THE PHONETICS AND PHONOLOGY OF VOCATIVE INTONATION IN TOKYO JAPANESE

Haruo Kubozono and Ai Mizoguchi

NINJAL
kubozono@ninjal.ac.jp; amizoguchi@ninjal.ac.jp

ABSTRACT

This paper examines the phonetic and phonological structure of vocative intonation in Tokyo Japanese, of which very little is known in the literature. Based on our auditory observations and acoustic experiments, we report that Tokyo Japanese uses three distinct pitch patterns termed Patterns α, β, and γ, when calling people by their names or kinship terms. Pattern α is the neutral pattern used in a wide range of contexts and shows a markedly greater pitch range than the corresponding declarative sentence. Patterns β and γ, in contrast, are different from the declarative form in involving an additional H tone and HL tone, respectively, on the final syllable. Pattern β is specifically used when the speaker attempts to draw attention from someone in the distance. Pattern γ is used when the speaker shows affection to the hearer. None of these patterns triggers neutralization of the lexical accent distinctions in the language.

Keywords: vocative intonation, question intonation, Tokyo Japanese, accent neutralization, pitch accent

1. INTRODUCTION

Languages fall into two groups depending on whether they have a particular morphological or grammatical marker for the vocative case of people's names and kinship terms. The first group, represented by Latin, has a special vocative marker, as illustrated in (1a). Bulgarian, Korean and many other languages belong to the same group [8]. On the other hand, there are also many languages that do not have such an overt morphological marker. This group is represented by modern English, as shown in (1b).

(1) a. Latin
Brutus (nominative) vs.
Et tu, Brute (vocative) 'You, too Brutus'
b. English
Brutus (nom.) vs. You, too, Brutus. (vocative)

Present-day Japanese belongs to the second group and displays no overt grammatical marker for the vocative case. Because of this, one and the same sentence often becomes semantically ambiguous in Japanese, as illustrated in (2) and (3). The sentence in (2) is a declarative sentence in which the subject *obaachan* 'grandma' is ambiguous between the vocative and nominative case. The sentence in (3) is an interrogative sentence but is still ambiguous between the two cases.

(2) obaachan kita-yo 'Grandma came'
a. (He) is here, Grandma (vocative)
b. I (Grandma) am here. (nominative)

(3) obaachan genki? 'Grandma well'
a. Are you well, Grandma? (vocative)
b. Is Grandma well? (nominative)

This raises a question of how the vocative meaning is phonetically conveyed and is different from non-vocative meanings. Although potentially very important and interesting, this question has attracted little or no serious attention in Japanese phonetics and phonology, even in the prosodic study of standard Tokyo Japanese. With this background, this paper examines the phonetic and phonological structure of vocative intonation in Tokyo Japanese, in comparison with those of declarative and interrogative sentences.

Looking beyond Japanese, one finds a certain amount of work on the vocative prosody in other languages. One actually finds two phonetic parameters that are often used for vocative marking. One is vowel lengthening, which is observed in Southern Sierra Miwok [2] and Nivkh [7]; see [3] and [8] for an overview. A second phonetic feature is intonation, or postlexical pitch features, that is very often used in languages without specific vocative forms such as English [5] and German [4]; see also [1] for the vocative intonation in Polish.

2. BASIC PATTERNS

Our auditory observations of TV dramas and daily conversations for many hours revealed that native speakers of Tokyo Japanese use three distinctive pitch patterns when calling people such as their friends and family members. These three patterns are termed Patterns α, β, and γ in this study. In phonetic terms, these three patterns differ from each other as well as the declarative form (one-word statement) in the

following respects. First, Pattern α looks similar to the declarative pattern in an overall shape but is crucially different in exhibiting an enhanced pitch contour, i.e. an expanded pitch range (Figs. 1a vs. 1b). Pattern β (Fig. 1c), in contrast, is different from the declarative pattern (Fig. 1a) in involving an additional H(igh) tone on the final syllable. Pattern γ (Fig. 1d) is further different from Pattern β in exhibiting an additional pitch fall (HL) on the same syllable. Pattern β looks quite similar to the pitch contour of interrogative sentences in the same language [6].

Figure 1a: Accented name *Na'oya* (statement)

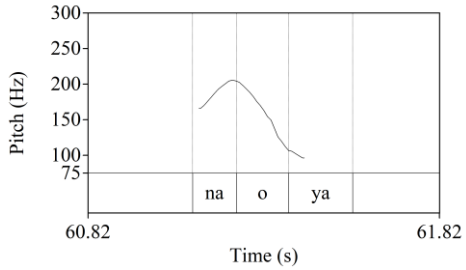


Figure 1b: Accented name *Na'oya* (vocative α)

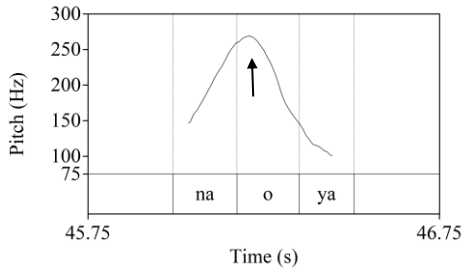


Figure 1c: Accented name *Na'oya* (vocative β)

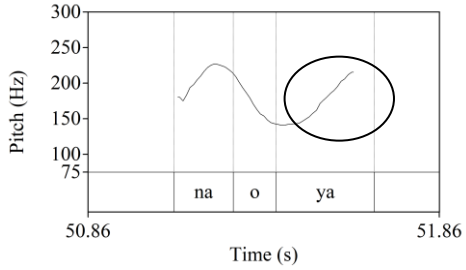
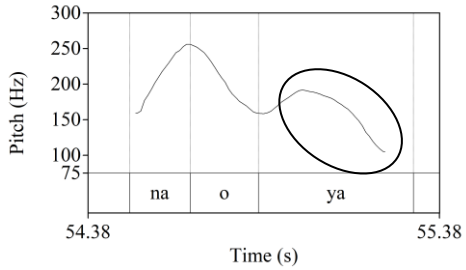


Figure 1d: Accented name *Na'oya* (vocative γ)



The unaccented counterparts of the accented names in Fig. 1 also exhibit corresponding pitch patterns for the vocative, as illustrated in Fig. 2a-d: Fig. 2a (declarative sentence), 2b (Pattern α), 2c (Pattern β), and 2d (Pattern γ). Again, Pattern β looks quite similar to the pitch contour of interrogative sentences in the same language [6].

Figure 2a: Unaccented name *Naomi* (statement)

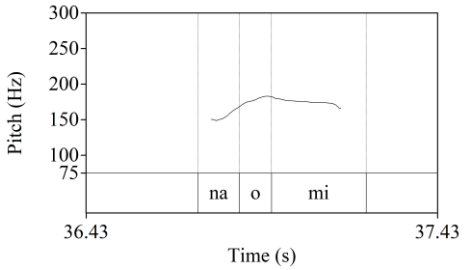


Figure 2b: Unaccented name *Naomi* (vocative α)

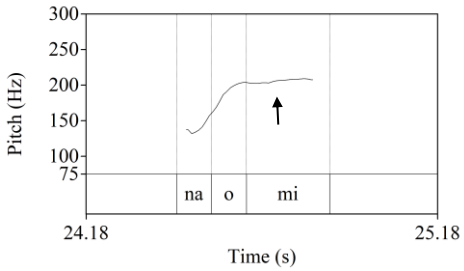


Figure 2c: Unaccented name *Naomi* (vocative β)

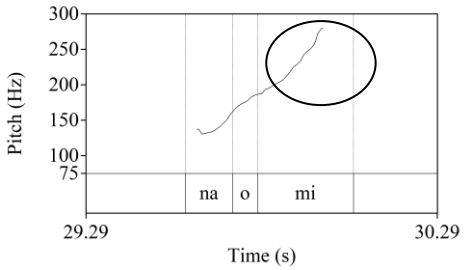
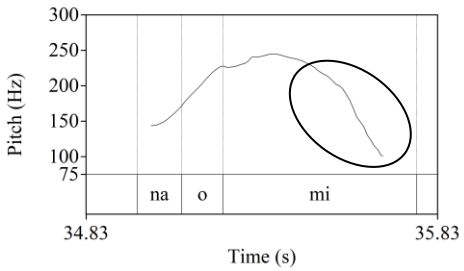


Figure 2d: Unaccented name *Naomi* (vocative γ)



A comparison between Figs. 1b-d and Figs. 2b-d clearly shows that the lexical pitch contrast between accented and unaccented nouns is well preserved. In other words, the two lexical patterns are not neutralized since accented words always exhibit an abrupt pitch fall, whereas their unaccented counterparts lack this phonetic feature.

3. PRODUCTION EXPERIMENT

The phonological observations mentioned above raise an interesting question of how the three patterns of vocative intonation are distinctively used by language users. In order to answer this question, we conducted a production experiment with two native speakers of Tokyo Japanese as subjects (one male speaker aged 32 and one female speaker aged 75). We prepared a set of people's names and similar nouns often used for calling such as *mama* 'Mom', *sensei* 'teacher', and *obaachan* 'Grandma'. These test words are balanced as much as possible with respect to the lexical accent classes (accented vs. unaccented), word length (two to five moras long), and the type of the word-final syllable, heavy (bimoraic) vs. light (monomoraic). Some examples are given in (4), where apostrophes indicate word accent, or the locus where an abrupt pitch fall occurs.

(4) test words
 a. accented names
  2 moras: Yu'u, Ke'n, Yu'mi, Ma'ma
  3 moras: Yu'uma, Ta'roo, Na'oya
  4 moras: sense'i 'teacher'
 b. unaccented names
  2 moras: Tada, kimi 'you'
  3 moras: Kyooda, Inoo, Naomi
  4 moras: obachan 'aunt'

We also considered various contexts in which the speaker is calling the name, as summarized in (5).

(5) a. simply drawing attention of a person
 b. calling someone nearby
 c. calling her from the distance
 d. confirming that she is present nearby
 e. showing concern or worry for her
 f. speaking to her after a long absence
 g. blaming her for what she has done
 h. asking her a favour
 i. mourning for her at her funeral

In the experiment, we presented each subject with the list of personal names including those in (4) and asked how she would call the person in each of the nine contexts in (5). This production experiment yielded the results in (6), with relatively little inter-speaker differences.

(6) a. Pattern α is used in a wide range of contexts for both word accent types, suggesting that this is the default or unmarked pattern of vocative intonation.
 b. Pattern β is used typically when the speaker wants to confirm the addressee's presence (5d) and when she is concerned about the addressee (5e).
 c. Pattern γ is typically used in contexts where the speaker is calling emotionally, such as in (5f-i), and where she is calling from a distance, i.e. (5c).

4. ACOUSTIC MEASUREMENTS AND DISCUSSION

In addition to the production experiment described above, we conducted an acoustic analysis to examine the phonetic differences between declarative and vocative intonation, on the one hand, and among the three patterns of vocative intonation, i.e. Patterns α, β, and γ, on the other.

This experiment involved two native speakers of Tokyo Japanese, one male (35 years old) and one female speaker (44 years old). We asked each subject to produce the three vocative patterns plus the corresponding declarative sentence with the pair of trimoraic names *Na'oyo* (accented) and *Naomi* (unaccented). In total, we asked them to pronounce each test word with each of the four pitch patterns ten times, which produced 160 tokens altogether (2 test words x 4 pitch patterns x 10 repetitions x 2 speakers).

In the analysis that followed, we measured (i) the pitch value at the onset of the word, (ii) the value of the first pitch peak, (iii) the value of the valley after the first peak, and (iv) the value of the second peak. The third value was obtained only from the accented word, while the fourth value was obtained from the vocative patterns β and γ of the accented word (Figs. 1c and 1d).

4.1. Accented words

We first compared the declarative (statement) pitch pattern with the vocative pattern α, i.e. Figs. 1a vs. 1b, with respect to the extent of pitch rise from the word onset (V1) to the first peak (P1) as well as the extent of pitch fall from the peak (P1) to the offset of the word (V2). This comparison revealed that the vocative pattern involves a much greater pitch rise than the statement. This difference was statistically significant for both speakers as shown in Table 1 (female speaker) and Table 2 (male speaker): the figures in the parentheses denote standard deviations.

Table 1: Comparison of statement and vocative α (Hz)

	P1-V1	P1-V2
Statement	36.46 (10.08)	89.20 (11.31)
Vocative α	116.75 (7.14)	92.13 (29.50)
t	$t(18)=19.79$	$t(18)=0.29$
p	<0.001*	0.774

Table 2: Comparison of statement and vocative α (Hz)

	P1-V1	P1-V2
Statement	17.69 (3.34)	45.65 (9.56)
Vocative α	69.09 (15.31)	49.93(16.05)
t	*t*(18)=7.10	*t*(18)=0.73
p	<0.001*	0.478

We then compared vocative patterns β and γ with respect to (i) the degree of pitch rise from the word onset to the first peak (P1-V1), (ii) the degree of pitch fall from the first peak to the following valley (P2-V2), and (iii) the degree of pitch rise from the valley to the second peak (P2-V2). Statistical results of this comparison are given in Table 3 (female speaker).

Table 3: Comparison of vocative β and γ (Hz)

	P1-V1	P1-V2	P2-V2
Vocative β	103.60 (26.35)	128.68 (33.15)	33.53 (29.53)
Vocative γ	227.72 (20.38)	148.25 (56.47)	-73.20 (33.14)
t	*t*(18)=11.78	*t*(18)=0.94	*t*(18)=7.60
p	<0.001*	0.357	<0.001*

The results in this table show that vocative patterns β and γ differ crucially from each other in the degree of the first pitch rise and that of the second rise: pattern γ exhibits a much steeper pitch rise than pattern β towards the first peak, while pattern β displays a steeper pitch rise towards the second peak.

4.2. Unaccented words

We compared the statement contour and the three vocative contours for the unaccented name *Naomi*, too (Figs. 2a-d). To understand the basic differences between these four types of contours, we examined the degree of pitch rise from the word onset to the peak (P1-V1). This comparison yielded the results in Table 4.

Table 4: Comparison of vocative β and γ (Hz)

	P1-V1 (male speaker)	P1-V1 (female speaker)
Statement	36.90(2.50)	62.51(25.38)
Vocative α	40.43(5.41)	74.91(15.84)
Vocative β	57.87(7.81)	86.53(21.49)
Vocative γ	92.74(11.24)	183.26(18.05)

The extent of pitch rise is smallest in the statement and largest in vocative pattern γ, with vocative patterns α and β coming between these two in this order. The results of T-test are given in the same table: * indicates that the difference between the two respective patterns is statistically significant.

5. COMPARISON WITH OTHER CASES AND SENTENCE TYPES

In the foregoing discussion, we have hinted that vocative intonation is sometimes similar to the phonetic forms of other types of cases and sentences. This is true only in some restricted contexts. In the declarative sentence in (2), for example, the vocative case in (2a) can be clearly differentiated from the nominative case in (2b) if the intonational patterns β and γ are used for the vocative case: the two meanings in (2a,b) are differentiated between the patterns in Fig. 1c/d (vocative) and the pattern in Fig. 1a (statement).

However, the prosodic difference between the two cases becomes subtle if the vocative form in (2a) is pronounced with the prosodic pattern in Fig. 1b. The phonetic difference becomes especially subtle if the nominative form in (2b) receives some semantic emphasis/focus and, consequently, is pronounced with an expanded pitch range as in Fig. 1b.

Basically the same observation can be made of the interrogative sentence in (3), which is also ambiguous between the vocative and nominative cases.

6. CONCLUSIONS

This study examined vocative (calling) intonation in Tokyo Japanese in comparison with the intonation of declarative sentences. It has revealed that this language uses three distinct pitch patterns for the vocative meaning: Pattern α simply enhances the pitch (range) of the declarative intonation, Pattern β involves an extra high pitch on the final syllable, and Pattern γ introduces a new pitch fall in the same syllable. Acoustic measurements showed that these three melodies are distinct from each other in phonetic terms, too. Native speakers choose from these three melodies in different pragmatic contexts although they sometimes use two or three melodies in one and the same pragmatic context. Finally, the lexical pitch contrast between accented and unaccented words is not lost in any context.

7. ACKNOWLEDGEMENTS

The work reported here was supported by the JSPS KAKENHI Grant Numbers 26244022, 16H06319 and 17K18502 as well as the NINJAL collaborative research project 'Cross-linguistic Studies of Japanese Prosody and Grammar'.

8. REFERENCES

[1] Arvaniti, A., Żygis, M., Jaskuła, M. 2016. The phonetics and phonology of the Polish calling melodies. *Phonetica*, 73, 338–361.

[2] Broadbent, S. M. 1964. *The Southern Sierra Miwok Language*. (University of California Publications in Linguistics, 38.) Berkeley: University of California Press.

[3] Daniel, M., Spencer, A. 2008. The vocative – An outlier case. In A. L. Malchukov and A. Spencer (eds.), *The Oxford Handbook of Case*, 626–634. Oxford: Oxford University Press.

[4] Grice, M., Baumann, S., Benzmüller, R. 2005. German intonation in autosegmental-metrical phonology. In S-A. Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 55–83.

[5] Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

[6] Kubozono, H. 2018. Focus prosody in Kagoshima Japanese. To appear in R. Goedemans, J. Heinz, H. van der Hulst (eds.), *The Study of Word Stress and Accent: Theories, Methods and Data*. 323-345. Cambridge: Cambridge University Press.

[7] Panfilov, V. Z. 1962. *Grammatika Nivxskogo Jazyka*. [The Grammar of the Nivkh Language]. Moscow: Nauka.

[8] Sonnenhauser, B., Noel A.H., P. (eds.) (2013) *Vocative! Addressing between System and Performance*. Berlin: Mouton de Gruyter.