tones and Break Indices, version 3.1) (Jun 2000) proposed that the intonational structure of Seoul Korean comprises two prosodic units, Intonation Phrase (IP) and Accentual Phrase (AP). An IP can possess one or multiple APs, and an AP can possess one or multiple phonological words (w). An IP is signaled with a final lengthening and (phrasal-final) boundary tone (e.g., L%, H%, LH%, HL%, LHL%, HLH%, HLHL%, LHLH%, LHLHL%). An AP is signaled with AP-initial tones (i.e., L, H, and L+) and AP-final tones (i.e., La, Ha, and L+). In terms of the AP-initial tones, L and H are realized on the first syllable of an AP according to the quality of an AP-initial segment, and +H is optionally realized on the second syllable of an AP. When an AP-initial segment is an aspirated or tense stop, an H is realized on the first syllable; otherwise, L is realized on the first syllable. Regarding the AP final tones, La and Ha are tones for the AP-final syllable, and L+ is the optional tone for the penult syllable of an AP. **Figure 1** presents the schematical structure of Seoul Korean intonation. Moreover, **Figure 2** describes possible F0 contours of APs with an initial syllable of lenis velar stop /k/.



**Figure 1:** Intonational structure of Seoul Korean (*adapted from Jun 2000*).



**Figure 2:** Schematic F0 contours of tonal patterns of short AP with an initial segment of a velar lenis stop /k/ (*adapted from Jun 2000*).

With this background in Seoul Korean intonation research, we controlled the AP-initial segment in our study design, which influences AP-initial tones. While analyzing the production data, we also scrutinized the AP and IP boundaries, closely adhering to the findings established in Seoul Korean intonation research. The details of our study design and the analyses are presented in Section 3. Again, the primary objective of this study is to investigate how the phonetic realization of prosodic prominences is affected by different types of focus, namely identificational and contrastive focus, within the context of MACs. To achieve this goal, we examined how the semantic relations between two accusative NPs in MACs interact with the focus types arising from three discourse contexts: default reading, NP1-contrast, and NP2-contrast contexts.



## 3 Method

### 3.1 Participants

Sixteen native speakers of Seoul Korean (eight females and eight males) participated in the current study. All were university students living in Seoul. The average age of the participants was 24.69 years (s.d. = 3.11 years). They were born and raised in the metropolitan areas of Seoul, Incheon, or Gyeonggi Province. The length of participants' residence outside of the target dialect areas was controlled as less than two years. None of the speakers reported any hearing or speech problems.

### 3.2 Experimental materials and procedure

There were two target words of two-syllable sequences: /koki/'meat' and /koni/'duck.<sup>3</sup> The word-initial syllables of the target items were controlled as lenis velar stop (/k/) in the /o/ context (/ko/) to control possible segmental effects on Voice Onset Time (VOT) and F0 dimensions. The test words were included in the carrier sentences, creating Type-subtype and Subtype-type phrases as described in **Table 1**.<sup>4</sup>

(a)	Type-subtype				
	Mimi-ka	koki-ruul	koni-ruul	p*alli	mAk-As*-A.
	Mimi-NOM	meat-ACC	duck-ACC	quickly	eat-PST-DECL
	"Mimi quickly ate koki (meat), koni (duck)."				
(b)	Subtype-type				
	Mimi-ka	koni-ruul	koki-ruul	p*alli	mAk-As*-A.
	Mimi-NOM	duck-ACC	meat-ACC	quickly	eat-PST-DECL
	"Mimi quickly ate koni (duck), koki (meat)."				

**Table 1:** Target words within carrier sentences as a function of semantic relations.<sup>5</sup>

<sup>3</sup> The binominal name of /koni/ is *cygnus columbianus*. We use duck, which is a superspecies of /koni/, in the current study for the sake of convenience.

<sup>4</sup> The current study was conducted with only one type of speech materials, the bisyllabic words beginning with /ko/, to control the phonetic environments of the target items. As one reviewer pointed out, further studies are needed to investigate the effects with different types of speech materials.

<sup>5</sup> The MACs that we used as the stimuli are commonly used in colloquial speech and are judged as grammatical by native speakers. Their properties have been extensively discussed from diverse theoretical perspectives (see, among others, Chae & Kim 2008; Park 2013; Yoon 2015; Kim 2020; 2023).

When the participants arrived at the sound recording studio in Hanyang Institute for Phonetics and Cognitive Science of Language (HICPS), they completed a demographical questionnaire about their language and residential backgrounds. Afterward, the researcher presented a brief introduction to the experimental procedure. An interview was then performed with demographical questions from Q-GEN-II (Labov 1984). The purpose of the interview session was to get the participants used to being in the new laboratory environment and to adjust the recording settings (e.g., proper location of the microphone and recording levels) for each participant. Then, he checked whether the participants were familiar with the subtype target word /koni/. Most participants answered that they had heard of it, though two of them were not sure about its meaning.<sup>6</sup> For those participants, he explained the meaning in comparison to its hypernym, /koki/, and other hyponyms of /koki/, such as /twɛzi/ ‘pork’, /so/ ‘beef’, and /taLk/ ‘chicken’. The recording session was conducted in a sound-attenuated room with a SHURE KSN44 microphone and a Tascam Hd-P2 digital recorder at a sampling rate of 44.1 kHz.

Following Fodor’s (1998; 2002) IPH and the methodology of Jun (2003) and Cho et al. (2019), the participants produced the carrier sentences in three different discourse contexts as in **Table 2**. In default reading, participants were first asked to read the sentence three times silently while the target sentences were presented in Korean orthography on a monitor. After the third silent reading, the participants read the sentence aloud as naturally as possible without any additional direction. In contrastive focus contexts, the carrier sentences were located in mini-discourse situations yielding NP1-contrast and NP2-contrast. The participants listened to the prerecorded question of whether Mimi ate some food, which was expressed by a specific pair of complex noun phrases. Then, the participants were asked to answer the question based on the correct answer presented in Korean orthography on a screen. For example, if they hear the question, “*Did Mimi quickly eat ‘vegetable,’ duck?*”, then they will see the answer, “*No. Mimi quickly ate ‘meat,’ duck.*” presented on a computer screen and give a verbal answer to the question based on what they see. The auditory stimuli were prerecorded by two speakers of Seoul Korean born and raised in Seoul (a female and a male speaker). These prosodic contexts allowed us to compare the realizations of the default reading with two contrastive focused conditions.

---

<sup>6</sup> We analyzed the data to assess the impact of the two participants on our data. The results were consistent regardless of whether we included or excluded these participants. Therefore, we decided to retain the two participants in our final analysis.

(a)	Default reading
A:	Mimi-ka koki-lul koni-lul p*alli mAk-As*-A. Mimi-NOM meat-ACC duck-ACC quickly eat-PST-DECL “Mimi quickly ate koki (meat), koni (duck).”
(b)	NP1-contrast context
Q:	Mimi-ka <b>yats<sup>h</sup>ε-lul</b> koni-lul p*alli mAk-As*-ni? Mimi-NOM vegetable-ACC duck-ACC quickly eat-PST-QUE “Did Mimi quickly eat <b>yats<sup>h</sup>ε (vegetable)</b> , koni (duck)?”
A:	ani. Mimi-ka <b><u>koki-lul</u></b> koni-lul p*alli mAk-As*-A. no. Mimi-NOM meat-ACC duck-ACC quickly eat-PST-DECL “No. Mimi quickly ate <b><u>koki (meat)</u></b> , koni (duck).”
(c)	NP2-contrast context
Q:	Mimi-ka koki-lul <b>twεzi-lul</b> p*alli mAk-As*-ni? Mimi-NOM meat-ACC pork-ACC quickly eat-PST-QUE “Did Mimi quickly eat koki (meat), <b>twεzi (pork)</b> ?”
A:	ani. Mimi-ka koki-lul <b><u>koni-lul</u></b> p*alli mAk-As*-A. no. Mimi-NOM meat-ACC duck-ACC quickly eat-PST-DECL “No. Mimi quickly ate koki (meat), <b><u>koni (duck)</u></b> .”

**Table 2:** Example sentences with Type-subtype phrases as a function of discourse contexts.

In the first block, the default reading was given to the participants in random order. The second block consisted of random-ordered sentences of NP1-contrast and NP2-contrast contexts. Each block was repeated three times with randomized orders of the experiment sentences. To prevent any possible learning effects on the default reading, none of the second block recordings preceded those of the first block. In total, 288 tokens (2 semantic relations  $\times$  3 prosodic conditions  $\times$  3 repetitions  $\times$  16 participants) were recorded as well as 54 filler sentences containing other types of MACs.<sup>7</sup> One of the authors (a trained Korean ToBI transcriber) reviewed the prosodic realizations of the target sentences in NP1-contrast and NP2-contrast conditions (i.e., the realizations of intended

<sup>7</sup> One reviewer pointed out that the lack of non-MAC fillers in the experimental design may be a possible limitation of our study. We chose to use only MAC fillers to reduce the duration of the production task and avoid fatigue effects on the participants. Future research could explore how different types of filler affect the prosodic realizations of MACs.

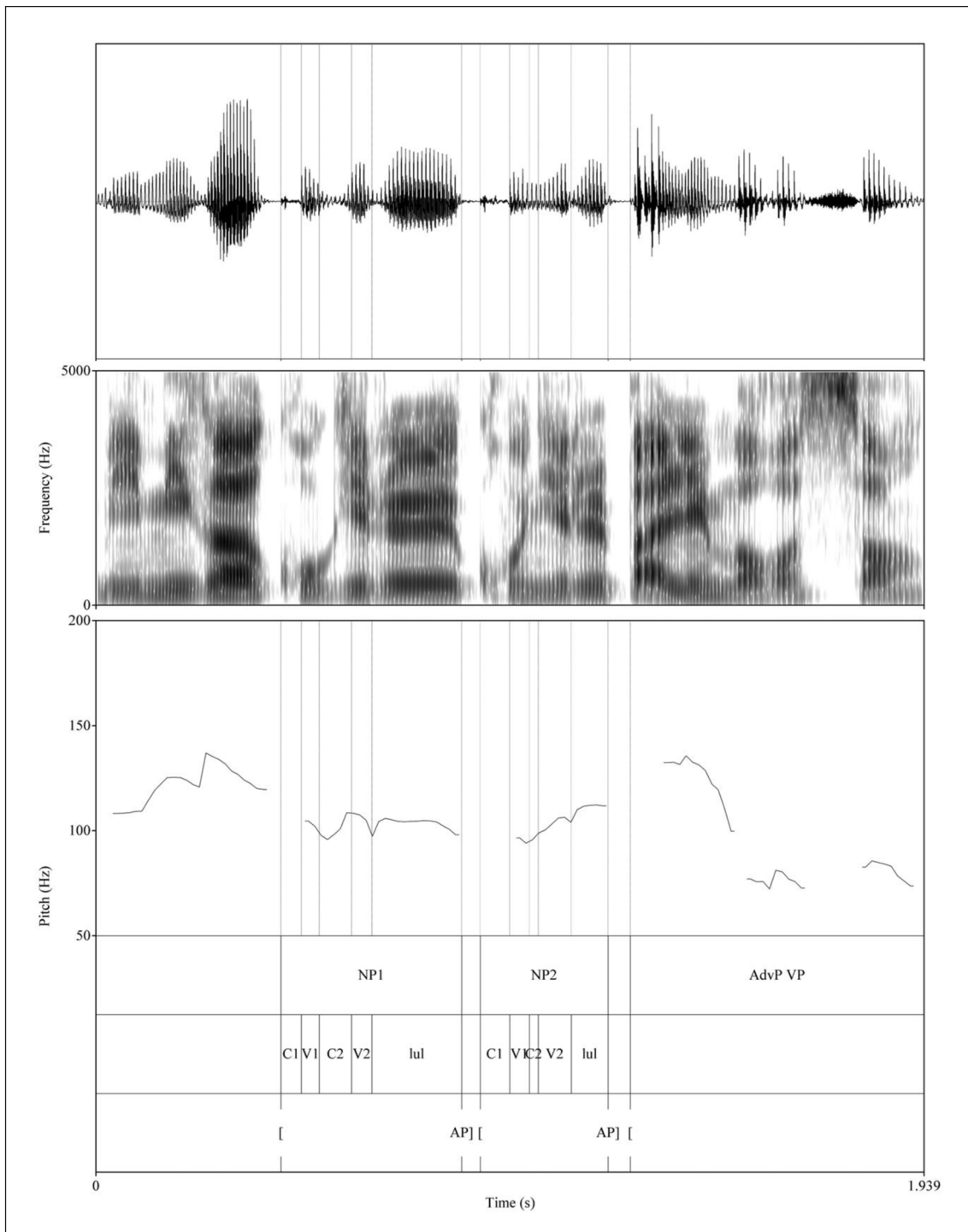
prominence marking) given that NP1-contrast and NP2-contrast conditions were intended to be produced with a focus on the targeted NPs. He checked the realization of the focus in terms of the acoustic characteristics of the focused noun phrases in relation to durational and tonal patterns following the criteria suggested by previous literature (Jun 2000; Xu & Xu 2005; Cho et al. 2011; Choi et al. 2020). As a result, we discarded 37 tokens (12.85%) because they had insufficient or additional realizations of focus in the sentences; 16 tokens were from Subtype-type MACs sentences and 21 from Type-subtype MACs ones. Results from a t-test show that the percentage of discarded tokens was similar in these two MAC sentence types ( $t(1) > 0.01, p = .998, 95\% \text{ CI: } -22.05, 22.04$ ), suggesting that sentence type was not likely to influence our decision.

### 3.3 Data annotation and acoustic measurements

The recording was first labeled in terms of phonemic and phrasal duration in Praat (Boersma & Weenink 2016), as shown in **Figure 3**. The duration of the first consonant and vowel in target nouns was measured. In  $C_1V_1C_2V_2$  sequences,  $C_1$  duration was measured from the onset of stop release to the onset of the second formant (F2) seen in spectrograms, given that  $C_1$  was controlled as /k/. The vowel portions were set from the onset to the offset of F2. Next, the duration of the noun phrases, the first NP (i.e., NP1) and the second NP (i.e., NP2) of the sentences, was measured. For example, when a target sentence was *Mimi-ka koki-lul koni-lul p\*alli mal-k-as\*-a* ‘Mimi quickly ate meat, duck’, *koki-lul* was an NP1, *koni-lul* was an NP2, and *p\*alli* was an adverb phrase (AdvP). The NP1 and NP2 durations were from the onset of stop release to the offset of the voice bar of the syllable-final /l/. Furthermore, the strength of the prosodic boundary at the left edges of the NPs (i.e., NP1, NP2, and AdvP) was labeled as either AP or IP, following Jun (2000). This labeling was mostly based on the cues such as phrasal-final tones, phrase-final lengthening given that an IP is marked by a boundary tone and final lengthening, while an AP is marked by a phrasal tone but not by final lengthening. When there were no clear acoustical manifestations, the Korean ToBI transcriber’s sense of juncture was the last resort for the labeling. (3) illustrates the labeling of the target constituents in the Type-subtype MAC sentence.

- (3) Mimi-ka [<sub>NP1</sub> koki-lul] [<sub>NP2</sub> koni-lul] [<sub>AdvP</sub> p\*alli] mal-k-as\*-a.  
 Mimi-NOM meat-ACC duck-ACC quickly eat-PST-DECL  
 ‘Mimi quickly ate meat, duck.’

F0 curves of the entire speech samples were obtained by Voicesauce (Shue 2010; Shue et al. 2011) with the STRAIGHT algorithm (Kawahara et al. 1998). The raw values of F0 in Hz were converted into semitones by applying the formula  $12[\log_2((F0 \text{ in Hz})/100)]$ . F0 contours of NP1 and NP2 were extracted from the entire contours of the experimental sentences. Furthermore, the F0 values of the  $V_1$  midpoint were obtained from the entirety of the F0 data.



**Figure 3:** An example waveform, spectrogram, pitch contour, and annotations for *Mimi-ka koki-lul koni-lul p\*alli mak-as\*-l* 'Mimi quickly ate meat, duck'.

### 3.4 Statistical analysis

We analyzed data using R (R Core Team 2020). For the strength of boundary marking, zero-inflated poisson log-linear models were fitted with `pscl` package (Zeileis et al. 2008; Jackman 2020). Also, a series of mixed-effects models with `lme4` package (Bates et al. 2013; 2015) was performed to test the effects of Semantic relations (with two levels: Subtype-type = -1, Type-subtype = 1), NP locations (with two levels: NP1 = -1, NP2 = 1), and Prosody (with three levels: Default = [-1, -1], NP1-contrast = [1, 0], NP2-contrast = [0, 1]) on the acoustic measurements (C1 duration and F0 values at the V1-mid points). The independent variables in mixed-effects modeling were coded with sum coding to compare each level with the grand mean. In addition, a by-participant random intercept was included to account for individual differences and repeated measurements (see Table 6 in Appendix for the full syntax for the statistical analysis). Furthermore, we tested the effect of Semantic Relations on each acoustic measurement in each condition separately, so that we could explore possible interactions among three dependent variables (i.e., Semantic Relations, NP Locations, and Prosody). For example, the effects of Semantic Relations on C1 duration were compared regarding their relations with NP Locations and Prosody (e.g., Semantic Relations effects on C1 duration on NP1 vs. Semantic Relations effects on C1 duration on NP2). The models were compared in a stepwise manner using the backward elimination method with the alpha level set at 0.05. Finally, the F0 curves of NPs (i.e., NP1s and NP2s) were examined with Smoothing Spline Analysis of Variance (SSANOVA) using `gss` (Gu 2014) package. SSANOVA is a statistical method for analyzing data varying in a complex series of time that has been widely used in research with articulatory data (Davidson 2006; Lee-Kim et al. 2013; Bennett et al. 2018) and F0 (Derrick & Schultz 2013; Yiu 2015). In order to take possible interactions into account, the SSANOVA models were specific to each Prosody condition reflecting a comparison of NP Locations and Semantic Relations. This analysis allowed each semantic relation to have a separate offset across the normalized time. The best-fit curves from the analysis represented the F0 contours of each condition across speakers and repetitions. In our analysis, the shaded bands indicated 95% of the Confidential Interval of each offset. When the confidence interval of the two contours did not overlap, we interpreted that the difference between the two F0 contours was significant. Therefore, tonal realizations of the noun phrases were compared regarding the types and locus of focus.

## 4 Results

### 4.1 The strength of the boundary

We performed a series of descriptive and inferential statistical analyses to test focus effects on prosodic phrasing (i.e., whether an NP-initial boundary was produced as an AP, or IP-initial

position)<sup>8</sup> in relation to information structure (i.e., identificational focus) and discourse contexts (i.e., default vs. contrastive focus). We fit the saturated log-linear model to the three-way contingency tables. Based on the frequency distribution of the outcome, the models were fit under a Poisson assumption in which sample zeroes are permissible (Agresti 2003). The saturated model included fixed effects of Boundaries (AP vs. IP), Semantic Relations (Type-subtype vs. Subtype-type), Prosody (Default, NP1-contrast, NP2-contrast), and their interactions.

#### 4.1.1 NP1

Regarding NP1, the log-linear model analysis was not performed because the boundaries at NP1-initial positions were always realized as APs. The three-way contingency table of boundary realization of NP1-initial positions is presented at the top of **Table 3**.

#### 4.1.2 NP2

In terms of NP2, a series of log-linear model comparisons was conducted to test whether the boundary marking at the left edge of NP2 was different based on the effects of Semantic Relations and Prosody as well as possible interactions among variables. The frequencies of boundary realizations were summarized in the middle of **Table 3**, and the log-linear modeling processes were described in **Table 7** at Appendix. The most parsimonious model included the main effects of Boundary, Prosody, Semantic Relations (see Table 6 in Appendix for the description of the parsimonious model). The estimated parameters for the most parsimonious log-linear model are presented in **Table 8** at Appendix. With regard to the count model (See **Table 7a**), the only statistically significant effect was Boundary, which indicated that the odds ratio of having an IP boundary at the left edge of NP2 was 0.18 times that of the AP boundary ( $z = -6, p < .001$ ). Furthermore, none of the parameters were statistically significant when we considered the zero-inflation model (See **Table 7b**). The comparison of AIC and BIC values showed that the zero-inflation model (AIC: 60.94, BIC: 65.79) outperformed the count model (AIC: 81.33, BIC: 83.76). Therefore, we can conclude that most of the boundaries at the left edge of NP2 were realized as AP boundaries, and these tendencies in the data were regardless of the effects of Semantic Relations and Prosody.

---

<sup>8</sup> We excluded the possibility of prosodic word (W) since there was no observation of constituents initiating a prosodic word. Intermediate Phrases (ip) were not considered in this paper because the current model of K-ToBI (Jun 2000) does not include it.

### 4.1.3 AdvP

With regard to AdvP, the boundary at the left edge was always realized as an AP, as summarized at the bottom of **Table 3**. Because there was no other boundary marking, an additional statistical analysis was not required.

Location	Semantic relations	Prosody	Boundary realization		
			AP	IP	Total
NP1-initial	Subtype-type	Default	48	0	48
		NP1-contrast	40	0	40
		NP2-contrast	39	0	39
		Total	127	0	127
	Type-subtype	Default	48	0	48
		NP1-contrast	38	0	38
		NP2-contrast	38	0	38
		Total	124	0	124
NP2-initial	Subtype-type	Default	40	8	48
		NP1-contrast	40	0	40
		NP2-contrast	39	0	39
		Total	119	8	127
	Type-subtype	Default	41	7	48
		NP1-contrast	38	0	38
		NP2-contrast	38	0	38
		Total	117	7	124
AdvP-initial	Subtype-type	Default	48	0	48
		NP1-contrast	40	0	40
		NP2-contrast	39	0	39
		Total	127	0	127
	Type-subtype	Default	48	0	48
		NP1-contrast	38	0	38
		NP2-contrast	38	0	38
		Total	124	0	124

**Table 3:** Three-way contingency tables of boundary realizations in NP1-initial, NP2-initial, and AdvP-initial positions.



## 4.2 C1 duration and F0 in V1

To test focus effects on phonetic manifestations in relation to information structure (i.e., identificational focus) and discourse contexts (i.e., default vs. contrastive focus), the linear mixed-effects model was performed. Following the suggestion of Pedhazur (1997) that deviation coding is advantageous in interpreting interaction effects, the predictor variables were coded using deviation coding: Semantic Relations (0: Subtype-type, 1: Type-subtype), NP Locations (0: NP1, 1: NP2), and Prosody (0: Default, 1: NP1-contrast, 2: NP2-contrast).

### 4.2.1 C1 duration (VOT)

The saturated model was not statistically different from a model eliminating the three-way interaction term ( $\chi^2(2) = 1.58, p = .453$ ), so the three-way interaction term was excluded from the model. The model comparisons removing each interaction revealed that the interactions between Semantic Relations and NP Locations ( $\chi^2(1) = 23.17, p < .001$ ) and between NP Locations and Prosody ( $\chi^2(2) = 80.87, p < .001$ ) were significant. However, removing the interaction effect between Semantic Relations and Prosody did not show significant statistical results ( $\chi^2(2) = 3.13, p = .209$ ). All in all, the best-fitted model included the two-way interaction terms of Semantic Relations  $\times$  NP Locations and NP Locations  $\times$  Prosody, as well as their main effects (see Table 6 in Appendix for the description of the best-fitted model). The summary of the mixed-effects modeling and the best-fitted model is presented in Table 4 and Table 5.

Models	$df^a$	AIC	BIC	logLik	deviance	$\chi^2$	$\chi^2 df$	$p$
<b>The saturated model</b>	14	4223	4282	-2098	4195			
<b>The three-way interaction term</b>								
– Boundary:Semantic Relations:Prosody	12	4221	4271	-2098	4197	1.58	2	.453
<b>The two-way interaction term</b>								
– Semantic Relations:NP Locations	11	4242	4288	-2110	4220	23.17	1	<.001
– NP Locations:Prosody	10	4297	4340	-2139	4277	80.87	2	<.001
– Semantic Relations:Prosody	10	4220	4262	-2100	4200	3.13	2	.209

**Table 4:** Summary of linear regression model analysis results on C1 duration.  
a. This indicates the degrees of freedom between models.

Parameter estimate	Estimate	Std. error	df	<i>t</i>	<i>p</i>
Intercept	66.69	1.65	14.88	40.31	<.001
Semantic Relations[1] <sup>a, b</sup>	0.82	0.68	481.65	1.2	.23
NP Locations[1] <sup>c</sup>	-2.03	0.68	480.95	-2.96	.003
Prosody[1] <sup>d</sup>	-5.59	0.94	482.22	-5.95	<.001
Prosody[2] <sup>e</sup>	0.77	0.99	483.64	0.78	.435
Semantic Relations[1]:NP Locations[1]	3.28	0.68	480.97	4.82	<.001
NP Locations[1]:Prosody[1]	0.64	0.94	481	0.68	.497
NP Locations[1] :Prosody[2]	7.72	0.98	481.11	7.86	<.001

**Table 5:** Estimated parameters and related statistics for most parsimonious linear regression model for C1 duration.

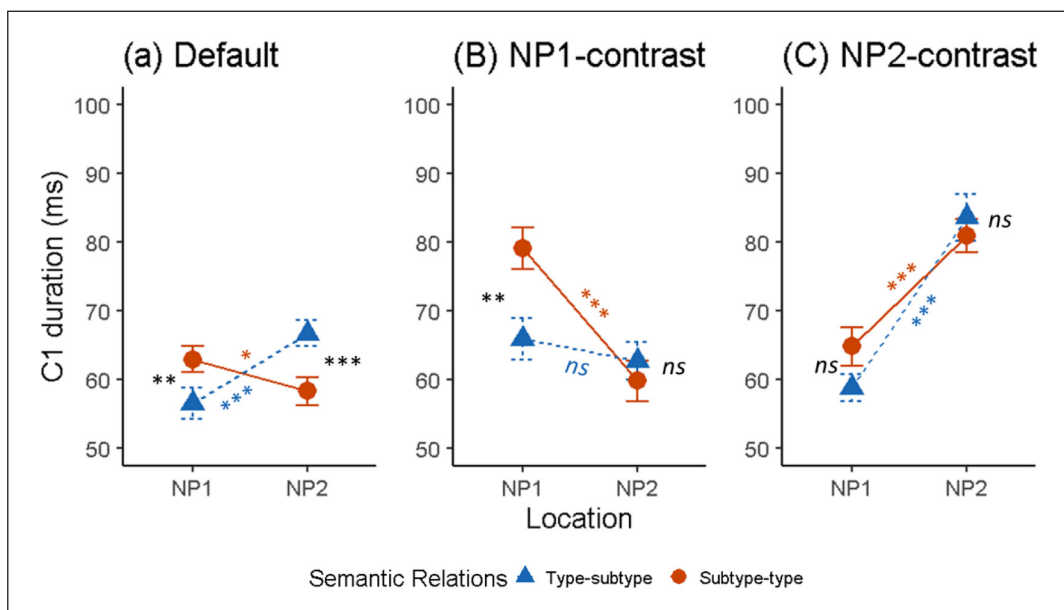
- a. The 1 represents the first level of the factor.
- b. The reference level of Semantic Relations was Subtype-type.
- c. The reference level of NP Locations was NP1.
- d. The reference level of Prosody was Default.
- e. The second level of Prosody was NP1-contrast.

Although the three-way interaction term was not statistically significant, it is noteworthy that the two-way interaction between Semantic Relations and NP Locations seems to be further affected by the effect of Prosody. **Figure 4a** shows the mean C1 duration in default reading, separated by Semantic Relations and NP locations. It was hypothesized that the identificational-focused constituents show longer C1 duration. As hypothesized, C1 duration under identificational focus was always longer than that in its counterpart. Analysis with mixed-effects models also confirmed this observation. C1 duration of Subtype-type NP1s was significantly longer than that of Type-subtype NP1s ( $\beta = 3.32$ ,  $SE = 1.07$ ,  $\chi^2(1) = 9.18$ ,  $p = .002$ ). In the meantime, C1 duration of Subtype-type NP2s was significantly shorter than that of Type-subtype NP2s ( $\beta = -4.15$ ,  $SE = 1.21$ ,  $\chi^2(1) = 11.19$ ,  $p < .001$ ). Regarding the difference between NP1s and NP2s, both Subtype-type ( $\beta = 2.31$ ,  $SE = 1.14$ ,  $\chi^2(1) = 4.06$ ,  $p = .044$ ) and Type-subtype ( $\beta = -5.22$ ,  $SE = 1.27$ ,  $\chi^2(1) = 15.4$ ,  $p < .001$ ) showed statistically significant results, but with opposite directionality.

In NP1-contrast contexts, we predicted that lengthening driven by contrastive focus would be more pronounced than lengthening driven by identificational focus. **Figure 4b** showed a substantial amount of lengthening in NP1s and durational compression in NP2s. C1 duration of both Subtype-type NP1s ( $\beta = 8.28$ ,  $SE = 1.56$ ,  $\chi^2(1) = 23.59$ ,  $p < .001$ ) and Type-subtype NP1s ( $\beta = 4.88$ ,  $SE = 1.76$ ,  $\chi^2(1) = 7.18$ ,  $p = .007$ ) was significantly longer than the counterparts in Default condition. The difference between Type-subtype and Subtype-type was significant in

NP1s ( $\beta = 6.92$ ,  $SE = 1.78$ ,  $\chi^2(1) = 13.5$ ,  $p < .001$ ) whereas, C1 duration of Subtype-type NP2s was not significantly different from that of Type-subtype NP2s ( $\beta = -0.87$ ,  $SE = 1.74$ ,  $\chi^2(1) = 0.26$ ,  $p = .607$ ). With regard to the difference between NP1s and NP2s, C1 duration of NP1s was significantly longer than that of NP2s in Subtype-type relations ( $\beta = 9.78$ ,  $SE = 1.8$ ,  $\chi^2(1) = 24.6$ ,  $p < .001$ ), but not in Type-subtype relations ( $\beta = 1.62$ ,  $SE = 1.71$ ,  $\chi^2(1) = 0.9$ ,  $p = .344$ ).

In NP2-contrast contexts, we hypothesized that contrastive-focused constituents would show more robust lengthening than their corresponding constituents. **Figure 4c** demonstrated that the C1 duration under contrastive focus was always longer than that in NP1s, as hypothesized. C1 duration of Subtype-type NP1s was not significantly different from that of Type-subtype NP1s ( $\beta = 2.96$ ,  $SE = 1.52$ ,  $\chi^2(1) = 3.73$ ,  $p = .054$ ). Similarly, C1 duration of Subtype-type NP2s was not significantly different from that of Type-subtype NP2s ( $\beta = -1.7$ ,  $SE = 1.37$ ,  $\chi^2(1) = 1.49$ ,  $p = .222$ ). With regard to the difference between NP1s and NP2s, C1 duration of NP1s was significantly shorter than that of NP2s both in Subtype-type relations ( $\beta = -8.33$ ,  $SE = 1.52$ ,  $\chi^2(1) = 24.58$ ,  $p < .001$ ) and in Type-subtype relations ( $\beta = -12.4$ ,  $SE = 1.63$ ,  $\chi^2(1) = 38.58$ ,  $p < .001$ ).



**Figure 4:** C1 duration in relation to Semantic Relations and NP Locations in (a) Default reading, (b) NP1-contrast, and (c) NP2-contrast contexts. Error bars represent standard errors.

#### 4.2.2 F0 at the midpoint of V1

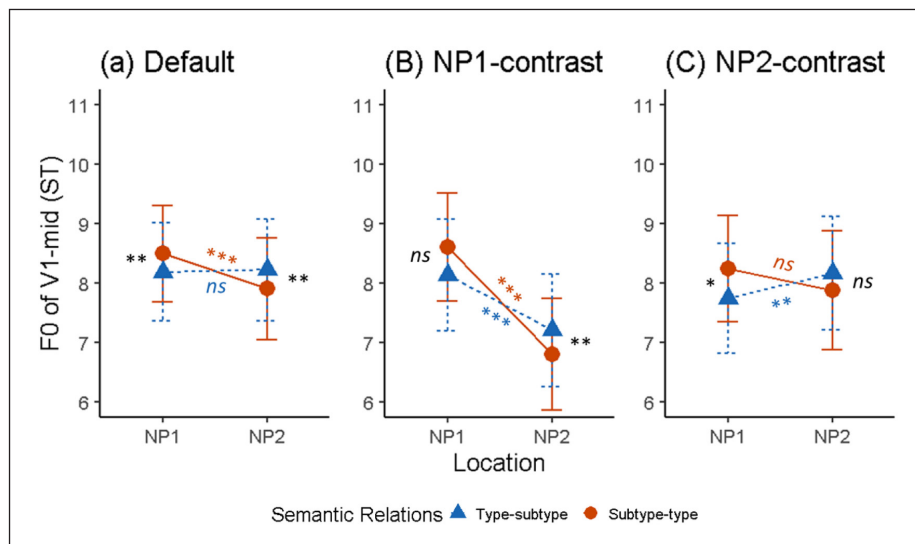
Regarding the best-fitted model of F0 of V1-mid, the full model was not statistically different from a model eliminating the three-way interaction term ( $\chi^2(2) = 0.49$ ,  $p = .782$ ); thus, the

three-way interaction term was excluded from the model. The interaction between Semantic Relations and NP Locations was significant ( $\beta = 0.14$ ,  $SE = 0.03$ ,  $\chi^2(1) = 18.53$ ,  $p < .001$ ). A series of separate comparisons revealed that the effects of Semantic Relations were significant in both NP1 ( $\beta = 0.14$ ,  $SE = 0.04$ ,  $\chi^2(1) = 9.4$ ,  $p = .002$ ) and NP2 ( $\beta = -0.16$ ,  $SE = 0.06$ ,  $\chi^2(1) = 8.11$ ,  $p = .004$ ), with the opposite directionality. On the one hand, the F0 of V1 in the Subtype-type NP1 (8.57 ST) was significantly higher than that in the Type-subtype NP1 (8.3 ST). On the other hand, the F0 of V1 in the Subtype-type NP2 (7.81 ST) was significantly lower than in the Type-subtype NP2 (8.14 ST). Furthermore, the interaction effect was significant between NP Locations and Prosody ( $\chi^2(2) = 84.81$ ,  $p < .001$ ). With regard to the interaction between Semantic Relations and Prosody, removing the interaction effect did not reach statistically significant results ( $\chi^2(2) = 2.24$ ,  $p = .327$ ). All in all, the best-fitted model included the interaction terms of Semantic Relations  $\times$  NP Locations and NP Locations  $\times$  Prosody and their main effects (see Table 6 in Appendix for the description of the best-fitted model). Table 9 and Table 10 in Appendix summarize the mixed-effects modeling and the best-fitted model.

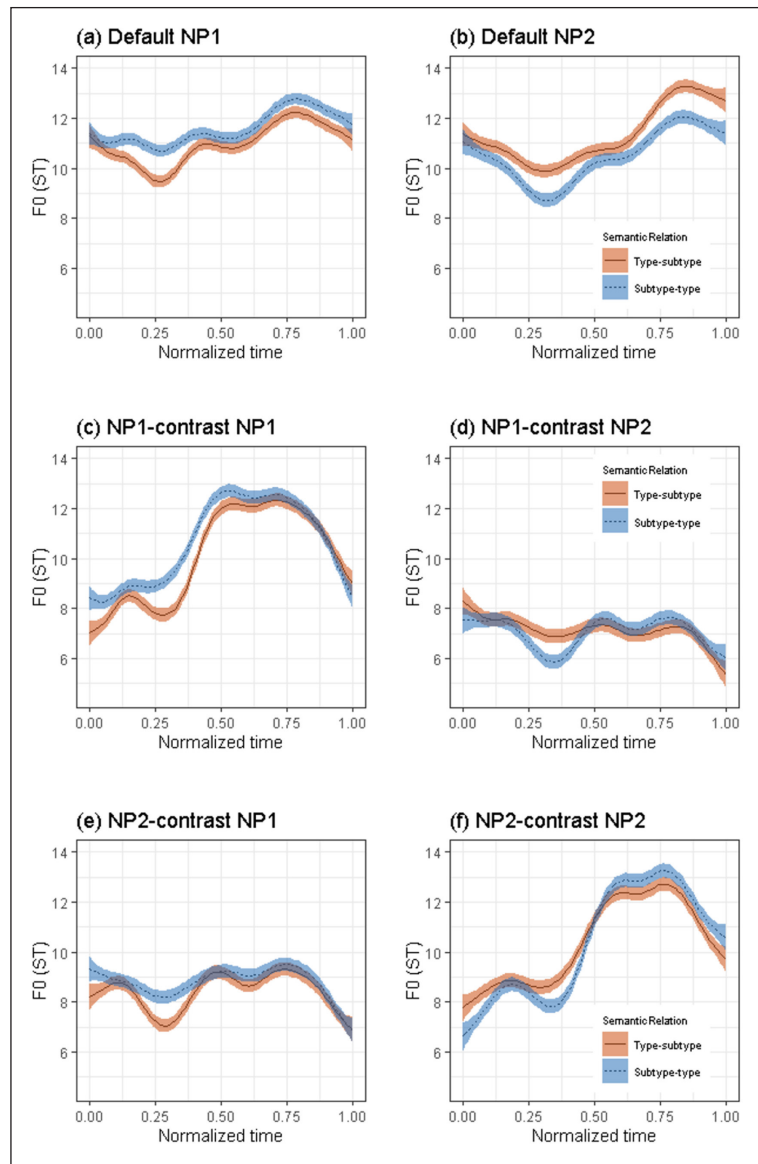
Notwithstanding that the three-way interaction term did not show a significant result, it is worth noting that the two-way interaction between Semantic Relations and NP Locations seems to further interact with the effect of Prosody. In default reading, it was hypothesized that F0 of V1 under identificational focus would be higher than that in the corresponding condition. **Figure 5a** showed that speakers pronounced V1 with a higher F0 if the words were under identificational focus. In the meantime, F0 of V1 in Subtype-type NP1s was significantly higher than that in Type-subtype NP1s ( $\beta = 0.15$ ,  $SE = 0.05$ ,  $\chi^2(1) = 10.32$ ,  $p = .001$ ), but, F0 of V1 in Subtype-type NP2s was significantly lower than that in Type-subtype NP2s ( $\beta = -0.16$ ,  $SE = 0.04$ ,  $\chi^2(1) = 10.64$ ,  $p = .001$ ). Regarding the difference between NP1s and NP2s, Subtype-type relations showed significant results ( $\beta = 0.29$ ,  $SE = 0.06$ ,  $\chi^2(1) = 18.54$ ,  $p < .001$ ), but Type-subtype relations did not ( $\beta = -0.02$ ,  $SE = 0.05$ ,  $\chi^2(1) = 0.14$ ,  $p = .712$ ).

**Figure 6a** and **Figure 6b** present the F0 contours of NP1s and NP2s in the Default reading. The time was normalized from 0 (the onset of NPs) to 1 (the offset of NPs) on the x-axis. The shaded polygons around the smoothing splines represent 95% Bayesian confidence intervals of the predictions. In NP1, the difference between the two semantic relations could be found throughout the entire portion of NP1s, and the difference between Type-subtype and Subtype-type was maximized in the first syllable portion. In Default NP2, the difference was more extensive than that in Default NP1 and maintained until the end of NP2s. The F0 contours of NP1s throughout two semantic relations showed the realizations of an L tone at the NP-initial position and Ha tone at the NP-final position. Regarding the NP-initial tone of NPs, L + L seemed to be realized, yielding LLHa.

In NP1-contrast contexts, it was expected that higher F0 would be realized in V1 under contrastive focus compared with identificational focus or in the defocused condition. **Figure 5b** demonstrates that F0 of V1 with contrastive focus was always higher than that in NP2s. F0 in NP1s was significantly higher than that in NP2s both in Subtype-type relations ( $\beta = 0.84$ ,  $SE = 0.07$ ,  $\chi^2(1) = 79.86$ ,  $p < .001$ ) and in Type-subtype relations ( $\beta = 0.53$ ,  $SE = 0.07$ ,  $\chi^2(1) = 41.44$ ,  $p < .001$ ). With regard to the identificational focus driven by the information structure of NP1s and NP2s, the phonetic realizations on the F0 dimension were not as robust as those in the default reading. F0 of V1 in Subtype-type NP1s was not significantly different from that in Type-Subtype NP1s ( $\beta = 0.08$ ,  $SE = 0.05$ ,  $\chi^2(1) = 2.52$ ,  $p = .113$ ). However, F0 of V1 in Subtype-type NP2s was significantly lower than that in Type-subtype NP2s ( $\beta = -0.22$ ,  $SE = 0.07$ ,  $\chi^2(1) = 9.61$ ,  $p = .002$ ). The SSANOVA analysis presented in **Figure 6c** and **Figure 6d** also indicated that the difference between the Type-Subtype and Subtype-type MACs was not considerable in NP1 or NP2. In contrast, the difference between NP1s and NP2s was robust, as demonstrated by the F0 contours, which indicates the effects of contrastive focus. One of the distinct patterns of NP1-focus context is that the second syllables of the contrast-focused constituents were realized with H tone. That is, the AP-initial tone under contrastive focus was realized with L + H preceding the AP-final Ha tone, as presented in **Figure 6c**. In addition, the analysis revealed that the F0 contour of NP2s was compressed as LLa, demonstrated by the flattened contours. These flattened contours manifest evidence of post-focus compression (see the results of NP2 in **Figure 6d** in terms of F0 contours of NP1-contrast NP1 and NP2).



**Figure 5:** F0 of V1-mid in relation to Semantic Relations and NP Locations in (a) Default reading, (b) NP1-contrast, and (c) NP2-contrast contexts. Error bars represent standard errors.



**Figure 6:** An SSANOVA of F0 (ST) calculated separately for NP1 – NP2 sequences plotted by Semantic Relations in (a-b) Default reading, (c-d) NP1-contrast, and (e-f) NP2-contrast contexts. In (a), the first, second, and third syllables approximately correspond to the intervals between 0 and 0.25, 0.25 and 0.6, and 0.6 and 1, respectively. In (b), the first, second, and third syllables approximately correspond to the intervals between 0 and 0.3, 0.3 and 0.7, and 0.7 and 1, respectively. In (c), the first, second, and third syllables approximately correspond to the intervals between 0 and 0.3, 0.3 and 0.6, and 0.6 and 1, respectively. In (d), the first, second, and third syllables approximately correspond to the intervals between 0 and 0.3, 0.3 and 0.7, and 0.7 and 1, respectively. In (e), the interval from 0 to 0.3 is the first syllable interval, 0.3-0.6 is the second syllable interval, and 0.6-1 is the third syllable interval. In (f), the interval from 0 to 0.3 is the first syllable interval, 0.3-0.7 is the second syllable interval, and 0.7-1 is the third syllable interval.

In NP2-contrast contexts, the observation of higher F0 from the constituents with contrastive focus was hypothesized. **Figure 5c** showed that F0 of V1 under contrastive focus was higher than in NP1s when the semantic relation between NPs was Type-subtype. When such a relation was Subtype-type, the F0 of V1 in NP1s and NP2s was similar. F0 of V1 in Subtype-type NP1s was significantly higher than that in Type-subtype NP1s ( $\beta = 0.16$ ,  $SE = 0.08$ ,  $\chi^2(1) = 4.56$ ,  $p = .033$ ), while F0 of V1 in Subtype-type NP2s was not significantly different from that in Type-subtype NP2s ( $\beta = -0.09$ ,  $SE = 0.07$ ,  $\chi^2(1) = 1.72$ ,  $p = .19$ ). Concerning the difference between NP1s and NP2s, F0 in NP1s was not significantly different from that in NP2s in Subtype-type relations ( $\beta = 0.01$ ,  $SE = 0.09$ ,  $\chi^2(1) = 0.02$ ,  $p = .901$ ), but differed significantly in Type-subtype relations ( $\beta = -0.21$ ,  $SE = 0.08$ ,  $\chi^2(1) = 7.59$ ,  $p = .006$ ). The SSANOVA analysis results presented in **Figure 6e** and **Figure 6f** suggest that the difference between Subtype-type and Type-subtype MACs in NP2-contrast contexts was marginal, as represented in the contours and confidence intervals of two semantic relations that mostly overlapped in NP1 and NP2. Nevertheless, the difference in F0 contour between NP1s and NP2s was substantial. F0 contour under contrastive focus was markedly higher than that in the defocused NP1s. The defocused NP1s were realized with LLa, which is presented in **Figure 6e**. The flattened F0 contours of NP1s demonstrate that pre-focus compression existed in the NP2-contrast context.

### 4.3 Results Summary

This paper investigated how the prosodic realizations of MACs in Seoul Korean are affected by identificational and contrastive focus. We conducted a production experiment and measured the strength of boundaries, C1 duration, F0 of V1, and F0 contours of nouns in MACs under different focus conditions. Our results revealed that identificational focus, which was triggered by semantic factors, lengthened C1 duration and raised the F0 of V1 when they were in the focused words. Contrastive focus also marked the focused words with longer C1 duration and higher F0 of V1, as well as caused pre- and post-focus compression. The results coming from the SSANOVA analysis further confirmed these patterns. However, neither identificational nor contrastive focus showed any effects on the strength of prosodic boundaries. A more intriguing finding was observed in the interaction of two types of focus. We discovered that contrastive focus overrides identificational focus in terms of C1 duration. Regarding F0 contours, we found distinct AP-initial tones for two focus conditions: L for identificational focus and L + H for contrastive focus.

## 5 Discussion

### 5.1 Prosodic marking of identificational focus

The results showed identificational focus effects induced by semantic factors most clearly in default reading. Even though Seoul Korean lacks culminative prominence marking such as the

pitch accent of head-prominence languages, the identificational focus was prosodically marked at the left edges of the boundary of focused constituents. The present results demonstrate that identificational focus induced longer duration and higher F0, as reflected in lengthened VOT of the phrase-initial lenis stop /k/ and heightened F0 in Subtype NPs (i.e., *koni-lul* ‘duck-ACC’). This extends the previous findings of identificational focus effects in VP-initial positions (Jun et al. 2006; Jun & Kim 2007) by showing such effects even in VP-medial positions. Although the VP-initial position is the preferred locus of VP focus because of the tendency of Seoul Korean to show left-headedness (Jun 2005), the realization of identificational focus was indeed sensitive to the information structure established by the taxonomic hierarchy between two NPs in MACs.

As to the prosodic phrasing, the identificational-focused constituents always initiate a new AP, at the least. This is in line with the results of Jun et al. (2006) in which all of their participants in the default reading condition produced each word phrased in a separate AP. This was ascribed to the fact that all the information in the target sentences was new to the readers in the default reading condition. In the current study, although NP1s always initiated a new AP, there existed 15 instances (out of 96 tokens in total) in which speakers initiated a new IP at the left edge of NP2s. Nonetheless, the statistical analysis results confirmed that the probability of prosodic marking with IP boundaries was significantly lower than (precisely 0.18 times of) that of prosodic marking with AP boundaries. Moreover, phrasing patterns of NP2-initial positions were not associated with the semantic relations between NP1s and NP2s. Therefore, we can conclude that most tokens initiated a new AP at the left edge of NP2s regardless of their semantic relations with NP1s. This may be ascribed to the fact that NP1s and NP2s refer to an identical referent. If the boundaries between NP1s and NP2s were larger than an AP boundary, listeners might find it difficult to discern that NP1s and NP2s are signaling the identical referent. Therefore, locating an IP boundary between NP1s and NP2s may be detrimental for decoding the intended relations of NPs. In sum, these results suggest that in most cases, AP boundaries were realized at the left edge of NP1s and NP2s, and these patterns were consistent regardless of the information structure of NPs (i.e., Type-subtype vs. Subtype-type MACs). Therefore, the hypothesis that speakers will initiate a new AP for the constituents with identificational focus was attested by the results of the current study in accordance with the results of Jun & Lee (1998).

## 5.2 Prosodic marking of contrastive focus

Another key finding of the current work was the prosodic marking of contrastive focus on durational and F0 dimensions. The VOT of phrasal-initial lenis velar stops was lengthened in contrastive-focused NPs compared with that in unfocused NPs. Moreover, F0 of V1s was higher when prominence driven by contrastive focus was marked on NPs. These correspond to the results of the previous studies (Lee & Xu 2010; Cho et al. 2011; Choi et al. 2020), showing longer duration and higher F0 of the focused constituents in Seoul Korean.



With regard to the prosodic phrasing, all of the NPs in the NP1- and NP2-contrast conditions initiated a new AP. These results support the hypothesis that speakers will initiate a new AP at the left edge of the focused constituents. However, the results did not attest the hypothesis that speakers will tend to locate a higher prosodic boundary at the left edge of the focused words. This is because the prosodic phrasing in contrastive-focus and identificational-focus contexts was not disparate from that in the defocused contexts. Therefore, the current study results suggest that neither identificational focus nor contrastive focus necessarily requires initiating a new IP for the focused accusative constituents in Seoul Korean MACs.

One crucial finding of the current study is that F0 contours of NPs in both pre- and post-focus positions showed compression of pitch range and pitch amplitude in the contrastive focus conditions. The contrastive focus induced a lowering and flattening of the entire F0 contours of pre-focus and post-focus NPs. Our findings agree with the results of Oh (2001) and Jeon & Nolan (2017) in that both pre-focus and post-focus F0 compression were observed in Seoul Korean. In addition, the compression patterns of the current study are overall consistent with the results of Jun & Lee (1998) and Yang et al. (2015), which reported general tendencies of compressions in pre- and post-focus constituents. Nonetheless, the results of the current study were different from those of Lee & Xu (2010) where they did not observe pre-focus compression in Seoul Korean. This may be due to the difference in experimental tasks. The task used by Lee & Xu (2010) was answering a wh-question, which is categorized as informational narrow focus based on the previous literature (Gussenhoven 1983; Féry 2013). According to Féry's scale of focus strength, informational narrow focus is located higher than identificational focus and lower than contrastive focus. Therefore, the discrepant patterns of pre-focus and post-focus compression found in Lee & Xu (2010) would be understood in relation to the scale of focus strength, suggesting further research comparing such focus compression effects in regard to diverse types of focus.

### 5.3 Interaction between identificational and contrastive focus

The results showed interaction effects between two focus types, identificational and contrastive focus. In the default reading, we found that C1 duration was longer, and F0 at the V1-mid was higher when NPs were realized with identificational focus. This trend was further confirmed by the F0 contours displayed in **Figure 5b**, which showed generally enhanced F0 contours for NPs in positions of identificational focus. However, our results suggested that the effects of identificational focus are influenced by the presence of contrastive focus. Specifically, when the target items were placed in the NP1-contrast condition, identificational focus effects on NP1s in terms of F0 dimensions (F0 at V1-mid and F0 contours) were not statistically significant. This pattern was consistent with the identificational focus effects in the NP2-contrast condition, where C1 duration and F0 measurements induced by identificational focus did not show statistical significance in NP2 positions. The absence of identificational focus effects in contrastive-focused

NPs can be attributed to what may be considered a ‘ceiling effect.’ In fact, the duration of contrastive-focused C1s (79.14 ms in Subtype-type NP1, 65.31 ms in Type-subtype NP1, 77.22 ms in Subtype-type NP2, and 80.61 ms in Type-subtype NP2) exceeded even the VOT of IP-medial focused lenis stops (approximately 65 ms) reported in a previous study (Choi et al. 2020),<sup>9</sup> which had participants of similar age ranges as our study. This suggests that there might be limited room within the contrastive-focused words, particularly in the NP2-contrast condition, for additional C1 lengthening resulting from identificational focus. This ‘ceiling effect’ phenomenon aligns with findings from prior literature (Cho 2006; Cho et al. 2011), which discussed the relationship between prominence-driven and boundary-driven prosodic strengthening. The present study demonstrates that such a ceiling effect can be found between different types of prominence-induced prosodic strengthening.

The patterns of F0 contours also interacted with the focus types. In terms of contrastive-focused NPs, the F0 of the second syllable was as high as the F0 of the third syllable (**Figure 6c** and **Figure 6f**). Note that, in identificational-focused NPs, the F0 of the second syllable was as high as that of the first syllable, yielding the F0 of the third syllable as the highest in NPs (**Figure 6a** and **Figure 6b**). These differences in F0 contours may be related to the different types of focus: identificational vs. contrastive focus. In the contrastive focus conditions, the AP-initial tone of the focused constructions was realized as L+H, yielding the peak F0 within the AP closer to the target of narrow focus: the nouns. On the other hand, in the default reading condition, the AP-initial tone of the focused constructions was realized as L, and the F0 peak of a Ha boundary tone was adjacent to the right edge of the APs. Considering that all other aspects of the target sentences are equal, the inconsistent prosodic realizations between identificational and contrastive focus may be ascribed to the fact that the contrastive-focused NPs contain the information that had to be corrected in the given discourse contexts. We look forward to seeing further research on diverse focus types and speakers’ preferences for their prosodic realizations.

## 6 Conclusion

In this paper, we discussed the prosodic realization of identificational focus and contrastive focus in Seoul Korean. Specifically focusing on the interaction between identificational and contrastive focus, we have provided concrete evidence for how speech variations coming from information structure are systematically realized in fine-grained phonetic details. The prominence-induced prosodic strengthenings coming from various sources (identificational focus and contrastive focus) were phonetically distinguishable in terms of tonal patterns, duration, and F0, supporting the view that the two types of focus (i.e., identificational and contrastive) are distinctive notions.

---

<sup>9</sup> Note that the target items of Choi et al. (2020) included an alveolar lenis stop (i.e., /t/), but the current study includes a velar lenis stop (i.e., /k/).

Although both identificational and contrastive focus showed longer duration and higher F0 on the target of the prosodic strengthening, such utilization was further conditioned following the sources of focus as manifested on the strength of durational and F0-related cues, and the different tonal patterns as well as pre- and post- focus compression patterns. These results align with the formal distinction between two focus types driven from the linguistic (syntactic) analysis. This distinction further proposes a suggestion to explain the contradictory patterns of focus effects found in previous literature on Seoul Korean. All in all, the current study demonstrates the importance of studying the interplay between phonetics, semantics, and syntax in the prosodic production of the intended speech sounds.

---

## Appendix

Dependent variable	Model formula
Frequency	Frequency ~ Boundary + Semantic Relations + Prosody
C1 duration	C1_dur ~ Semantic Relations + NP Locations + Prosody + Semantic Relations:NP Locations + NP Locations:Prosody + (1 Participant)
F0 (st)	F0_st ~ Semantic Relations + NP Locations + Prosody + Semantic Relations:NP Locations + NP Locations:Prosody + (1 Participant)

**Table 6:** The descriptions of the parsimonious models.

Models	$df^a$	LogLik	$\chi^2$	$\chi^2 df^b$	$p$
<b>The saturated model</b>	24	-20.41			
<b>The three-way interaction term</b>					
- Semantic Relations:Prosody:Boundary	20	-20.41	<0.01	4	1
<b>The two-way interaction term</b>					
- Semantic Relations:Prosody	16	-20.44	0.06	4	1
- Boundary:Semantic Relations	14	-20.47	0.05	2	.976
- Boundary:Prosody	10	-20.47	<0.01	4	1
<b>The parsimonious model<sup>c</sup></b>			79.19	8	<.001

**Table 7:** Summary of log-linear model analysis on the boundary realization of NP2-initial positions.

- This indicates the degrees of freedom of each model.
- This indicates the residual deviances from the summary of each model.
- This indicates the degrees of freedom between models.
- The parsimonious model was compared with the null model.

Parameter estimate	Estimate	Std. error	$z$	$p$
a. Count model coefficients (poisson with log link)				
Intercept	3.71	0.13	29.14	<.001
Boundary – IP	-1.69	0.28	-6.00	<.001
Prosody – NP1-contrast	-0.04	0.16	-0.24	.812
Prosody – NP2-contrast	-0.05	0.16	-0.32	.750
Semantic Relations – Topic-type	-0.02	0.13	-0.19	.850

(Contd.)

Parameter estimate	Estimate	Std. error	z	p
b. Zero-inflation model coefficients (binomial with log link)				
Intercept	-7.379e + 01	2.249e + 05	0	1
Boundary – IP	4.954e + 01	1.691e + 05	0	1
Prosody – NP1-contrast	4.902e + 01	1.960e + 05	0	1
Prosody – NP2-contrast	4.902e + 01	1.960e + 05	0	1
Semantic Relations – Topic-type	-9.532e-11	1.418e + 05	0	1

**Table 8:** Estimated parameters and related statistics for most parsimonious log-linear model for NP2-initial positions.

Models	df <sup>a</sup>	AIC	BIC	logLik	deviance	$\chi^2$	$\chi^2$ df	p
<b>The saturated model</b>	14	1253	1312	-612	1225			
<b>The three-way interaction term</b>								
– Boundary:Semantic Relations:Prosody	12	1249	1312	-612	1225	0.49	2	.782
<b>The two-way interaction term</b>								
– Semantic Relations:NP Locations	11	1266	1312	-622	1244	18.53	1	<.001
– NP Locations:Prosody	10	1330	1372	-655	1310	84.81	2	<.001
– Semantic Relations:Prosody	10	1247	1290	-614	1227	2.24	2	.327
<b>The parsimonious model</b>	10	1247	1290	-614	1227	2.24	2	.327

**Table 9:** Summary of results of linear regression model analysis on F0 (st) of V1-mid.  
a. This indicates the degrees of freedom between models.

Parameter estimate	Estimate	Std. error	df	t	p
Intercept	8.21	1.47	15	5.58	<.001
Semantic Relations[1] <sup>a</sup>	-0.01	0.03	475	-0.33	.742
NP Locations[1] <sup>b</sup>	0.24	0.03	475	7.14	<.001
Prosody[1] <sup>c</sup>	-0.01	0.05	475	-0.12	.904
Prosody[2] <sup>d</sup>	-0.12	0.05	475	-2.43	.016
Semantic Relations[1]:NP Locations[1]	0.14	0.03	475	4.3	<.001
NP Locations[1]:Prosody[1]	-0.1	0.05	475	-2.22	.027
NP Locations[1]:Prosody[2]	0.44	0.05	475	9.22	<.001

**Table 10:** Estimated parameters and related statistics for most parsimonious linear regression model for F0 (st) of V1-mid.

- The reference level of Semantic Relations was Subtype-type.
- The reference level of NP Locations was NP1.
- The reference level of Prosody was Default.
- The second level of Prosody was NP1-contrast.

## Abbreviations

NOM = nominative, ACC = accusative, PST = past, DECL = declarative, QUE = question

## Data availability

The supporting data for the study are available at: <https://doi.org/10.17605/OSF.IO/6VF8R>.

## Ethics and consent

The protocol of the current study was revised and approved by the Institutional Review Board at the University of Wisconsin-Milwaukee (IRB#: 19.102).

## Acknowledgements

This work has greatly benefited from the insightful feedback and suggestions of Sara Finley, Juliet Stanton, and three anonymous reviewers. We are thankful to Taehong Cho for his support in granting us access to his lab, which enabled us to collect data for our research. We presented the pilot study of this research at the Hanyang International Symposium on Phonetics and Cognitive Sciences of Language 2019 (HISPhonCog 2019), where we received valuable comments from Taehong Cho, Sun-Ah Jun, and Jason Shaw. We also appreciate the helpful remarks of Anne Pycha and Jae Yung Song on an earlier draft of this paper. Our gratitude extends to Jieun Lee for her suggestions on the target items.

## Competing interests

The authors have no competing interests to declare.

---

## References

- Agresti, Alan. 2003. *Categorical data analysis* 2nd ed. Hoboken, NJ: John Wiley & Sons.
- Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Bates, Douglas & Maechler, Martin & Bolker, Ben & Walker, Steven. 2013. lme4: Linear mixed-effects models using Eigen and S4. Retrieved from <http://cran.r-project.org/package=lme4>
- Beckman, Mary E. & Pierrehumbert, Janet B. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309. DOI: <https://doi.org/10.1017/S095267570000066X>
- Bennett, Ryan & Ní Chiosáin, Máire & Padgett, Jaye & McGuire, Grant. 2018. An ultrasound study of Connemara Irish palatalization and velarization. *Journal of the International Phonetic Association* 48(3). 261–304. DOI: <https://doi.org/10.1017/S0025100317000494>

- Boersma, Paul & Weenink, David. 2016. Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org/>
- Bolinger, Dwight L. 1961. Contrastive accent and contrastive stress. *Language* 37(1). 96–96. DOI: <https://doi.org/10.2307/411252>
- Bolinger, Dwight. 1972. Accent is predictable (if you're a mind-reader). *Language* 48(3). 644–644. DOI: <https://doi.org/10.2307/412039>
- Chae, Hee-Rahk and Ilkyu Kim. 2008. A clausal predicate analysis of Korean multiple nominative constructions. *Korean Journal of Linguistics* 33(4). 869–900.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, Charles N. (ed.), *Subject and Topic*, 25–55. New York, NY: Academic Press.
- Chafe, Wallace. 1987. Cognitive constraints on information flow. In Tomlin, Russell S. (ed.), *Coherence and grounding in discourse*, 21–51. Philadelphia, PA: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/tsl.11.03cha>
- Cho, Taehong. 2006. Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In Goldstein, Louis & Whalen, Douglas Harry & Best, Catherine T. (eds.), *Laboratory phonology 8*, 519–548. New York, NY: Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110197211.3.519>
- Cho, Taehong. 2015. Language effects on timing at the segmental and suprasegmental levels. In Redford, Melissa A. (ed.), *The handbook of speech production*, 505–529. Hoboken, NJ: John Wiley & Sons, Inc. DOI: <https://doi.org/10.1002/9781118584156.ch22>
- Cho, Taehong. 2016. Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass* 10(3). 120–141. DOI: <https://doi.org/10.1111/lnc3.12178>
- Cho, Taehong. 2022. The phonetics-prosody interface and prosodic strengthening in Korean. In Cho, Sungdai & Whitman, John (eds.), *The Cambridge handbook of Korean linguistics*, 248–293. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781108292351.010>
- Cho, Taehong & Kim, Dong Jin & Kim, Sahyang. 2019. Prosodic strengthening in reference to the lexical pitch accent system in South Kyungsang Korean. *Linguistic Review* 36(1). DOI: <https://doi.org/10.1515/tlr-2018-2008>
- Cho, Taehong & Lee, Yoonjeong & Kim, Sahyang. 2011. Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics* 39(3). 344–361. DOI: <https://doi.org/10.1016/j.wocn.2011.02.005>
- Choi, Jiyoung & Kim, Sahyang & Cho, Taehong. 2020. An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PLoS ONE* 15(10). e0240682–e0240682. DOI: <https://doi.org/10.1371/journal.pone.0240682>
- Choi, Jiyoung & Lee, Jiyoung & Kim, Sahyang & Cho, Taehong. 2018. Prosodically-conditioned phonetic cue use in production of Korean aspirated vs. lenis stops. In Cho, Taehong & Kim, Sahyang & Kim, Jonny Jungyun & Kim, Say Young & Lee, Ki-Jeong (eds.), *Proceedings of Hanyang International Symposium on Phonetics and Cognitive Sciences of Language (HISPhonCog 2018)*, 121–123. Seoul: Hanyang Institute for Phonetics and Cognitive Sciences of Language.

- Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America* 120(1). 407–415. DOI: <https://doi.org/10.1121/1.2205133>
- Derrick, Donald & Schultz, Benjamin. 2013. Acoustic correlates of flaps in North American English. In *Proceedings of meetings on acoustics*, Vol. 19. 060260–060260. Acoustical Society of America. DOI: <https://doi.org/10.1121/1.4798779>
- Dik, Simon C. 1997. *Part 1: The structure of the clause*. (K. Hengeveld, Ed.) 2nd ed. New York, NY: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110218367>
- Dohen, Marion & Loevenbruck, Helene. 2004. Pre-focal rephrasing, focal enhancement and postfocal deaccentuation in French. In *Proc. Interspeech 2004*, 785–788. ISCA. DOI: <https://doi.org/10.21437/Interspeech.2004-296>
- Féry, Caroline. 2013. Focus as prosodic alignment. *Natural Language and Linguistic Theory* 31(3). 683–734. DOI: <https://doi.org/10.1007/s11049-013-9195-7>
- Féry, Caroline & Krifka, Manfred. 2008. Information structure: Notional distinctions, ways of expression. In Sterkenburg, Piet van (ed.), *Unity and diversity of languages*, 123–135. Philadelphia, PA: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/z.141.13kri>
- Fodor, Janet Dean. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27(2). 285–319. DOI: <https://doi.org/10.1023/A:1023258301588>
- Fodor, Janet Dean. 2002. Prosodic disambiguation in silent reading. In Hirotani, Masako (ed.), *Proceedings of the North East Linguistics Society*, Vol. 32. 113–132. New York, NY: GLSA, Department of Linguistics, University of Massachusetts, Amherst. Retrieved from <https://scholarworks.umass.edu/nels/vol32/iss1/8>
- Gu, Chong. 2014. Smoothing spline ANOVA models: R package gss. *Journal of Statistical Software* 58(5). 1–25. DOI: <https://doi.org/10.18637/jss.v058.i05>
- Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19(2). 377–417. DOI: <https://doi.org/10.1017/S0022226700007799>
- Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511616983>
- Halliday, Michael Alexander Kirkwood. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics* 3(2). 199–244. DOI: <https://doi.org/10.1017/S0022226700016613>
- Jackman, Simon. 2020. pscl: Classes and methods for R developed in the political science computational Laboratory. R package version 1.5.5.1. Retrieved from <https://github.com/atahk/pscl/>
- Jeon, Hae-Sung & Nolan, Francis. 2017. Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology* 8(1). 1–30. DOI: <https://doi.org/10.5334/labphon.48>
- Jun, Sun-Ah. 1993. *The phonetics and phonology of Korean prosody*. Columbus, OH: The Ohio State University dissertation.
- Jun, Sun-Ah. 2000. *K-Tobi Labelling Conventions*. Los Angeles, CA.
- Jun, Sun-Ah. 2003. Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* 32(2). 219–249. DOI: <https://doi.org/10.1023/A:1022452408944>



- Jun, Sun-Ah. 2005. Prosodic typology. In Jun, Sun-Ah (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 430–458. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199249633.003.0016>
- Jun, Sun-Ah. 2006. Intonational phonology of Seoul Korean revisited. In Vance, Timothy J & Jones, Kimberly (eds.), *Japanese/Korean Linguistics*, Vol 14. 15–26. Stanford, CA: CSLI Publications.
- Jun, Sun-Ah. 2014. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In Jun, Sun-Ah (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 520–539. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199567300.003.0017>
- Jun, Sun-Ah & Fougeron, Cécile. 2000. A phonological model of French intonation. In Botinis, Antonis (ed.), *Intonation: Analysis, modelling and technology*, Vol. 15. 209–242. Dordrecht: Springer. DOI: [https://doi.org/10.1007/978-94-011-4317-2\\_10](https://doi.org/10.1007/978-94-011-4317-2_10)
- Jun, Sun-Ah & Kim, Hee-Sun. 2007. VP focus and narrow focus in Korean. In Trouvain, Jürgen & Barry, William J (eds.), *Proceedings of the XVIth International Congress of Phonetic Sciences*, 1277–1280. Saarbrücken: Universität des Saarlandes.
- Jun, Sun-Ah & Kim, Hee-Sun & Lee, Hyuck-Joon & Kim, Jong-Bok. 2006. An experimental study on the effect of argument structure on VP focus. *Korean Linguistics* 13(1). 89–113. DOI: <https://doi.org/10.1075/kl.13.04saj>
- Jun, Sun-Ah & Lee, Hyuck-Joon. 1998. Phonetic and phonological makers of contrastive focus in Korean. In *Proceedings of the 5th International Conference of Spoken Language Processing (ICSLP 98)*, 1087. Sydney. DOI: <https://doi.org/10.21437/ICSLP.1998-151>
- Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1(1–3). 181–194. DOI: <https://doi.org/10.1515/thli.1974.1.1-3.181>
- Katz, Jonah & Selkirk, Elisabeth. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87(4). 771–816. DOI: <https://doi.org/10.1353/lan.2011.0076>
- Kawahara, Hideki & Cheveigne, de Alain & Patterson, Roy D. 1998. An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: revised TEMPO in the STRAIGHT-suite. In Mannell, R. H & Robert-Ribes, Jordi (eds.), *Proc. Fifth International Conference on Spoken Language Processing (ICSLP 1998)*, 0659. Sydney: Australian Speech Science and Technology Association, Incorporated (ASSTA). DOI: <https://doi.org/10.21437/ICSLP.1998-555>
- Kenesei, István. 2006. Focus as identification. In Molnár, Valéria & Winkler, Susanne (eds.), *The Architecture of Focus*, 137–168. Berlin: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110922011.137>
- Kim, Okgi. 2020. Reformulative multiple accusative constructions as vacuous reformulative appositives. In *Proceedings of the Chicago Linguistic Society* 56, 243–256.
- Kim, Okgi. 2023. Korean reformulative multiple accusative construction as vacuous reformulative apposition. *Linguistic Research* 40(2). 217–244. DOI: <https://doi.org/10.17250/khisli.40.2.202306.003>
- Kiss, Katalin É. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273. DOI: <https://doi.org/10.1353/lan.1998.0211>

- Klok, Jozina Vander & Goad, Heather & Wagner, Michael. 2018. Prosodic focus in English vs. French: A scope account. *Glossa: A Journal of General Linguistics* 3(1). DOI: <https://doi.org/10.5334/gjgl.172>
- Kratzer, Angelika & Selkirk, Elisabeth. 2007. Phase theory and prosodic spellout: The case of verbs. *Linguistic Review* 24(2–3). 93–135. DOI: <https://doi.org/10.1515/TLR.2007.005>
- Labov, William. 1984. Field methods of the project on linguistic change and variation. In Baugh, John & Sherzer, Joel (eds.), *Language and Use*, 28–53. Englewood Cliffs, NJ: Prentice-Hall.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620607>
- Lee, Yong-cheol. 2017. Prosodic focus in Seoul Korean and South Kyungsang Korean. *Linguistic Research* 34(1). 133–161. DOI: <https://doi.org/10.17250/khisli.34.1.201703.005>
- Lee, Yong-cheol & Cho, Sunghye. 2020. Focus prosody varies by phrase-initial tones in Seoul Korean: Production, perception, and automatic classification. *Languages* 5(4). 64. DOI: <https://doi.org/10.3390/languages5040064>
- Lee, Yong-cheol & Xu, Yi. 2010. Phonetic realization of contrastive focus in Korean. In Hasegawa-Johnson, Mark (ed.), *Proc. Speech Prosody 2010*, 033. Chicago, IL.
- Lee-Kim, Sang-Im & Davidson, Lisa & Hwang, Sangjin. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4(2). 475–511. DOI: <https://doi.org/10.1515/lp-2013-0015>
- Lewis, David. 1979. Scorekeeping in a language game. In Bäuerle, Rainer & Egli, Urs & Stechow, Arnim von (eds.), *Semantics from Different Points of View*, 172–187. Berlin: Springer. DOI: [https://doi.org/10.1007/978-3-642-67458-7\\_12](https://doi.org/10.1007/978-3-642-67458-7_12)
- Molnár, Valéria. 2001. Contrast from a contrastive perspective. In Kruijff-Korbayvá, Ivana & Steedman, Mark (eds.), *Information structure, discourse structure and discourse semantics: Workshop Proceedings*, 99–114. Helsinki: The University of Helsinki.
- Neeleman, Ad & van de Koot, Hans. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11(2). 137–189. DOI: <https://doi.org/10.1007/s10828-008-9018-0>
- Oh, Mi-Ra. 2001. Focus and prosodic structure. *Speech Sciences* 8(1). 21–31.
- Park, Chongwon. 2013. Metonymy in grammar. *Functions of Language* 20(1). 31–63. DOI: <https://doi.org/10.1075/fol.20.1.02par>
- Pedhazur, Elazar J. 1997. *Multiple regression in behavioral research: explanation and prediction* 3rd edition. Fort Worth, TX: Harcourt Brace College Publishers.
- R Core Team. 2020. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Rochemont, Michael S. 1986. *Focus in Generative Grammar*. Philadelphia, PA: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/sigla.4>

- Shattuck-Hufnagel, Stefanie & Turk, Alice E. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247. DOI: <https://doi.org/10.1007/BF01708572>
- Shue, Yen-Liang. 2010. *The voice source in speech production: Data, analysis and models*. Los Angeles, CA: University of California, Los Angeles dissertation.
- Shue, Yen-Liang & Keating, Patricia & Vicenik, Chad & Yu, Kristine. 2011. Voicesauce: A program for voice analysis. In Lee, Wai-Sum & Zee, Eric (eds.), *Proceedings Of The 17th International Congress Of Phonetic Sciences*, Vol. 3. 1846–1849. Hong Kong.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In Munitz, Milton K & Unger, Peter (eds.), *Semantics and Philosophy*, 197–213. New York, NY: New York University Press.
- Steube, Anita. 2001. Correction by contrastive focus. *Theoretical Linguistics* 27. 215–249. DOI: <https://doi.org/10.1515/thli.2001.27.2-3.215>
- Vallduví, Enric & Vilkuna, Maria. 1998. On rheme and kontrast. In Culicover, Peter & McNally, Louise (eds.), *The limits of syntax*, Vol. 29. 79–108. Leiden: Brill. DOI: [https://doi.org/10.1163/9789004373167\\_005](https://doi.org/10.1163/9789004373167_005)
- Xu, Yi & Xu, Ching X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33(2). 159–197. DOI: <https://doi.org/10.1016/j.wocn.2004.11.001>
- Yang, Anqi & Cho, Taehong & Kim, Sahyang & Chen, Aojun. 2015. Phonetic focus-marking in Korean-speaking 7- to 8-year-olds and adults. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, 0673. Glasgow: The University of Glasgow.
- Yiu, Suki. 2015. Intonation of statements and questions in Cantonese English: acoustic evidence from a smoothing spline analysis of variance. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, 1018. Glasgow: The University of Glasgow.
- Yoon, James Hye Suk. 2015. Double nominative and double accusative constructions. In Brown, Lucien & Yeon, Jaehoon (eds.), *The handbook of Korean linguistics*, 79–97. West Sussex: John Wiley & Sons, Inc. DOI: <https://doi.org/10.1002/9781118371008.ch5>
- Zeileis, Achim & Kleiber, Christian & Jackman, Simon. 2008. Regression models for count data in R. *Journal of Statistical Software* 27(8). 1–25. DOI: <https://doi.org/10.18637/jss.v027.i08>
- Zimmermann, Malte & Onea, Edgar. 2011. Focus marking and focus interpretation. *Lingua* 121(11). 1651–1670. DOI: <https://doi.org/10.1016/j.lingua.2011.06.002>

