Downstep in Japanese Revisited : Morphology Matters

著者(英)	Hyun Kyung HWANG, Manami HIRAYAMA
journal or publication title	NINJAL Research Papers
number	21
page range	15-23
year	2021-07
URL	http://doi.org/10.15084/00003434

国立国語研究所論集 (NINJAL Research Papers) 21: 15–23 (2021)
ISSN 2186-1358

Downstep in Japanese Revisited: Morphology Matters

HWANG Hyun Kyung[a] HIRAYAMA Manami[b]

[a]University of Tsukuba /
Adjunct Researcher, Theory & Typology Division, Research Department, NINJAL [–2017.03] /
Project Collaborator, NINJAL
[b]Seikei University

Abstract

It has been widely acknowledged that an f0 after an accented word is noticeably lower than after an unaccented word in Japanese (i.e., downstep). In pursuing research concerning the syntax-phonology interface, Selkirk and Tateishi (1991) highlight the lack of downstep at the left edges of maximal projections of syntactic categories (XPs), and propose that the left edges of XPs are mapped onto the left edges of the major phrase boundaries that block downstep. Kubozono (1991) provides empirical results that are different from Selkirk and Tateishi (1991). This study aims to test the effect of different lexical categories on downstep, as the distinction between different categories is frequently ignored in the literature regarding Japanese downstep.

The results from a production test suggest the possibility that distinct lexical categories influence the presence or absence of downstep in Japanese. Specifically, downstep occurred in the noun condition but not in the adjective condition with *-i* endings, which casts doubt on Selkirk and Tateishi's generalization. We propose that the blocking of downstep is due to the perceived unnaturalness associated with the morphology of adjectives. Moreover, this research reveals an important implication regarding the methodology: lexical category and morphological forms should be carefully controlled for studying Japanese downstep.*

Keywords: lexical category, production test, downstep, Japanese

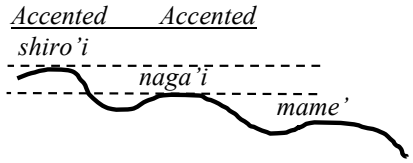
1. Introduction

It has been widely acknowledged that the fundamental frequency (henceforth f0) peak of the material following an accented word is noticeably lower than that following an unaccented word in Japanese (Poser 1984, Pierrehumbert and Beckman 1988, Kubozono 1989, among others). This f0 lowering effect triggered by a preceding accented item is referred to as *downstep*. In pursuing research concerning the syntax-phonology interface, Selkirk and Tateishi (1991) highlight the lack of downstep at the left edges of maximal projections of syntactic categories (henceforth XPs), and propose that the left edges of XPs are mapped onto the left edges of the major phrase boundaries that block downstep. More interestingly, they compare two lexical categories: adjectives (A) and nouns (N) ([A_1 [$\underline{A}_2$ N]] vs. [N_1 [$\underline{N}_2$ N_3]]), and report that downstep occurs only in A_2, not in N_2. Based on this observation, Selkirk and Tateishi claim that N_2 is at the left edge of a maximal projection NP, whereas [A_2 N] does not constitute a maximal projection.

* We would like to thank the audience at LabPhon 15 and Haruo Kubozono for their valuable comments. The work reported in this article was partially supported by the NINJAL collaborative research project 'Cross-linguistic Studies of Japanese Prosody and Grammar' (project leader: Haruo Kubozono) and JSPS KAKENHI grant #19K00613.

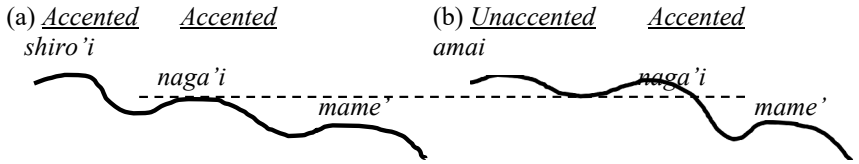
Methodologically, Selkirk and Tateishi's diagnosis is rather *syntagmatic* because the mean f0s of two adjacent phrases are compared within a single utterance. As displayed in the schematic f0 curves (1), the presence of downstep is determined within a sentence. Apostrophes indicate lexical pitch accents. For instance, in a phrase *shiro'i naga'i mame*' 'long white beans (literally, white long bean(s))', where there are two accented phrases *shiro'i* and *naga'i* in a row, and thus the first accented phrase *shiro'i* potentially triggers the second phrase *naga'i* to be downstepped, the second accented phrase *naga'i* is considered downstepped if the pitch is lower in this phrase when compared to the preceding phrase.

(1) Schematic pitch curves in syntagmatic downstep



Traditionally, however, the presence or absence of downstep has been tested paradigmatically by comparing the f0 peaks of targets following accented and unaccented words. In the paradigmatic description of downstep in Japanese (e.g., Poser 1984, Kubozono 1989, Kubozono 2007), the presence of downstep is diagnosed by comparing f0 curves in two sentences/phrases. For example, the *naga'i* phrase in *shiro'i naga'i mame*' 'white long beans' is considered downstepped if the f0 peak in this phrase is lower when compared to an equivalent phrase from another sentence that has the same structure except that the preceding phrase is unaccented, e.g., *amai naga'i mame*' 'long sweet beans', where the first phrase *amai* is unaccented. This is illustrated in (2). Whether or not the phrase *naga'i* is downstepped can be diagnosed only by comparing the f0 peak in that phrase with the f0 peak in the second phrase in (b), in which it must not be downstepped due to the preceding unaccented phrase. The phrase *naga'i* is downstepped in (2a) if its pitch is lower than the pitch in *naga'i* in (2b).

(2) Schematic pitch curves in the paradigmatic downstep



It is not a trivial issue to diagnose the presence or absence of downstep using the syntagmatic diagnostic since it is not obvious how much f0 lowering is necessary as an indication of downstep. Thus, we mainly adopt the paradigmatic diagnostic in this study, although the syntagmatic diagnostic has been used in much of the literature on Japanese downstep (Selkirk and Tateishi 1991, Nagahara 1994, Hirotani 2005).

Regardless of the methodology, the distinction between different categories is frequently ignored in the literature on Japanese downstep. One tested pair from Kubozono (1989) is given in (3). In (3a), the trigger *ao'i* is an adjective, while the target *momen* is a noun. Further, the target *marui* in the utterance for comparison (3b) is an adjective, which is a different category from that

in (3a).

(3) A test set in Kubozono (1989)
   a. ao'i momen-no ori'mono
      blue cotton-GEN fabric
      'blue cotton fabric'
   b. amai marui me'ron
      sweet round melon
      'sweet round melon'

Kubozono (1991) investigated two speakers with more controlled phrases, in particular [N [N N]] (*a'ni-no men'no eri'maki* '(my) brother's cotton muffler') and [A [A N]] (*ao'i o'okina eri'maki* 'a blue, big muffler'), both of which showed downstep (see section 4 for more discussion). Selkirk and Tateishi (1991) mention that Adjective Phrase (AP) boundaries do not block downstep unlike Noun Phrase (NP). Thus, the literature does not agree on the presence or absence of downstep in a particular category. Testing the effect of different lexical categories on downstep is an important line of research, and a more strictly controlled paradigmatic investigation with a larger set of data is needed to confirm the probable influence of lexical categories. In the current study, we investigate the effect of lexical categories, particularly N and A, paradigmatically. The results indicate that these two conditions influence downstep differently in Japanese. In addition, it implies that a new approach that is capable of accounting for a fuller range of data is necessary.

2. Methods

To investigate whether there is a difference in the downstep patterns depending on the lexical categories, particularly in phrases involving adjectives and those involving nouns, a production experiment was conducted. This section describes the design and data used in the experiment.

2.1 Material

In each of the A and N conditions, two phrases with accented triggers and two phrases with unaccented "triggers" were compared. In examples (4) and (5), (4) involves adjectival modifiers, and (5) involves nominal modifiers. In both (4) and (5), sentences in (a) and (b) have the same phonological lengths (counted by mora) and right-branching structure [X [X X]]. The mora is a phonological unit in Japanese, and (C)V and coda consonants are counted as one mora (see, for example, Kubozono 1999, for the mora in Japanese). In the adjectival modifier condition, as in (4), a noun phrase containing two adjectives modifies the head noun [A_1 [A_2 N]], for example, *shiro'i naga'i mame'* 'long white beans' (literally 'white long bean(s)') (4a). Sentences (4a) and (4b) differ in terms of accentuation on A_1, which is the trigger of downstep. In (4a), the trigger (*shi-ro'i*) is accented, while it (*amai*) is unaccented in (4b). The f0s of targets are compared to investigate paradigmatic downstep. A lower f0 in the target with an accented trigger (*naga'i* in (4a)) than in the target with an unaccented trigger (*naga'i* in (4b)) is the indication of the presence of downstep.

(4) Adjectival modifiers

	Trigger	*Target*	*Head N*	
a. a'ni-wa	shiro'i	naga'i	mame'-o	utta.
brother-TOP	white	long	bean-ACC	sell.PAST
'My brother sold long white beans.'				
b. a'ni-wa	amai	naga'i	mame'-o	utta.
brother-TOP	sweet	long	bean-ACC	sell.PAST
'My brother sold long sweet beans.'				

In the phrases with nominal modifiers [N_1 [N_2 N]], as in (5), two nouns, N_1 and N_2, modify the head noun, for example, *no'mo-no na'ra-no mame'* 'Nomo's beans from Nara' (literally 'Nomo's Nara's bean(s)'). The particle *-no* that attaches to a noun represents the modification relation between nouns. Once more, (5a) and (5b) differ in terms of whether the trigger is accented (5a) or unaccented (5b). The two modifying nouns (N_1 and N_2) are proper nouns, with N_1 being either a surname or a given name and N_2 being *Nara*, a city name. Proper nouns were used because common nouns were difficult to find to control for experimental factors, including the segmental structure, number of moras, and accentuation.

(5) Nominal modifiers

	Trigger	*Target*	*Head N*	
a. a'ni-wa	no'mo-no	na'ra-no	mame'-o	utta.
brother-TOP	Nomo-GEN	Nara-GEN	bean-ACC	sell.PAST
'My brother sold Nomo's beans from Nara.'				
b. a'ni-wa	ono-no	na'ra-no	mame'-o	utta.
brother-TOP	Ono-GEN	Nara-GEN	bean-ACC	sell.PAST
'My brother sold Ono's beans from Nara.'				

There were two phrases for each of the types in (4) and (5) in the experiment. Therefore, a total of eight phrases were tested.[1]

2.2 Recording and measurements

The [A_1 [A_2 N]] and [N_1 [N_2 N]] phrases discussed in Section 2.1 were put in a carrier phrase *a'ni-wa __ -o utta* 'my brother sold ___'. The sentences were divided into two lists so that each list contained one item of a pair of accented and unaccented sentences. Each list had 18 randomized distractor sentences; these other sentences were for other projects and are not discussed

[1] The other test sentences are given below.

1. Adjectival modifiers [A_1 [A_2 N]]

a.	a'ni-wa	uma'i	naga'i	mame'-o	utta.	'My brother sold long delicious beans.'
	brother-TOP	good	long	bean-ACC	sell.PAST	
b.	a'ni-wa	marui	naga'i	mame'-o	utta.	'My brother sold long round beans.'
	brother-TOP	round	long	bean-ACC	sell.PAST	

2. Nominal modifiers [N_1 [N_2 N]]

a.	a'ni-wa	ma'mi-no	na'ra-no	mame'-o	utta.	'My brother sold Mami's beans from Nara.'
	brother-TOP	Mami-GEN	Nara-GEN	bean-ACC	sell.PAST	
b.	a'ni-wa	yano-no	na'ra-no	mame'-o	utta.	'My brother sold Yano's beans from Nara.'
	brother-TOP	Yano-GEN	Nara-GEN	bean-ACC	sell.PAST	

in this paper. Six speakers from Tokyo or nearby areas participated in the study (two males and four females, mean age 30.3, range 24–33). They repeated the two lists eight times. Thus, there were 384 tokens recorded in total. The recording was performed in a room with sound-attenuating walls. We used a Marantz PMD661 recorder with a sampling rate of 44.1 kHz and 24-bit quantization, and a unidirectional dynamic microphone (SHURE SM10A, frequency response: 50–15,000 Hz).

We measured peak f0 of each of the phrases in [A_1 [A_2 N]] and [N_1 [N_2 N]] plus the subject/topic phrase *a'ni-wa* as the reference point, as illustrated in (6), in which the boundaries are marked with a vertical bar (|). The f0 measurements were performed by using Praat and running a script called *ProsodyPro* (Xu 2013) on the intervals made manually. A screenshot indicating the intervals (numbered 1 to 4 for the phrases) is presented in Figure 1. We measured the peak f0s so that our work is comparable with the data in the literature where the paradigmatic diagnosis was utilized (Kubozono 1989, 1991).

(6) Peak f0s in four intervals

	Reference	*Trigger*	*Target*	*Head N*	
a.	| a'ni-wa	| shiro'i	| naga'i	| mame' | -o	utta.
	brother-TOP	white	long	bean-ACC	sell.PAST
	'My brother sold long white beans.'				
b.	| a'ni-wa	| no'mo-no	| na'ra-no	| mame' | -o	utta.
	brother-TOP	Nomo-GEN	Nara-GEN	bean-ACC	sell.PAST
	'My brother sold Nomo's beans from Nara.'				

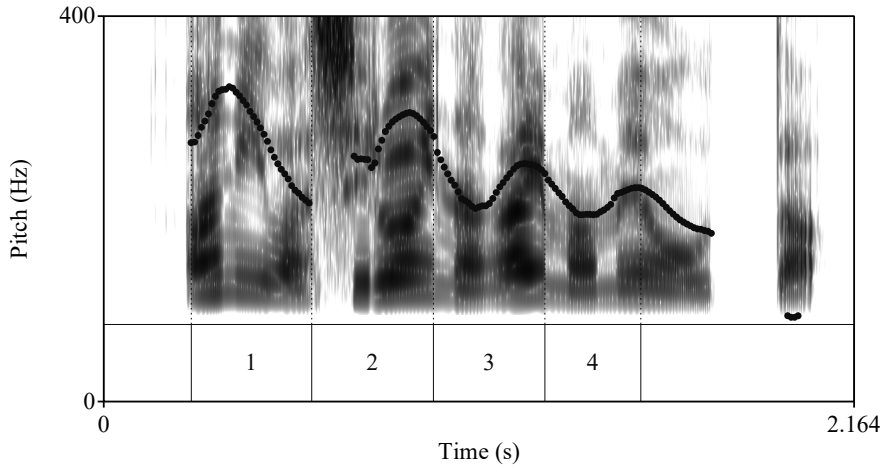


Figure 1 The spectrogram and pitch contour of (6a) *a'niwa shiro'i naga'i mame'o utta* uttered by Speaker RA, with intervals indicating the reference phrase (1), the trigger phrase (2), the target phrase (3) and the head noun (4).

For statistical analyses, linear mixed-effects models were used. The speaker and item were entered into the model as random effects; random intercepts and random slopes were included for the speaker and random intercepts were included for the item (Barr et al. 2013). Linear mixed-effects analysis with R (R Core Team 2014) was performed on the relation between the peak f0s of

the target XPs and the accentedness of the trigger XP (i.e., accented vs. unaccented).

3. Results

The influence of accentedness (A vs. U) on downstep differed between NP and AP. Figure 2 displays the average peak f0s of the four conditions: 2 categories (NP/AP) × 2 accentedness (A/U). As for N (right), the peak f0s of the targets in the unaccented condition (NP_U; dotted line) were remarkably higher than those in the accented counterparts (NP_A; solid line), suggesting that the targets were downstepped. As for A (left), on the other hand, no such difference was found in the target between accented (AP_A; solid line) and unaccented (AP_U; dotted line) conditions; the difference is quite minimal.

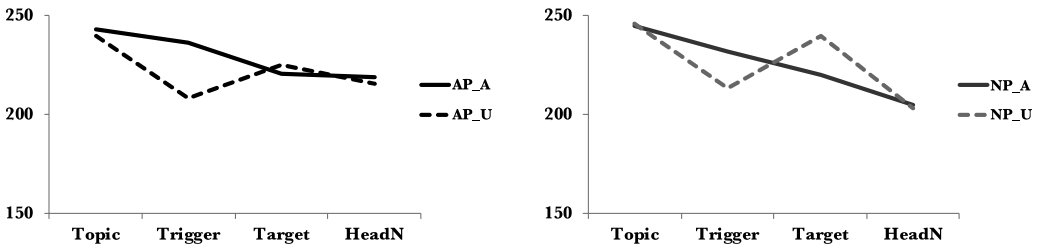


Figure 2 Average peak f0s of the four phrases in the AP (left) and NP (right) conditions

These observations were confirmed in the statistical analyses, where A and N did indeed differ in terms of the effect of accentedness. The f0s of the targets between the accented and unaccented tokens for A and N were compared individually, thereby revealing that the unaccented targets (AP_U & NP_U) yielded quite different F0s depending on their lexical category (see Figure 2). In N, there was a significant effect of the accent (Table 1). The peak f0 of the target phrase is higher in phrases where the trigger is unaccented than in phrases where the trigger is accented. As indicated in Table 1, the model estimates it to be higher by 19.936 Hz. This difference was statistically significant (t = 4.934, p = 0.003742). Hence, downstep occurred in the N condition.

Table 1 Results of the statistical analysis for NP

Factor	β-coefficient	t	p
Intercept	219.646	10.634	0.000119 ***
Trigger unaccented	19.936	4.934	0.003742 **

In contrast, there was no effect of the accent in the results involving A (Table 2). The accent difference on the trigger AP does not have a significant effect on the f0 of the target AP (t = 1.143, p = 0.304720), indicating that downstep did not occur in this condition.

Table 2 Results of the statistical analysis for AP

Factor	β-coefficient	t	p
Intercept	220.285	10.490	0.000136 ***
Trigger unaccented	4.911	1.143	0.304720

4. Discussion and conclusions

The current study explored the influence of the parts of speech in the presence or absence of downstep in Japanese. Specifically, we compared [N_1 [N_2 N]] and [A_1 [A_2 N]] to examine whether N_2 and A_2 undergo downstep given the accented triggers N_1 and A_1, respectively. The results of a production study suggest that distinct lexical categories make a significant difference in terms of downstep: downstep was observed at N_2 in [N_1 [N_2 N]] (which replicates Kubozono's (1991) results in a comparable structure), whereas it was blocked at A_2 in [A_1 [A_2 N]].

The results with adjectives appear to contradict those in Kubozono's (1991) results, where downstep was observed in the phrase [*ao'i* [*o'okina eri'maki*]] 'a blue, big muffler'. It is possible that the seemingly contradictory results can simply be attributed to the speaker/age difference. However, there may not be any contradiction between the two studies if we carefully consider the test phrases and the morphology of the words in those phrases.

In fact, the consideration suggests two possible analyses of the results with respect to the influence on downstep. In Japanese morphology, the words used in the adjective condition in this study (*shiro'i*, *naga'i*, *uma'i*, *marui*, *amai*) are usually analyzed as adjectives while *ookina* 'big' may allow more than one analysis. One possible category it belongs to is *rentaishi* (e.g., Masuoka and Takubo 1992: 55). *O'okina* also displays some properties common to adjectival nouns (*keiyoudoushi*) (Kageyama 1993: 25). Thus, if we assume that *o'okina* is not in the same lexical category as words used in this study, the different results between this study and Kubozono's (1991) can mean that different lexical categories affect the presence and absence of downstep. Still another possibility for *o'okina* in terms of its part of speech is that it is a form of adjective (p.c. Haruo Kubozono). If we assume this, then the different results between the two studies can be said to show that different morphological forms within the category (here adjective) may affect downstep: the *-i* form of the adjective blocks downstep, whereas the *-na* form does not.

What does the morphological distinction mean linguistically with respect to the presence and absence of downstep? The difference between N and A may be attributed to the syntactic characteristics that are exhibited by each category, particularly A. That is, what matters is probably not the lexical category *per se*, but the different usages of adjectives—attributive vs. predicative. In fact, Kuno (1973) separates the two usages and considers the attributive use as a relative clause. However, Hirayama and Hwang (2019) and Hirayama, Hwang, and Kato (2019) find that the left edge of relative clauses does not block downstep, which does not support this line of explanation if we analyze the attributive use of an adjective as a relative clause.

Alternatively, perceived naturalness may account for the different patterns between the morphological differences. For example, a noun phrase that has two adjectives with *-i* ending individually modifying the head noun ([A-i_1 [A-i_2 N]]), as in (2a) [*shiroi* [*nagai mame*]] 'long white beans', does not sound quite natural, though not ungrammatical, in Japanese. We propose that because of this unnaturalness, speakers tend to insert a prosodic phrase boundary between the two adjectives, which results in a pitch reset by blocking downstep. Indeed, the structure where the first adjective appears in the *-te* (or gerundive) form (i.e., [[A-*te* A] N]) sounds more natural (e.g., [[*shiroku-te nagai*] *mame*] 'white and long beans') when two adjectives are used to modify a noun. Similarly, when the second word is *o'okina* with *-na* in the end, as in phrases by Kubozono (1991), the phrase does not sound unnatural at all. In addition, phrases in the N condition [N-*no* [N-*no* N]] (Kubozono 1991, this study) do not give rise to any perceived unnaturalness when Japanese speakers parse them.

Furthermore, the empirical results provided in the current study appear to contradict those reported by Selkirk and Tateishi (1991). In particular, the occurrence of downstep in the N condition corroborates the argument by Kubozono (1988, 1989, 1991) and Pierrehumbert and Beckman (1988), which casts doubt on Selkirk and Tateishi's generalizations concerning the syntax-phonology interface. As mentioned earlier, they argue that major phrase boundaries in Japanese are inserted at the left edges of XPs. Given this generalization, a pitch reset rather than a downstep is expected in the structure $[N_1\ [N_2\ N]]$. One source of the discrepancy may be due to the different methodologies: downstep was diagnosed based on the syntagmatic analysis by Selkirk and Tateishi (1991). On the other hand, the paradigmatic analysis in this study and by Kubozono (1991) found that NP did not block downstep, suggesting that a new approach is necessary that is capable of accounting for a fuller range of data.

Moreover, this research has an important implication regarding the methodology of Japanese phonology. In testing downstep, sentences are commonly used without considering different lexical categories or different morphological forms. However, these results indicate that f0s are affected by the morphology of materials, suggesting that the morphological category should be carefully controlled in the study of Japanese downstep.

References

Barr, Dale J., Roger Levy, Christoph Scheepers and Harry J. Tily (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–278.

Hirayama, Manami and Hyun Kyung Hwang (2019) Relative clause and downstep in Japanese. In: Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose and Matt Zaslansky (eds.), *Supplemental Proceedings of the 2018 Annual Meeting on Phonology*. (https://doi.org/10.3765/amp.v7i0.4466). Washington, DC: Linguistic Society of America.

Hirayama, Manami, Hyun Kyung Hwang and Takaomi Kato (2019) Lexical category in downstep in Japanese. In: Sasha Calhoun, Paola Escudero, Marija Tabain and Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences*, 2851–2855.

Hirotani, Masako (2005) *Prosody and LF: Processing Japanese Wh-questions*. Doctoral dissertation, Amherst: University of Massachusetts.

Kageyama, Taro (1993) *Bunpoo-to gokeisei* (Grammar and word formation). Tokyo: Hituzi Syobo.

Kubozono, Haruo (1988) *The organization of Japanese prosody*. Doctoral dissertation, University of Edinburgh.

Kubozono, Haruo (1989) Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6(1): 39–67.

Kubozono, Haruo (1991) Modeling syntactic effects on downstep in Japanese. *Papers in Laboratory Phonology* II, 368–387. Cambridge: Cambridge University Press.

Kubozono, Haruo (1999) Mora and syllable. In: Natsuko Tsujimura (ed.) *The handbook of Japanese linguistics*, 31–61. Oxford: Blackwell.

Kubozono, Haruo (2007) Focus and intonation in Japanese: Does focus trigger pitch reset? In: Shinichiro Ishihara (ed.) *Interdisciplinary studies on information structure 9*, Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure (WPSI 2), 1–27. Potsdam: University of Potsdam.

Kuno, Susumu (1973) *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Masuoka, Takashi and Yukinori Takubo (1992) *Kiso nihongo bunpou (Kaitei-ban)* (Basic Japanese grammar (revised edition)). Tokyo: Kurosio Publishers.

Nagahara, Hiroyuki (1994) *Phonological phrasing in Japanese*. Doctoral dissertation, University of California, Los Angeles.

Pierrehumbert, Janet B. and Mary E. Beckman (1988) *Japanese tone structure*. Cambridge, MA: MIT Press.

Poser, William J. (1984) *The phonetics and phonology of tone and intonation in Japanese*. Unpublished doctoral dissertation, MA: MIT.

R Core Team (2014) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL http://www.R-project.org/.
Selkirk, Elisabeth and Koichi Tateishi (1991) Syntax and downstep in Japanese. In: Carol Georgopoulos and Roberta Ishihara (eds.) *Interdisciplinary approaches to language*, 519–543. Dordrecht: Kluwer Academic Publishers.
Xu, Yi (2013) ProsodyPro — A tool for large-scale systematic prosody analysis. In: *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013), Aix-en-Provence, France. 7–10.

日本語ダウンステップの再検討

——形態の影響——

ホワン ヒョンギョン[a]　　平山真奈美[b]
[a] 筑波大学／国立国語研究所 研究系 理論・対照研究領域 非常勤研究員［–2017.03］／
国立国語研究所 共同研究員
[b] 成蹊大学

要旨
日本語には，無アクセント語に後続する語の f0 よりアクセントのある語に後続する語の f0 が低くなる現象，いわゆるダウンステップがある。統語と音韻のインターフェースの観点から，Selkirk and Tateishi（1991）は，XP の左端が major phrase の左端に写像され，これによりダウンステップがブロックされると主張する。しかし Kubozono（1991）のデータはこの主張に反する。本研究では，従来の研究で特に取り沙汰されていなかったダウンステップにおける品詞の影響を検討した。
音声産出実験の結果，品詞が日本語のダウンステップの有無に影響する可能性があることが明らかになった。特に，名詞の連続する句でダウンステップが見られたが，形容詞のイ形が連続する句では見られなかった。先行研究の指摘同様，Selkirk and Tateishi（1991）の一般化には疑問が残る。ダウンステップが阻止される要因は，ある形態を使用するときに生まれる句全体の不自然さであると分析した。また，日本語のダウンステップ研究で形態をコントロールする必要があることが示唆された。

キーワード：品詞，音声産出，ダウンステップ，日本語