F0 perturbation in a "pitch-accent" language

Jiayin Gao[1,2,3], *Takayuki Arai*[1]

[1]Sophia University, Japan
[2]Japan Society for the Promotion of Science, Japan
[3]Lacito-LPP-LabEx EFL, France

jiayin.gao@gmail.com, arai@sophia.ac.jp

Abstract

This study investigates the relation between consonant voicing and F0 in modern Tokyo Japanese, as produced by young female speakers. In a tone language, the F0 perturbation related to onset voicing has been reported to be inhibited, so that F0 can be maximally used in tonal contrasts. According to this explanation, the same pattern should be found in Tokyo Japanese, as F0 should be maximally used to signal its "pitch-accent".

Contrary to this prediction, our data show that in Tokyo Japanese, for the initial mora, F0 is remarkably lower after voiced than voiceless stops, and this effect lasts till the final part of the mora. However, the F0 level of the mora endpoint is maintained at H or L so that the pitch-accent pattern is well preserved. We thus argue that the competing role of F0 in a pitch-accent language, or even a tone language does not necessarily impose limitations on the F0 perturbation effect.

We also found that voiced stops are very often phonetically voiceless in utterance-initial position, while being phonetically voiced in utterance-medial position. Therefore, we question whether Tokyo Japanese is undergoing an incipient tonogenesis, given that the VOT cue is giving place to an F0 cue.

Index Terms: Japanese, microprosody, F0 perturbation, pitch-accent, voicing, VOT, tonogenesis

1. Introduction

1.1. Consonant voicing and F0 perturbation

The consonant-F0 interaction is a well-documented phenomenon. In particular, it has been attested in a great number of languages that the F0 of the vowel onset is higher when following a voiceless sound than a voiced sound, hereafter the "F0 perturbation" effect (for a review, see [1, 2]). For convenience, the "voiceless-voiced" contrast here is used in a broad sense as in [3]. It embraces at least two types: aspiration or [spread glottis] contrast such as in English or German (see instead [4] and [1] for their proposal of a [stiff] feature), and true voicing or [voice] contrast such as in French or Spanish. [Spread glottis] and [voice] are considered by some as different phonetic implementations or phonetic features of one phonological feature [voice] [3], and by others as different phonological features [5]. Moreover, there has been some debate about whether these features are privative or binary.

In the above-mentioned languages, the voicing-induced F0 perturbation generally disappears no later than the vowel midpoint. In some tone languages, it has been found that F0 perturbation is more limited in duration (e.g., immediately after the stop release in Thai [6], only at the vowel onset in Yoruba [7]). The widely accepted explanation is that F0 perturbation is actively inhibited due to the competing role of F0 in signaling tonal contrasts, "so that the different tones will be maximally distinct perceptually." [7]

1.2. Tokyo Japanese

This study focuses on Tokyo Japanese. Japanese is traditionally described as having a voicing contrast. However, unlike a true voicing language, voiceless stops in Japanese are slightly aspirated in word-initial position [8]. Furthermore, recent studies showed a trend toward devoicing word-initial voiced stops in Tokyo Japanese [9]. The first goal of this study is to revisit the phonetic realization of the voicing contrast in Tokyo Japanese stops in terms of positive or negative VOT, and examine its variations according to the position in an utterance.

That Tokyo Japanese is a tone language or not is controversial. It is traditionally viewed as a "pitch-accent" language. However, given that pitch is necessary for determining the meaning of a word in a number of tonal minimal pairs, it may also be viewed as a restricted tone language (see [10]). On a sensible view, languages can be classified on a continuum ranging from dense tone languages (e.g., Chinese) to restricted tone languages (e.g., Swedish, Limburgian, Tokyo Japanese), depending on the functional load of tone. For the sake of simplicity, we stick here to the "pitch accent" label. This allows, quite simply, to describe a word or phrase (1) as accentless: pitch is gradually raised over the entire sequence, that is, the initial mora carries lower pitch than the following moras; (2) as carrying a non-initial accent: pitch is raised till a peak on the accented mora, drops abruptly on the following mora and remains low; (3) as carrying an initial accent: pitch starts high on the first mora, which is accented, drops abruptly on the following mora and remains low. For example, /ha.si.ga/ (*ga* is a particle for subject) means 'edge' if it carries LHH melody (accentless), 'bridge' if it carries LHL melody (accent on second mora), or 'chopsticks' if it carries HLL melody (accent on first mora, or initial accent). The second goal of our study is to examine the impact of the phonemic role of F0 in Tokyo Japanese on voicing-induced F0 perturbation.

Our working hypothesis is as follows: if F0 perturbation is inhibited by the competing use of F0 for tone in tone languages, similarly, it should also be limited in a pitch-accent language such as Tokyo Japanese.

2. Experiment

The experiment was designed to assess the realization of the voicing contrast in Tokyo Japanese, in word-initial and word-medial stops. In this paper, we focus on the stop VOT and F0 in the following moraic vowel.

2.1. Method

2.1.1. Participants

We report results from 9 female speakers aged 23 in average (20 to 27), all born and raised in Tokyo area except one who lived in Hiroshima before entering elementary school in Tokyo. They were recruited from Sophia University and received a prepaid gift card for their time. One speaker judged her English level as good, while all the others reported a basic or intermediate English level. Analyses of male speakers are in progress.

2.1.2. Speech material

We report results from a selection of 37 (near-) minimal pairs with a voicing contrast for labial, dental and velar stop onsets, plus 4 words with /m/ onset as a reference, making a total of 78 words, with target onsets in word-initial or word-medial positions. They were elicited in isolation in the first session, and in the carrier sentence "sore o _ to iu" ('(I) call it _') in the second session. Each word contains 2 to 3 syllables, and 2 to 5 moras. Within each word pair, words of comparable frequency of usage were chosen.

Two pitch-accent contexts were used: (1) accentless (38 words), that is, low pitch on the first mora and high pitch on the following moras, and (2) initial accent (40 words), that is, high pitch on the first mora and low pitch on the following moras. The target mora (that with the target onset) may thus carry a low pitch (pitch-accent context "L"), or a high pitch (pitch-accent context "H"). Each utterance was repeated twice, making a total of 312 utterances (78 words*2 sessions*2 repetitions) for each speaker.

This experiment design also gave the possibility of studying variations in three prosodic positions. Word-initial onsets produced in isolated words (session 1) are defined as "utterance-initial (UI, 64 words)"; word-initial onsets produced in the carrier sentence (session 2) are defined as "utterance-medial & word initial (UM, 64 words)"; word-medial onsets, either in isolation or in the carrier sentence are simply defined as "word-medial (WM, 28 words)". Both UI and UM positions contain 31 accentless and 33 initial-accent words, and WM position contains 7 accentless and 7 initial-accent words repeated in the two sessions. An example is given in Table 1.

Table 1: *Example of target onsets /t-d/ in 3 prosodic positions.*

Utterance-initial (UI)	**Utterance-medial Word-initial (UM)**	**Word-medial (WM)**
taikeː (form)		ʥi**t**ai (font)
	sore o **t**aikeː to iu	sore o ʥi**t**ai to iu
daikeː (trapezoid)		ʥi**d**ai (epoch)
	sore o **d**aikeː to iu	sore o ʥi**d**ai to iu

2.1.3. Procedures and analyses

Each speaker was recorded individually in a sound-proof room, reading utterances presented on a laptop screen in a different random order. Instructions were given in Japanese. Sino-Japanese words were presented in *kanji* (Chinese characters) with *katakana* (Japanese writing) in parentheses.

For stop onsets, we measured their VOTs. For F0, the vowel of the mora beginning with the target onset consonant was divided into 50 equal time intervals, and the mean F0 of each time interval was measured. Raw F0 values were then normalized using a within-speaker z-score transformation, so that between-speaker variations were minimized. We used SS-ANOVA (smoothing-spline ANOVA) models [11] to compute the averaged F0 curves and their 95% bayesian confidence levels. This method has been used to compare tongue curves from ultrasound imaging [12] as well as F0 curves [13]. If two curves do not overlap with each other at a given time point, they are significantly different at this time point.

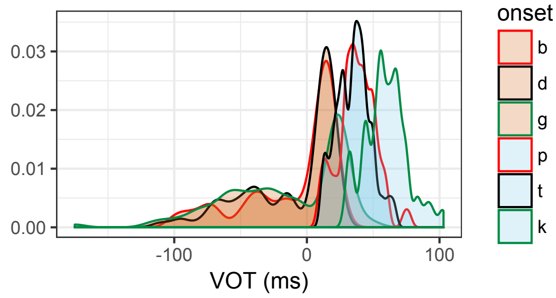

Figure 1: *Kernel density estimation of VOT distributions of utterance-initial stops.*

Moreover, words that were produced with a deviant pitch-accent due to idiosyncratic realizations, based on perceptual judgment and visualization of the pitch contours, were excluded from all analyses.

3. Results

In this section, we will briefly report the VOT data and focus on the F0 data.

3.1. VOT

As shown in Table 2, voiceless stops have medium-lag VOTs word-initially, in both UI and UM positions, but short-lag VOTs word-medially. Moreover, VOTs are in average slightly longer in UI than UM, and in L than H pitch-accent contexts.

Table 2: *Mean (standard deviation) VOTs in ms according to place of articulation, prosodic positions and pitch accent.*

	UI		**UM**		**WM**	
	L	H	L	H	L	H
/p/	41(16)	35(14)	37(16)	32(12)	15(6)	15(6)
/t/	37(15)	34(12)	29(13)	28(11)	15(3)	13(2)
/k/	61(17)	57(16)	51(13)	50(12)	25(7)	20(7)

Concerning voiced stops, 67% of them are devoiced in UI position. In this position, only one of the nine speakers produced prevoiced stops more than 60% of the time. More than half of the speakers produced them as short-lag more than 75% of the time. As the prosodic hierarchy lowers, the devoiced proportion decreases (24% in UM, 1% in WM).

Figure 1 further shows the VOT distributions for utterance-initial stops. VOTs of the short-lag /b,d,g/ and medium-lag /p,t,k/ overlap, but differ significantly, $\chi^2(1) = 54.0$, $p < .0001$, estimated by linear mixed-effect models (VOT∼PhonologicalVoicing, random=speaker+item).

3.2. F0

Since voiced stops are variably realized with or without phonetic voicing, they will be referred to as "phonologically voiced" or /b,d,g/. They will be referred to as "devoiced" when they are phonetically voiceless, and "prevoiced" when they are phonetically voiced. /p,t,k/ refers to phonologically voiceless

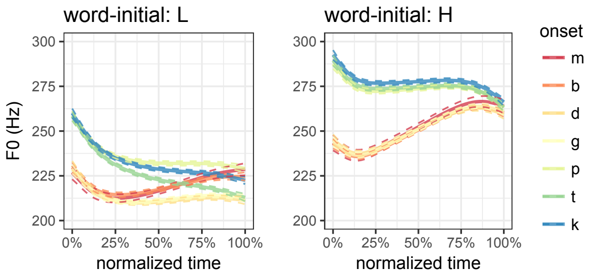

Figure 2: *F0 curves over the moraic vowel following word-initial onsets (in both UI and UM) computed by SS-ANOVA.*

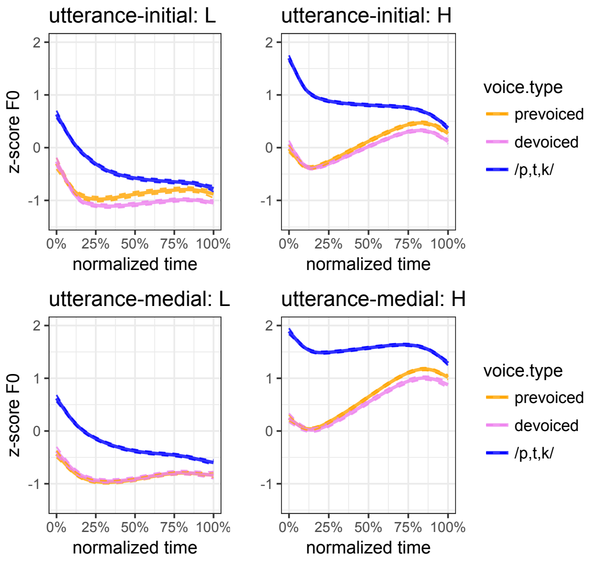

Figure 3: *Normalized F0 curves over the vowel following word-initial voiced vs. voiceless stops computed by SS-ANOVA.*

stops. The term "voicing (contrast)", without specification, will refer to "phonological voicing".

The F0 pattern varies according to the prosodic position. Word-initial voicing has a large effect on F0 in the following mora. Figure 2 shows the raw F0 curves computed by SS-ANOVA [11] over the moraic vowel following each onset in word-initial position, including utterance-initial and -medial positions. The F0 contour is globally falling after /p,t,k/ onsets and rising after /b,d,g/ as well as /m/ onsets. A pitch shift of about 40 Hz in H context, and of about 25 Hz in L context can be observed at vowel onset between the two series. The F0 difference diminishes gradually but is maintained until about 3/4 of the vowel. Statistical comparisons will be made on the normalized F0 data in the following.

In order to assess the effect of phonetic voicing vs. phonological voicing, Figure 3 shows the normalized F0 curves using z-score transformation over the moraic vowel following /p,t,k/ and /b,d,g/ pooled over place of articulation, the latter further separated into "devoiced" and "prevoiced", in word-initial position (UI & UM). The F0 curves plotted by SS-ANOVA models indicate that F0 is significantly higher after /p,t,k/ than after /b,d,g/ nearly until the end of the vowel. (The mean vowel duration in the initial mora is 107 ms.) Prevoiced /b,d,g/ exhibits slightly higher F0 contours than devoiced /b,d,g/ in the latter

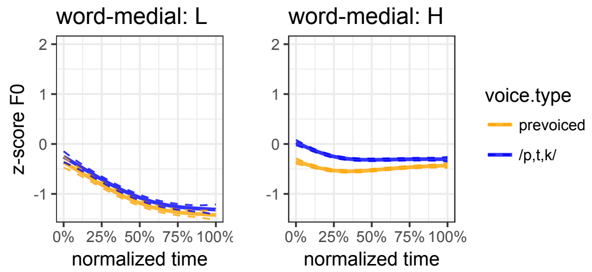

Figure 4: *Normalized F0 curves over the vowel following word-medial voiced vs. voiceless stops computed by SS-ANOVA.*

part of the vowel, except in utterance-medial L pitch-accent context. This might suggest a slight tradeoff between low pitch and phonetic voicing. Furthermore, the F0 perturbation effect is larger in H than L pitch-accent contexts. This corroborates reports on other languages (Engish, [1]; French and Italian, [2]).

In word-medial position, the pattern is very different. Figure 4 shows the z-score F0 curves over the moraic vowel following word-medial voiceless vs. voiced stops. Recall that the latter are phonetically voiced. Onset voicing, both phonological and phonetic, has little effect on F0. If a slight F0 perturbation is observed in H context, the effect seems negligible in L context.

4. Discussion

4.1. True voicing or aspirating

Phonological arguments, notably different rules involved in the voicing in compounds, or *Rendaku*, have been put forward in favor of an active [voice] feature in Japanese [14, 15]. However, our phonetic evidence show that modern Tokyo Japanese differs from both "true voicing" and "aspiration" types. The /p,t,k/ series are less aspirated than in an aspirating language, and the /b,d,g/ series are less consistently voiced than in a true voicing language. As a comparison, utterance-initial voiced stops in French have consistiently negative VOTs [16]. (Some gender difference has been reported by [9] and observed from our preliminary results on male speakers, which we plan to report in the near future.) In line with [17] and [18], this might suggest that phonologically active features do not necessarily match phonetic properties.

From an evolutionary point of view, [9]'s cross-generational comparison suggests a trend towards devoicing /b,d,g/ in Tokyo Japanese. Our young female speakers show consistent results, suggesting that the VOT cue is gradually losing its weight, although this has not lead to a complete merger between the two stop series in terms of VOT.

4.2. F0 perturbation, pitch-accent, and tone

Similar to languages with a "voiceless-voiced" contrast, as defined in 1.1, Tokyo Japanese exhibits higher F0 after voiceless than voiced sounds, but only in word-initial position. Contrary to our predictions, in this position, the F0 perturbation is neither limited in magnitude nor in duration.

In word-initial position, despite this great F0 perturbation effect, we observe that the use of F0 in signaling pitch-accent is not compromised, so that the pitch-accent pattern is still well preserved. At the onset of the word-initial mora, the F0 height is at about the same level after /b,d,g/ stops for H context and after /p,t,k/ stops for L context. However, at the endpoint of the

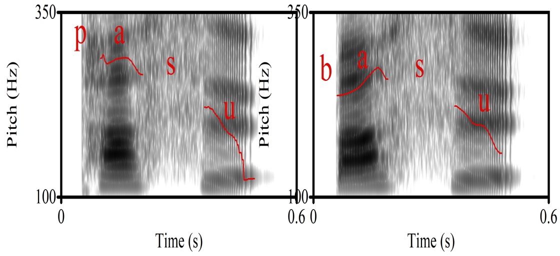

Figure 5: *Illustration of F0 curves of /pasu/ vs. /basu/, both carrying an initial accent.*

word-initial mora, F0 curves converge at a higher F0 level for H context, and at a lower F0 level for L context. This suggests that for the initial mora, the F0 height at the mora offset has more importance than at the mora onset in signaling H or L. As shown by the minimal pair in Figure 5, the initial mora /ba/ (with a devoiced /b/) exhibits much lower F0 than /pa/ at the mora onset, but at the mora offset, the F0 difference is very weak. Importantly, the pitch-accent pattern, that is, initial H relative to the following L in this case, remains unaffected.

In word-medial position, voicing has little effect, if any, on the following F0. This may be due to one or both of the following reasons: (1) the phonetic voicing of /b,d,g/ is robustly present in this position, making F0 a somewhat unnecessary cue to voicing; (2) the entire F0 contour of non-initial moras may be important in signaling the relative pitch height with the previous mora and the following mora.

The question about F0 perturbation in a pitch-accent language has been raised by [19] in their study of Southern Kyungsang Korean, which has a Japanese-like pitch-accent system [20]. Their study showed an F0 perturbation effect at the onset as well as the midpoint of the following vowel, without providing further measurements after the vowel midpoint.

The evidence from Tokyo Japanese and Kyungsang Korean shows that the competing use of F0 in a phonemic contrast like pitch-accent does not necessarily impose limitations on the use of F0 in signaling an originally consonantal contrast. Similarly, the use of duration as one of the major cues signaling vowel length contrast in English does not limit its use in signaling a voicing contrast of the following obstruent. F0, being not only scalar, but also a long time-domain property, can be manipulated in multiple ways with respect to its absolute height, relative height, peak, contour, among others, as well as the timing of these parameters. Therefore, F0 does not necessarily participate in the *maximization*, but in the *sufficient distinction* of a tonal or pitch-accent contrast. This may be especially true for Japanese pitch-accent, as it is characterized by a two-way pitch height distinction, allowing more space for F0 perturbation. An alternative explaination, as pointed out by one anonymous reviewer, might be the low functional load of pitch accent in Tokyo Japanese, as shown by a low percentage of words distinguished only by their pitch accent [21] and a much weaker role of pitch accent than segments in word recognition [22, 23].

It is, however, worth mentioning that even in a dense tone language with a high functional load of tone, F0 perturbation is not necessarily inhibited. Otherwise, Middle Chinese and Vietnamese would not have undergone “tone split”, during which a high vs. low tone contrast has replaced the voicing contrast, each of the previous tones split into a high and low tone [24].

4.3. Another case of tonogenesis?

The differential use of cues to voicing depending on the position in a word is observed cross-linguistically. The Tokyo Japanese pattern is similar to Kyungsang Korean (although we have no knowledge about the word-medial position), as we mentioned above, but also to Seoul Korean and Shanghai Chinese, in that the primary cue is F0 height in word-initial position and VOT in word-medial position. The notable difference between the three languages is that stops have a two-way laryngeal contrast in Tokyo Japanese instead of a three-way laryngeal contrast in the other two languages.

One stop series, which we refer to here as “b,d,g”, is realized as phonetically voiced in word-medial position in all three languages. It is worth noting that it has been analyzed with different phonological features in each language. Shanghai Chinese is traditionally analyzed as a tone language, including a tonal height contrast co-occurring with a voicing contrast. “b,d,g” are analyzed as [+voice] in this approach [25]. Other features including [+slack] [26] and [+murmur] [27] have also been proposed. Seoul Korean, which lost its pitch-accent very long time ago, has been recently analyzed as a language undergoing tonogenesis (e.g., [28, 29]). Contrary to Shanghai Chinese and Japanese, the “b,d,g” series of Seoul Korean have long been analyzed as phonologically voiceless, while features related to the tensity like [tense] are often used to distinguish the two unaspirated series [30, 31]. As for Tokyo Japanese, as we summarized in 1.2, it has long been analyzed as having a [voice] contrast and a pitch-accent system. To our knowledge, no proposal has ever been made about a tonogenetic or “tone split” development in any Japanese variety.

Nonetheless, modern Tokyo Japanese does share some similarities with languages reported to be undergoing (incipient) tonogenesis or “tone split”, such as Seoul Korean, as we mentioned above, Dutch [32], Afrikaans [33], and Tamang [34]. Whether the VOT cue will eventually be replaced with an F0 cue in Tokyo Japanese needs more investigations in the years to come. However, the above-mentioned languages present a challenge to phonological feature theories with a purely static and synchronic analysis. We believe a dynamic and evolutionary perspective is needed to account for each transitional stage towards a potential tonal system.

5. Conclusions

Our study on Tokyo Japanese showed that F0 perturbation is not inhibited in a Japanese-like pitch-accent language. Meanwhile, we showed that the VOT cue is diminishing. Further investigations are needed to determine whether it is gradually replaced with an F0 cue. We believe that, following [35], studying the recurrent change from a voicing contrast to a tonal contrast may give us crucial insights on the understanding of the complex laryngeal features at both segmental and suprasegmental levels.

6. Acknowledgements

We wish to express our gratitude to Mafuyu Kitahara for his help in the selection of the speech material, and to Martine Mazaudon and Pierre Hallé for their comments on an earlier draft of this paper. This ongoing project is supported by JSPS KAKENHI Grant Number 17F17006.

7. References

[1] H. M. Hanson, "Effects of obstruent consonants on fundamental frequency at vowel onset in English," *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 425–441, 2009.

[2] J. Kirby and D. R. Ladd, "Effects of obstruent voicing on vowel F0: evidence from "true voicing" languages," *Journal of the Acoustic Society of America*, vol. 140, no. 4, pp. 2400–2411, 2016.

[3] P. A. Keating, "Phonetic and Phonological Representation of Stop Consonant Voicing," *Language*, vol. 60, no. 2, p. 286, 1984.

[4] M. Halle and K. N. Stevens, "A note on laryngeal features," *MIT Research Laboratory of Electronics Quarterly Research Report*, vol. 101, pp. 198–213, 1971.

[5] G. Iverson and J. Salmons, "Aspiration and laryngeal representations in Germanic," *Phonology*, vol. 12, no. 3, pp. 369–396, 1995.

[6] J. Gandour, "Consonant types and tone in Siamese," *Journal of Phonetics*, vol. 2, pp. 337–350, 1974.

[7] J.-M. Hombert, J. J. Ohala, and W. G. Ewan, "Phonetic Explanations for the Development of Tones," *Language*, vol. 55, no. 1, pp. 37–58, 1979.

[8] T. J. Riney, N. Takagi, K. Ota, and Y. Uchida, "The intermediate degree of VOT in Japanese initial voiceless stops," *Journal of Phonetics*, vol. 35, no. 3, pp. 439–443, 2007.

[9] M. Takada, E. J. Kong, K. Yoneyama, and M. E. Beckman, "Loss of prevoicing in modern Japanese /g , d , b/," in *Proceedings of the 18th International Congress of Phonetic Sciences*, 2015, pp. 1–5.

[10] L. M. Hyman, "How (not) to do phonological typology: the case of pitch-accent," *Language Sciences*, vol. 31, no. 2-3, pp. 213–238, 2009.

[11] C. Gu, "Smoothing Spline ANOVA Models: R Package gss," *Journal of Statistical Software*, vol. 58, no. 5, 2014.

[12] L. Davidson, "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance," *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 407–415, 2006.

[13] C.-T. Chuang, Y.-C. Chang, and F.-F. Hsieh, "Complete and not-so-complet tonal neutralization in Penang Hokkien," in *Proceedings of ICPLC*, 2013, pp. 54–57.

[14] J. Itô and A. Mester, "The Phonology of Voicing in Japanese: Theoretical Consequences for Morphological Accessibility," *Linguistic Inquiry*, vol. 17, no. 1, pp. 49–73, 1986.

[15] K. Nasukawa, "The representation of laryngeal-source contrasts in Japanese," in *Voicing in Japanese*, J. van de Weijer, K. Nanjo, and T. Nishihara, Eds. De Gruyter Mouton, 2005, pp. 71–87.

[16] A.-P. Benguerel, H. Hirose, M. Sawashima, and T. Ushijima, "Laryngeal Control in French Stop Production: a Fiberscopic, Acoustic and Electromyographic Study," *Folia Phoniatrica et Logopaedica*, vol. 30, no. 3, pp. 175–198, 1978.

[17] J. Mielke, *The Emergence of Distinctive Features*. Oxford University Press, 2008.

[18] M. Brunelle, "Tone perception in Northern and Southern Vietnamese," *Journal of Phonetics*, vol. 37, no. 1, pp. 79–96, 2009.

[19] M. Kenstowicz and C. Park, "Laryngeal Features and Tone in Kyungsang Korean: a Phonetic Study," *Studies in Phonetics, Phonology and Morphology*, vol. 12, no. 2, pp. 247–264, 2006.

[20] D. Lee and S. Davis, "On the Pitch-Accent System of South Kyungsang Korean: A Phonological Perspective," *Language Research*, vol. 45, no. 1, pp. 3–22, 2009.

[21] T. Shibata and R. Shibata, "Akusento wa doo'ongo o donoteido benbetsu shiuruno ka?: Nihongo, eigo, chuugokugo no baai [Can lexical accent distinguish homonyms in Japanese, English, and Chinese?]," *Keiryoo Kokugo-gaku [Mathematical Linguistics]*, vol. 17, pp. 317–237, 1990.

[22] H. W. Yamamoto and E. Haryu, "The role of pitch pattern in Japanese 24-month-olds' word recognition," *Journal of Memory and Language*, vol. 99, pp. 90–98, 2018.

[23] K. Tamaoka, N. Saito, S. Kiyama, K. Timmer, and R. G. Verdonschot, "Is pitch accent necessary for comprehension by native Japanese speakers? - An ERP investigation," *Journal of Neurolinguistics*, vol. 27, no. 1, pp. 31–40, 2014.

[24] A. G. Haudricourt, "Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient," *Bulletin de la Société de Linguistique de Paris*, vol. 56, no. 1, pp. 163–180, 1961.

[25] J. Gao and P. Hallé, "Phonetic and phonological properties of tones in Shanghai Chinese," *Cahiers de Linguistique Asie Orientale*, vol. 46, no. 1, pp. 1–31, 2017.

[26] Y. Chen, "How does phonology guide phonetics in segment-f0 interaction?" *Journal of Phonetics*, vol. 39, no. 4, pp. 612–625, 2011.

[27] M. Yip, "Tonal register in East Asian languages," in *The phonology of tone: The representation of tonal register*, H. van der Hulst and K. L. Snider, Eds. De Gruyter Mouton, 1993, pp. 245–268.

[28] D. J. Silva, "Acoustic evidence for the emergence of tonal contrast in contemporary Korean," in *Phonology*, vol. 23, no. 2, 2006, pp. 287–308.

[29] Y. Kang, "Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study," *Journal of Phonetics*, vol. 45, no. 1, pp. 76–90, 2014.

[30] C.-W. Kim, "On the Autonomy of the Tensity Feature in Stop Classification (with Special Reference to Korean Stops)," *Word*, vol. 21, no. 3, pp. 339–359, 1965.

[31] H. Kim, S. Maeda, K. Honda, and L. Crevier-Buchman, "The Mechanism and Representation of Korean Three-Way Phonation Contrast: External Photoglottography, Intra-Oral Air Pressure, Airflow, and Acoustic Data," *Phonetica*, vol. 75, pp. 57–84, 2018.

[32] A.-F. Pinget, "The actuation of sound change," Ph.D. dissertation, Utrecht University, 2015.

[33] A. W. Coetzee, P. Speeter, K. Shedden, W. Styler, and D. Wissing, "Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis," *Journal of Phonetics*, vol. 66, pp. 185–216, 2018.

[34] M. Mazaudon and A. Michaud, "Tonal contrasts and initial consonants: a case study of Tamang, a 'missing link' in tonogenesis," 2008.

[35] J. Blevins, *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press, 2007.