The role of L in the pitch accent system of Tokyo Japanese*

Phillip Backley        Kuniya Nasukawa
Tohoku Gakuin University          Tohoku Gakuin University

ABSTRACT. This paper demonstrates how L contributes in two different ways to the pitch pattern of nouns in Tokyo Japanese. First, L functions as a prosodic marker, signalling the left edge of prosodic word domains. And second, L (rather than H) functions as the marked tonal property. It is argued that a lexical accent is cued by a fall in pitch, represented as a lexical L on the accented mora preceded by an unspecified (default H tone) mora. At the segmental level, L also represents the voicing contrast in Japanese obstruents. The result is a characterisation of Japanese as a language in which L dominates throughout the phonology.

Keywords: Japanese, pitch accent, prosodic marker, default tone, tone dissimilation, obstruent voicing

1. Introduction

To represent tonal systems, intonation systems and pitch accent systems, the custom is to use two structural units, H(igh) and L(ow). These units capture (i) the overall shape of a word’s tonal pattern by referring to sequences of H and L, and (ii) the internal structure of individual (contour) tones by referring to how H and L combine. Because H and L exist as indivisible properties, they presumably constitute basic structural units in representations.

In some early feature systems, tonal properties were expressed as features, and, like all other features, they were taken to be segmental (Wang 1967). Since then, however, the consensus has been to view H and L as suprasegmental properties residing on independent tonal tiers (Leben 1971). Some tonal and pitch accent systems are analysed as having just active H, with L functioning as a default tone realised in the absence of H, while other systems involve active L (with H as the default). Yet others call upon both active H and active L. Our focus in this paper is the pitch accent system of Tokyo Japanese (TJ, hereafter), and in particular, the pitch patterns found in simplex Yamato Japanese nouns. We will make the unorthodox proposal that pitch accent in TJ requires L to be active, rather than H (Yoshida 1999) or HL (Haraguchi 1991). Furthermore, we claim that L functions in two different ways in TJ. Firstly, it operates as a prosodic boundary marker, its function being to signal the left edge of a prosodic word domain. Secondly, L serves as a lexical property which identifies the location of a noun’s lexical accent. These two functions of L mainly operate independently of each other, but in certain contexts to be described below they are seen to interact in an interesting way.

Accounting for the TJ pitch accent system solely in terms of L helps to characterise the phonology of TJ as a system in which L predominates. In §2 we describe the behaviour of L as a prosodic marker, and then in §3 we argue for L as the active pitch accent property in TJ. Finally, in §4 we briefly discuss an additional function of L in Japanese – namely, as the laryngeal property responsible for the voicing distinction in obstruents.
2. L as a prosodic marker

2.1. Expressing prosodic information

For some time it has been acknowledged that phonological information is not just segmental: for communication to succeed, listeners need access to prosodic information in addition to segmental information, where this prosodic information refers to the location of prosodic domains. Of course, words are mainly identified from their segments, but there is now strong evidence that a word’s prosodic structure also contributes to its perceptibility, i.e. knowing where prosodic domains begin and end can help listeners to recognise words easily and quickly. For example, Cutler and Norris (1988) show how listeners are able to parse and process continuous speech more efficiently if they know the location of prosodic domains such as the syllable, foot and word. To put this claim into perspective, we need only consider the contrary situation in which no prosodic information is available to hearers. Given such a scenario, hearers would have difficulty locating the edges of prosodic domains, meaning that they would not be able to break up running speech into individual words, and as a result, would fail to recover the linguistic message with any degree of efficiency.

The idea that phonological information is both prosodic as well as segmental raises two questions. The first is how hearers identify the location of prosodic boundaries in running speech. In recent work (Backley and Nasukawa 2009, Nasukawa and Backley 2012) we have shown how languages express prosodic information directly in segmental structure, making this information perceptible to hearers. In essence, a word’s segmental structure can reveal something about that word’s prosodic structure because certain types of segment – or, more precisely, certain properties within segments – are attracted to strong prosodic positions while other segment types naturally occur in weak positions. We assume that native-speaker knowledge includes knowing which segmental properties serve as cues to strong positions in any given language. The ability to distinguish between strong and weak positions is crucial for language processing because strong positions regularly mark the boundaries (typically, the left boundary) of prosodic domains. The importance of this information is underlined in psycholinguistics work by Jusczyk and others (e.g. Jusczyk 1999; Jusczyk et al. 1993), where it is claimed that, when processing speech, listeners focus instinctively on strong segmental properties. In other words, humans appear to be programmed to pay attention to strong segments, since these segments carry not just segmental information but also clues about the location of prosodic domains. In turn, this information helps them to process speech more efficiently.

The second question is how to identify strong segmental properties. Although strong segments always perform the same function of demarcating the edges of prosodic domains, languages differ as to which segments behave as strong. In general, strong properties are consonantal, as it is consonants which occupy the syllable onset and are thus typically domain-initial. More specifically, stops are the most frequently occurring onset consonants cross-linguistically, so there is a tendency for languages to mark prosodic domains using
stop-related properties such as aspiration (e.g. English), ejectiveness (e.g. Maidu (Penutian)) and obstruent voicing (e.g. Wiyot (Algic)). Examples are given below.

2.2. Strong segmental properties

Although aspiration is a well-known feature of English pronunciation, it is easy to overlook its phonological function as a prosodic marker: English aspiration is tied to strong positions, its distribution being restricted to foot-initial position (e.g. [pʰ]ity, [m][pʰ]örtant) and word-initial position (e.g. [pʰ]öetic, [pʰ]erhâps). In all cases it is non-contrastive, since it adds no segmental information; but it does convey structural information, as it demarcates the left edge of the foot or word domain. This is illustrated by alternating forms such as st[u][pʰ]idity~stū[p]id, where the appearance of aspiration is sensitive to prosodic structure: in st[u][pʰ]idity the stop p is aspirated because it is foot-initial, whereas in stū[p]id the same p is foot-internal and therefore un aspirated. Here, the absence of aspiration shows that p has no prosodic marking function to perform; thus the presence/absence of aspiration reflects the strength/weakness of its prosodic position. English aspiration may therefore be seen as a strong segmental property because it is tied to prosodically strong positions. Aspiration also serves as a boundary marker in other systems such as some dialects of Basque.

In other languages, segmental properties other than aspiration may act as prosodic markers. In Maidu the ejective stops [p’t’k’] occur in (prosodically strong) syllable onsets but not in (relatively weak) syllable codas. (1b) shows how ejective [p’t’k’] weaken to plain [p t k] in coda (i.e. pre-consonantal) position.

(1)  a. ejective stop (onset)  b. plain stop (coda)

jēp’im kawáju ‘stallion’  jēpsí ‘men’
batám hit’í ‘butter’  ħitpe ‘fat’
bolōk’om sólo ‘moccasin’  bolōkti ‘put on shoes’

Shuswap (Salishan) is another native American language in which ejectives serve as prosodic markers. Further examples are discussed in Backley and Nasukawa (2009).

In some languages, however, it is less obvious how segmental properties are used to highlight prosodic structure. In Japanese, for instance, there does not appear to be any consonantal property which has a prosodic function of the kind just described. On the other hand, there are some dialects of Japanese which display a tonal property – a low tone realised on vowels – which is closely tied to the beginning of word domains. We claim that this low tone functions as a prosodic domain marker. In the literature on pitch accent patterns in (Tokyo) Japanese it is observed that low tone regularly appears on the first mora of the word domain. This is most clearly seen in unaccented words, where an initial L(ow) tone is followed by a succession of H(igh) tones, as in hi (L) ‘sun’, hashi (LH) ‘edge’, kuruma (LHH) ‘car’, shiawase (LHHH) ‘happiness’. Yet although this pattern is well-documented, few attempts have been made to explain the reason why this initial L is present. For example, Pierrehumbert and Beckman (1988) simply observe that a boundary tone L marks the left edge of domains. A notable exception is Warner et al. (2010), who present
experimental evidence that the initial L tone (perceived as a rise in pitch from an initial L to the first H) provides a reliable cue to the beginnings of words and that this L helps listeners to segment continuous speech into separate words. In this paper we endorse this finding, since it supports our assumption that L in Japanese acts as a prosodic marker in the same way that, for example, aspiration does in English. In the case of Japanese, however, an interesting complication arises such that L also has a lexical function. In §3 we propose that Japanese pitch accent involves an active lexical L, which operates independently of the prosodic marker L described above.

3. L as a (lexical) pitch accent property

TJ has lexical tone, but because tones are always assigned to accented syllables, it is dubbed a pitch accent system rather than a tonal system. Some traditional descriptions of TJ pitch accent assume that both H and L are active (Haraguchi 1991). Yet if H and L both exist as lexical properties, then it may be inferred that both have equal linguistic status. Evidently, however, this is not the case. What this traditional approach fails to acknowledge is that H and L are in complementary distribution, suggesting it may be more appropriate to employ just one lexical property, either H or L, and to treat its complement value as a default or unspecified tone. Yoshida (1999) argues for an analysis based on a single lexical H, while in this paper we propose to describe TJ pitch accent with just a lexical L tone. This L-based approach helps to strengthen the overall character of Japanese as an ‘L language’, in which L rather than H is the phonologically active property. The predominance of L in TJ manifests itself in several ways: (i) L serves as a lexical tone property, (ii) L operates as a prosodic marker, and (iii) L represents voicing contrasts in obstruents. Our approach also concurs with findings in applied linguistics, where it is reported that learners of Japanese master pitch accent patterns more easily when presented with words in which the lexical accent is identified by the location of L tones rather than H tones (Watanabe 2011).

For comparison, we begin by outlining a standard analysis of TJ pitch accent in which both H and L are active. In the examples below, the words in (a), (b) and (c) are accented while the word in (d) is unaccented. Describing a mora/syllable as ‘accented’ tends to imply that it has prominence, and in TJ this prominence results from a pitch change: the pitch falls after the accented mora. So in traditional analyses the accented mora is assigned a lexical H, giving it high pitch, while L is assigned to the following mora to produce a distinctive fall. Deriving the correct pitch patterns for native nouns involves the four operations shown in figures (2) to (5) (adapted from Haraguchi 1991: 17). In (2), H is assigned to the lexically accented mora (indicated by *).

<table>
<thead>
<tr>
<th></th>
<th>a. 涙 ‘tear’</th>
<th>b. 卵 ‘egg’</th>
<th>c. 鏡 ‘mirror’</th>
<th>d. 車 ‘car’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>na.mi.da</td>
<td>ta.ma.go</td>
<td>ka.ga.mi</td>
<td>ku.ru.ma</td>
</tr>
<tr>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td></td>
</tr>
</tbody>
</table>
In (3), L is assigned to every mora to the right of H. The L tone then spreads rightwards.

(3) a. 涙 ‘tear’ b. 卵 ‘egg’ c. 鏡 ‘mirror’ d. 車 ‘car’

   *          *                   *
na.mi.da   ta.ma.go   ka.ga.mi   ku.ru.ma
|       |             |  |                 |                         |
H     L           H  L                 H     L                 H    L

In (4), the H tone is allowed to spread leftwards.

(4) a. 涙 ‘tear’ b. 卵 ‘egg’ c. 鏡 ‘mirror’ d. 車 ‘car’

   *          *                   *
na.mi.da   ta.ma.go   ka.ga.mi   ku.ru.ma
|       |            |  |                 |                         |
H     L          H  L                 H     L                 H    L

In (5), which illustrates a process generally referred to as initial lowering, the tone on the initial mora dissimilates from the tone on the second mora.

(5) a. 涙 ‘tear’ b. 卵 ‘egg’ c. 鏡 ‘mirror’ d. 車 ‘car’

   *          *                   *
na.mi.da   ta.ma.go   ka.ga.mi   ku.ru.ma
|       |            |  |                 |                         |
H     L   L     H  L    L           H    L  L             H    L

Although the operations in (2)–(5) produce the correct patterns, the pitch assignment process is rather complex. More importantly, it is largely stipulative, its format echoing the serial rule applications employed in early generative phonology. Furthermore, it leaves a number of questions unanswered. For example, no motivation is given for the dissimilation process in (5), thus failing to offer any insight into why the initial mora must be different from the following one. Also, although L behaves like a default tone in that it is never lexically assigned to an accented mora, it still spreads as if it were an active property. It seems appropriate, therefore, to re-examine the way TJ pitch accent patterns are represented. Below we argue that these patterns should be analysed using just one active property, but that this property should be L rather than H. In recognising L as the active tone in TJ, we claim that (i) only L can spread, and (ii) a word’s lexical accent is signalled not by a high tone but by a fall in pitch; this fall results from the realisation of a lexical L tone following an unspecified (by default, high-pitched) mora.¹

There are several advantages of adopting an L-based approach. Firstly, it is possible to capture the asymmetry between H and L: their linguistic behaviour suggests that they do not have equal status in TJ. Secondly, according to the tenets of linguistic thinking, positing one lexical property rather than two is preferred in the interests of analytical economy. Thirdly, recognising active L in the tonal system adds weight to the idea that Japanese is a language in which L plays a more central role than H, not only at the tonal level but also at the segmental level. The examples below illustrate how the L-based approach produces pitch patterns involving two different kinds of L tone, prosodic L and lexical L. As described in §2, each example word has a prosodic L marking its left boundary, realised as a low tone on
the leftmost vowel. Independently of prosodic L, each accented word also has a lexical L assigned to a given mora. Lexical L produces a falling tone on the accented mora, because it is always preceded by a toneless mora realised with a default H tone (i.e. the absence of L).

The word *tamago* 卵 [tamago] (LHL) ‘egg’ has two L tones, a prosodic L on the first mora and a lexical L on the third mora (cf. traditional analyses, in which it is assumed that H marks a lexical accent on the second mora). This lexical L is perceived as a fall in pitch. Notice how lexical L spreads rightwards when the word domain is extended by the addition of the subject marker *ga*.

(6) prosodic marker L [ L ] [ L ]
    [ L ]
    ta.ma.go ta.ma.go.ga
    |
    lexical L L L >>>
    pitch pattern L H L L H L L

The word *kagami* 鏡 [kagami] (LHH…L) ‘mirror’ also has two L tones, a prosodic L on the initial mora and a lexical L later in the word. But in this case lexical L is not realised until the fourth mora – it only becomes audible when the word domain extends by adding a particle such as *ga* (nominative marker), *o* (accusative marker) or *ni* (dative marker). This could be represented in several ways. One option, given in (7), assigns lexical L to the right boundary of the word domain (cf. boundary tones used in some intonation systems). L is then realised on an available tone-bearing mora such as *ga*. Another option would be to assume that nouns such as *kagami* have an empty CV structure at their right edge, giving *ka.ga.mi.CV*. If the V slot is empty (i.e. silent) then it cannot bear a tone, but if it is filled by segmental material from a particle such as *ga* then it can be phonetically realised and, in addition, can bear the low tone associated with lexical L.

(7) prosodic marker L [ L ] [ L ]
    [ L ]
    ka.ga.mi ka.ga.mi.ga
    |
    lexical L L
    pitch pattern L H H L H L L

The word *kuruma* 車 [kuruma] (LHH) ‘car’ is accentless, so no fall in pitch occurs anywhere in the word domain, even when the domain is extended through the addition of *ga*.

(8) prosodic marker L [ L ] [ L ]
    [ L ]
    ku.ru.ma ku.ru.ma.ga
    |
    lexical L
    pitch pattern L H H L H H H

Finally, the word *namida* 涙 [namida] (HLL) ‘tear’ has a lexical accent on the second mora; this is where a fall in pitch occurs and a series of L tones begins. But in this case we observe lexical L interacting with prosodic L because they occupy adjacent syllables. To perceive the low tone associated with lexical L, it must be preceded by a high tone (i.e. a fall in pitch is produced by a H-L sequence). But in *namida* the preceding mora is domain-
initial and therefore bears prosodic L, which obscures the following L tone and prevents it from being perceived. Indeed, the only way of phonetically realising the lexical L on the second mora is to suppress prosodic L in the initial mora – and if a low tone is removed then a default high pitch takes its place. Effectively, then, lexical L overrides prosodic L. Or, expressing this in terms of linguistic information, speakers give a higher priority to realising a lexical property L than a prosodic marking property L. On this basis, the dissimilation process in (5) appears to be motivated by the need to realise lexical material. Note that in namida-ga lexical L again spreads as an active property to right edge of the word domain.

(9) prosodic marker L [(L)               ]  [(L)                       ]  
    na.mi.da           na.mi.da.ga  

    lexical L          L >>>           L >>>>>>

    pitch pattern   H L L          H L L L  L

(* L L L)  (* L L L L L)

Among the example TJ words considered so far, none displays a pitch pattern in which lexical L is assigned to the initial mora. In principle, there is no reason to rule this out as a lexical possibility. In practice, however, we claim that this pattern never occurs because listeners would not be able to perceive the lexical L in question. Recall that the phonetic cue to a lexical L is a fall in pitch, and that to realise a fall, L must be preceded by a mora which is unspecified for tone (and is therefore realised as H). But clearly this is impossible if the lexically-accented mora is word-initial. So although initially-accented words constitute a lexical gap, there is a legitimate reason for their absence: if such words existed, language users would be unable to pronounce or perceive their initial accent.

4. The predominance of L in TJ

We noted earlier that an L-based approach to TJ pitch accent strengthens the view that Japanese should be characterised as an L-type language in which L (cf. H) has a central role in the phonology. We have shown how L contributes to pitch accent patterns in two ways, as a prosodic marker (§2) and a lexical tonal property (§3). But L also has a third function in Japanese (Nasukawa 2005; Backley and Nasukawa 2009), serving as the active property which distinguishes voiced obstruents (with L) from neutral obstruents (without L). It is known that languages with a two-way laryngeal contrast follow one of two typological patterns (Harris 1994; Iverson and Salmons 2006; Nasukawa 2010). In aspiration/H languages such as English, Korean and Swedish the ‘voicing’ contrast is cued by degrees of voicelessness and/or aspiration, rather than by voicing per se. But in voicing/L languages such as French, Russian and Japanese the laryngeal contrast really does reflect a contrast between voiced and voiceless. Thus L is active in Japanese in three ways: as a prosodic marker, a laryngeal property and a pitch accent property. The L-based approach developed here not only offers an interesting account of pitch accent patterns but it also reinforces the view that L dominates the phonology of Japanese as a whole.
Notes
* This paper was first delivered at the Phonology Forum 2012, Tohoku Gakuin University, Sendai, on 20 August 2012. We thank the participants and two anonymous reviewers for their helpful comments. This work was partially funded by the Japanese government (Grant-in-Aid for Scientific Research (B), grant 22320090).

1 Following Nasukawa (2005) we assume that a syllable in Japanese contains exactly one mora, making the ‘heavy syllables’ CVN and CVV bi-syllabic strings. On this basis it is possible for the N in CVN (e.g. *kan* (HL) ‘can’ vs. *kan* (LH) ‘senses’) and the second V in CVV (e.g. *ɕii* (HL) ‘deliberation’ vs. *ɕii* (LH) ‘arbitrariness’) to be assigned a lexical L tone.

References