Canadian Journal of Linguistics/Revue canadienne de linguistique 48(3/4): 357–383, 2003

The Development of Lexical Pitch Accent Systems: An Autosegmental Analysis

MITSUHIKO OTA
University of Edinburgh

1. Introduction

One of the challenging tasks that children face in learning a language is to unravel the phonological system that lies behind the observable pitch patterns of the target language. Even when they manage to extract the linguistically relevant aspects of fundamental frequency (F_0) from the ambient data, they still need to learn how the surface pattern relates to many of the phonological phenomena in natural language for which F_0 is the primary phonetic parameter. This set of phenomena will be referred to here as "pitch phonology". For most languages, this includes utterance-level intonation that signals sentence types (e.g., question, statement) and information structure (e.g., focus), as well as pitch contours that mark boundaries of syntactic units. For example, in North American English a high-rising intonation generally indicates a question, whereas phrasal pitch differentiates otherwise ambiguous surface structures such as [old [men and women] vs. [[old men] and women]. In addition to these intonational pitch phenomena that apply to sentences or phrases, some languages use pitch to encode lexical differences. In so-called "tone languages", such as the Chinese and Niger-Congo languages, pitch marks a paradigmatic lexical contrast over syllables. For example, in Mandarin Chinese, segmentally identical words such as *ma* 'mother', *ma* 'hemp', *ma* 'horse' and *ma* 'scold' are distinguished by four different pitch patterns: high level, high rising, low falling rising and high falling. In "pitch accent languages", such as Swedish, Japanese, and Basque, pitch signals a syntagmatic lexical contrast over

An earlier version of this work has been presented at the 27th Boston University Conference on Language Development. I am much indebted to Heather Goad, Bryn Harris, Bert Remijsen, Yvan Rose, and two anonymous *CJL* reviewers whose valuable comments led to significant improvements. Special thanks to Marilyn Vihman for providing me with access to her child language corpus and to Satsuki Nakai for assistance in transferring the data.

moras or syllables. For example, Una (Donahue 1997) has a contrast between ˈ*bita* 'truly' and *bi*ˈ*ta* 'final', manifested by a high pitch on the accented syllable.

Children initiate the learning process without knowing which of these pitch phonology systems they are to be exposed to. No a priori expectation can therefore be held as to whether a movement in pitch marks a phrase boundary, a question, a lexical tone, or a lexical pitch accent. Furthermore, it is not enough for them to discern whether the ambient language is a tone language, a pitch accent language or an "intonational language" (in which pitch is primarily a phonetic correlate of phrase- and sentence-level phonology). This is because in all these systems pitch participates in both lexical and non-lexical processes. Tone languages and pitch accent languages also have intonation that is independent of the lexical use of pitch (Inkelas and Zec 1988; Pierrehumbert and Beckman 1988; Bruce 1977). Intonational languages such as English and German also use a pitch accent (i.e., a local pitch contour) to mark certain lexically stressed syllables (Pierrehumbert 1980). Hence, the fact that a language uses, say, a falling contour to indicate a lexical contrast does not guarantee that any falling contour in the same language has a lexical function.

A model of pitch phonology acquisition, therefore, must be flexible enough to allow for a wide range of phonological processes that underlie the surface F_0 data, but also must be restrictive enough to constrain the type of hypotheses that the learner may construct. In this article, I explore a learning model that satisfies these conditions, based on the autosegmental theory of pitch phonology, first proposed for lexical tone (Leben 1973; Goldsmith 1976) and later extended to intonation and pitch accent phenomena (Pierrehumbert 1980; Ladd 1983, 1996; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). Under this framework, all pitch phonology phenomena, including lexical tone, lexical pitch accent and intonation, are surface manifestations of a linear structure that consists of a sequence of autosegmental tones associated with certain points in the segmental string (Ladd 1996). These autosegments are drawn from a universal inventory of tonal features (e.g., High (H) and Low (L)). Thus, a falling lexical tone, a lexical pitch accent with a falling contour, and a falling intonation can all have the same linear structure in the tonal tier, that is, a sequence of H and L level tones, which is related to the rest of the phonological structure in different ways. In the case of lexical tone, both the H and L are independently linked to a lexically specified tone-bearing unit. For lexical pitch accent, the H-L sequence may be a predetermined combination and only the position of the H (or L) is associated in the lexicon to a specified ("accented") syllable. The H and L that give rise to an intonational contour are linked to the segmental string at the phrasal or sentential level.

In direct contrast to this atomic view on pitch phonology is what Ladd (1996) calls the "overlay" approach, according to which F_0 patterns are composed not of a series of local tonal units but instead of global contour shapes and slopes associated with intonation, lexical tones and/or lexical pitch, all superimposed on each other (Fujisaki 1983; Gårding 1983, 1987). Under this view, the learner would have to search for the component functions that, when added together, give the complex function that best fits the observed pitch contour. This learning procedure seems to run into the type of computational difficulty described above, since the pitch contour the child hears can be a product of any number of component contours, the individual functions of which need to be extracted from the overall F_0 pattern.

The autosegmental approach, on the other hand, provides not only the flexibility but also the restrictiveness that we would like to see in a model of pitch phonology acquisition. In this model, the unit of analysis employed by the child in examining the F_0 input data is the set of tonal autosegments. The abstraction of quantitatively continuous F_0 information into strings of discrete units constrains the possible phonological structures that can be postulated for the data. At the same time, the learning procedure is robust enough to accommodate lexical tone, lexical pitch accent, intonation and the combination thereof, since it can be generalised across these descriptively different systems of pitch phonology. Rather than trying to establish the general type of phonological system, the learning mechanism searches for the relevant phonological categories (e.g., syllables, phrases, sentences) to which the tonal features are linked, and also determines how the association is conditioned (e.g., lexically, syntactically, semantically).

As a matter of fact, autosegmental theory has already been shown to be a viable model of the acquisition of lexical tone. Demuth (1992, 1993, 1995), for example, has examined child Sesotho and has demonstrated that the development pattern of the language's tonal system can be analysed as a process of discovering underlying tones and learning autosegmental rules that apply to the tones. Overapplication of general autosegmental constraints, such as the obligatory contour principle (OCP), can also explain some of the errors that the Sesotho-learning child commits. However, in order to show that the autosegmental approach is generalisable to any type of pitch phonology, we need to establish its applicability to systems other than lexical tone. The current study contributes to this research program by investigating the development of languages that have lexical pitch accent in addition to intonational phonology. I first lay out some general corollaries of the autosegmental theory on pitch phonology acquisition, and briefly review previous developmental findings from child Swedish that have a bearing on the predictions. I then proceed to take a detailed look at the development of pitch contours in Tokyo Japanese in order to test this model of acquisition. The article concludes by considering some of the implications of the findings for the idea that the acquisition of all pitch phonology phenomena involves assembling distinct tonal features in a linear structure rather than learning the overall shape of contours.

2. Some Consequences of the Autosegmental Acquisition of Pitch Phonology

An autosegmental theory of pitch phonology acquisition has several corollaries, three of which will be discussed here. First, if patterns in surface F_0 data are analysed into strings of discrete level tones, the overall contours of utterances are acquired componentially. This means that in some cases only certain tonal components of a complex contour may be acquired before the rest of the configuration. Indications of this learning pattern can be found in the early production of Swedish. In Stockholm Swedish, stressed syllables can carry one of the two pitch accent types, usually referred to as Accent I (or "acute") and Accent II (or "grave"). The pitch contours of Accent I and Accent II in citation form are readily distinguishable since the latter has a noticeable double-hump shape, as illustrated by the contour lines in (1).[1]

(1) Pitch contours of Stockholm Swedish words in citation form:

a. Accent I


númmer 'number'


lámm 'lamb'

b. Accent II

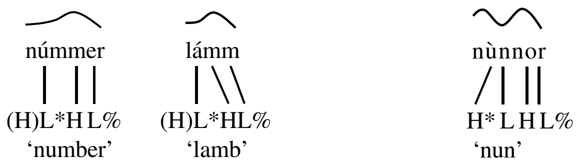

nùnnor 'nun'

Bruce (1977) has convincingly demonstrated that these contours are composites of a lexical pitch accent and sentence-level intonational features. The rise-fall at the end of both contours is the product of a post-stress peak on the focussed word in the prosodic phrase and a fall at the end of the utterance. Although these are intonational features, they apply equally to citation forms, as single-word utterances are also complete phrases and utterances. The difference between Accent I and Accent II comes from the lexical pitch accent, which is a high-low pitch contour in both cases but aligned differently with respect to the segmental string. For Accent I, the pitch peak is aligned before the stressed vowel. For Accent II, the peak appears on the stressed vowel. In autosegmental terms, these intonational and lexical components are the phrase accent H, the utterance-final L% ("%" indicates a boundary tone), and the lexical pitch accent HL, of which the central tone associated with the stressed vowel is L for Accent I (marked by an asterisk, i.e., L*), and H for Accent II (i.e., H*). The total effects of these component tones are shown in (2). In the citation form of an initially stressed Accent I word, the lexical H tone that precedes L* does not surface since it finds no segmental content to be linked to. The single-peaked contour arises from the L* pitch accent tone and the HL% intonational tones. The double peak of Accent II consists of the lexical pitch accent H*L and the intonational HL% tones.

[1] These figures are adapted from Bruce (1977) with some stylistic modification. The accent mark indicates the location of the stressed syllable. Examples of disyllabic and monosyllabic words are shown only for Accent I, since Accent II does not occur with monosyllabic lexical words.

(2) Autosegmental analysis of Stockholm Swedish words in citation form:

a. Accent I b. Accent II



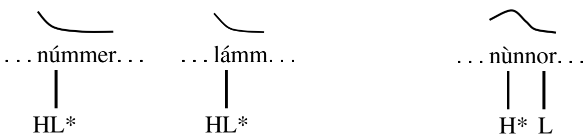
In acquiring these complex contours, the child may learn the global shape of the F_0 configuration as an unanalysed whole or, as the autosegmental model would predict, as a composition of tones. Evidence in favour of the latter can be found in the development of Accent II words. Engstrand, Williams, and Strömqvist (1991) examined the vocalisations of disyllabic Accent II words by children at 1;5, and discovered that they were differentiated from other vocalisations by a significant pitch rise into the second syllable, but not by a pitch fall toward the offset of the first syllable. In other words, only the second peak of the Accent II contour was reliably produced in contrast with other utterances. A possible interpretation of this finding is that the HL% (or at least the H) component of the Accent II contour has already been acquired by these subjects (hence the rise in the second syllable), but not the H*L (or at least the L) component of the lexical accent (hence the lack of differentiated falling contour in the first syllable). Consistent with this interpretation is the late mastery of lexical contrasts between Accent I and Accent II words, reported in a number of studies (e.g., Schmid 1986; Plunkett and Strömqvist 1992; Peters and Strömqvist 1996 and references therein), which suggests that the alignment of the lexical H(*)L(*) tones is acquired after the intonational component of the complex contour.

A second related corollary of the autosegmental approach is that the first aspects of pitch phonology to emerge are those whose surface contours consistently reflect the tonal structures. This is because the learning procedure involves the linking of level tones to prosodic categories, such as heads of words, edges of phrases, and utterances. Tonal segments are more likely to be put in place where a pitch pattern surfaces isomorphically within a certain domain of analysis, than where the tone-contour matching is inconsistent. If a particular contour pattern recurs regularly with phrases for instance, a corresponding tonal structure can be easily assigned at the phrase level, but if the surface contour pattern has a less regular surface contour pattern in the phrase, it is more difficult to recover the underlying tonal structure.

Again, the case of Swedish pitch phonology serves as an example. The analysis of the developmental data proposed above was that the phrase-level tone H, and possibly the utterance-level L%, emerge before the word-level lexical accent tones, H*L. According to the logic of the tone-contour matching, this is an expected outcome. The phrasal accent H and the terminal L% are both consistently realised as a post-stress peak in the phrase and a final fall in the utterance. On the other hand, the lexical HL tones do not always have a consistent surface contour pattern within their domain of association (i.e., the word). We can see this surface variability by comparing the contours of Accent I and Accent II words in isolation (given in (2)) with those in non-focal sentence position (shown in (3)).

(3) Autosegmental analysis of Stockholm Swedish words in non-focal sentence position:

a. Accent I b. Accent II



First, we notice that the HL* tone sequence in Accent I has a very different surface manifestation in (3a) from that in (2a). In (2a) the H lexical tone is not realised and the H phrasal tone that follows the lexical L* creates a rising contour. In contrast, (3a) lacks the phrasal H but has the segmental material to realise both the lexical H and L tones, giving rise to a falling contour. The H*L tone sequence in Accent II (3b) also has a surface contour that looks different from the corresponding portion of the contour in the single-word form (2b) because fewer tones are associated with *nùnnor* in the non-focal sentence position. While the pitch peak from H* is aligned with the stressed syllable in both cases, the low point that corresponds to the lexical L has shifted toward the end of the word in (3b), making its relation to a lexical tone less obvious. These discrepancies in the surface contours within the window of analysis render the tonal assignment in the word domain more opaque than that in the phrase domain, leading to the delayed acquisition of the full lexical tone assignment.

This also implies that the pitch peak after the stressed syllable must be analysed as a phrasal property when it emerges. Some supporting evidence can be found in the developmental data reported by Andersson (1990, cited in Plunkett and Strömqvist 1992:537). The child, Markus (1;10), applied the Accent II contour to disyllabic phrases (e.g., *gå vi* 'go we' and *få den* 'get it'). This overapplication error indicates that the child failed to notice that the high pitch phrasal accent must be realised within the stressed word, but he correctly understood that it is a feature of phrase-level intonation. In other words, the second half of the Accent II contour (i.e., the H tone) has been abstracted away from the double-hump shape and assigned to phrase phonology.

A third corollary of the autosegmental model is that there should be well-defined stages of development demarcated by the emergence of tonal targets in the contour. This can be illustrated by the hypothetical example contours in Figure 1. Imagine that the adult-like contours for the segmental strings CVCVCV and CVCVVCV have the shapes given at the endpoints of the learning sequences (a) and (b) in Figure 1. If these contours are acquired through autosegmental mechanisms, then the developmental pattern should exhibit stages that could be

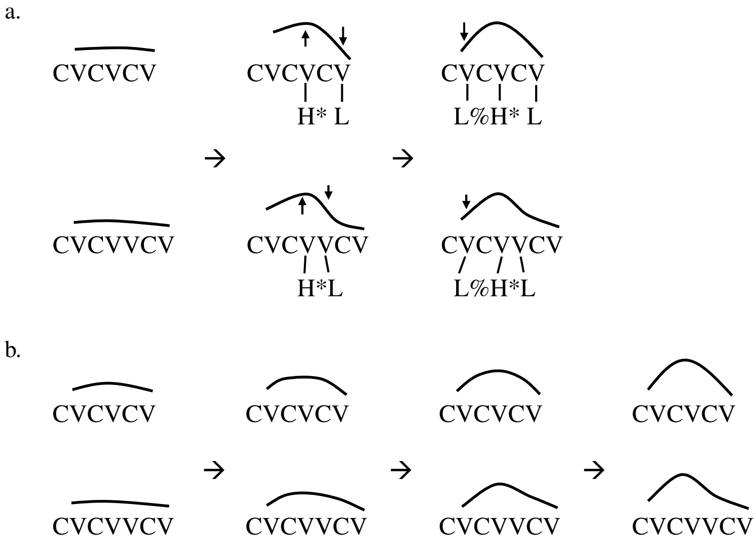


Figure 1: Two hypothetical patterns of pitch contour development. Sequence (a) represents a pattern consistent with the autosegmental model and sequence (b) a pattern that is inconsistent with the model.

easily described in terms of the addition of discrete tonal autosegments. This means that identifiable F_0 turning points (or "elbows") emerge where tonal targets are associated with tone-bearing units. For example, in sequence (a), the linking of H* to the medial syllable results in a pitch peak that is consistently aligned with the first vowel of that syllable, regardless of the structure of the syllable. The later addition of the phrasal L% is reflected in the lowering of the initial syllable. In contrast, the sequence in (b) shows a developmental pattern at odds with the view that tonal targets are involved in the acquisition of pitch contours. The shape and degree of slopes are being approximated globally and gradually, but there are no clear time points at which the F_0 elbows appear.

In the context of Swedish pitch development, this translates to the prediction that the first noticeable change in children's pitch contours occurs when H and L% become associated with the post-phrasal stress and the right boundary of the utterance (i.e., the appearance of a discernible pitch peak aligned with the post-stress syllable, followed by a distinct pitch fall). Around the same time, we should also begin to see another peak emerging on Accent II words, caused by the association of H* with the stressed syllable. Then we expect to see another contour change, reflecting the assignment of L* (the lowering of pitch in Accent I stressed syllables) and the placement of the lexical L (a pitch trough between the stressed and post-stress syllables in Accent II). Unfortunately, the available developmental data on Swedish pitch contours do not provide enough phonetic detail to verify these predictions.

While the studies on child Swedish reviewed here offer findings consistent with the autosegmental model of pitch phonology acquisition, none of them was specifically designed to test the model that is discussed here. As such, the data are open to alternative interpretations (see, for example, Peters and Strömqvist's 1996 explanation of the early emergence of the post-stress rise, which appeals to the perceptual salience and communicative functional load of the rise). The Swedish case is also complicated by the presence of dynamic stress, which involves amplitude and duration in addition to pitch, and also by the predictable pattern of stress assignment, which is defined by syllable weight and morphological structure. In the remainder of this article, I analyse some developmental data from Tokyo Japanese, another pitch accent language, in order to directly test the predictions of an autosegmental model of pitch phonology acquisition. Lexical accent in Tokyo Japanese is purely a matter of pitch, and its lexical assignment has, at least in nouns and functional items, no predictable pattern, thus making it a better test case for the learning model.

3. THE DEVELOPMENT OF PITCH CONTOURS IN JAPANESE

3.1. Pitch phonology in Tokyo Japanese

As with Stockholm Swedish, Tokyo Japanese uses pitch to mark both lexical accent and phrasing.[2] Lexical items are either accented (with one accented syllable) or unaccented (with no accented syllable). The only systematic phonetic correlate of the lexical accent is pitch; that is, accented syllables do not differ from others in duration or amplitude (Beckman 1986). When the accented syllable is CV:, CVV, or CVN, it exhibits a falling contour within itself, as can be seen in (4a).[3] When the accented syllable is CV, the pitch fall is realised over this and the following syllable (see (4b)). An utterance-final accented CV syllable, however, is only high pitched and does not show the dynamic pitch downfall observed in the preceding cases (see (4c)).

(4) Pitch contours of Tokyo Japanese words in citation forms:

a. akái 'red' b. néko 'cat' c. inú 'dog' d. buta 'pig' e. koori 'ice'

One important phrase-level pitch phenomenon is the rise that crosses the syllable boundary at the beginning of a phrase, which can be seen in (4a) and (4c–d). Although these examples are isolated words, they show this pattern because

[2] Japanese also has the garden-variety intonation pattern that marks sentence types (e.g., question, statement). As this aspect of pitch phonology is difficult to investigate without access to the speaker's utterance intention, it will be kept beyond the scope of this study.

[3] The acute accent mark indicates the location of the accent.

single-word utterances constitute phrases by themselves. As a consequence of this process, the unaccented word *buta* 'pig' in (4d) has a rising contour realised between the two syllables, even though it lacks an accented syllable. However, this type of cross-syllable rise does not occur when the phrase-initial syllable is accented as in (4b), or long as in (4e) (if the rhyme contains two sonorant moras, i.e., a long vowel, diphthong, or a short vowel followed by a nasal coda).

That the initial rise is a phrase-level property becomes apparent in longer phrases, as shown in (5).

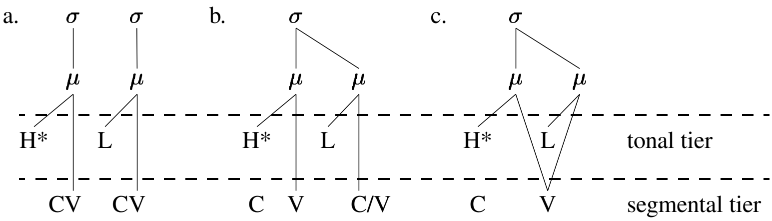
(5) Pitch contours of phrases in Tokyo Japanese:

a. kono néko-ga	b. kono inú-ga	c. kono buta-ga	d. kono koori-ga
this cat-NOM	this dog-NOM	this pig-NOM	this ice-NOM

In (5), the unaccented phrase-initial word *kono* 'this' shows a rise, but *buta* 'pig' (see (5c)) no longer has a rising contour, as it is not phrase-initial in this example. The final-accented *inú* still has a rise into the second syllable of the word (see (5b)) but a less prominent one. The falling contour associated with the lexical pitch accent in *néko* is unaffected in the phrasal context (see (5a)). The example in (5b) also shows the pitch fall of the lexical accent in *inú*, which is now realised between its final syllable and the following syllable (the nominative marker *ga*). The gradual downward trend of F_0 observed across the utterances in (5c) and (5d) is due to phonetic declination, which has no special phonological import in this context.

Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988) have proposed a now widely accepted autosegmental analysis of these patterns, which I adopt here with minor modifications. Under their analysis, the falling pitch accent is a tonal sequence of H*L, where the H* is associated in the lexicon with the left-most mora of an accented syllable, that is, the vowel in a CV(C) syllable, or the first portion of the vowel in a CV: or CVV syllable. These associations are illustrated in (6).[4]

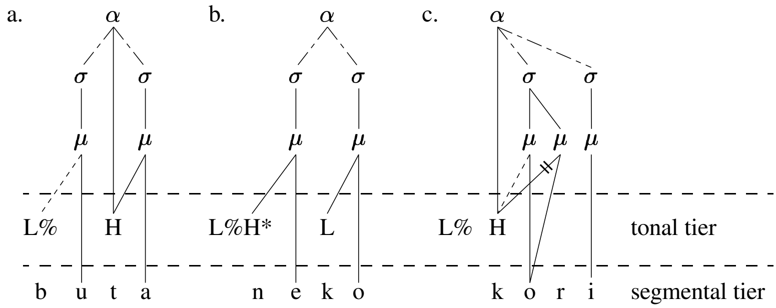
(6) Representation of pitch accent:



[4]The C/V notation in (6b) is representational shorthand to mean a C(onsonant) or V(owel). Pierrehumbert and Beckman (1988) are agnostic about the precise source of the lexical L. Here, I adopt the more traditional treatment of Poser (1984), according to which L is part of the underlying representation of the accent, associated with the mora following the one carrying H*. None of the analyses below hinges on this assumption.

The phrase-initial rising contour is due to two tonal autosegments assigned to phrasal units: the boundary L% tone, which is linked to the first mora of an accentual phrase, and the phrase default H tone, which is associated with the second sonorant mora of the phrase. This typically creates a sequence of L%H at the phrase-initial position. However, L% can be blocked by a high tone from linking to the first mora of the phrase. The high tone responsible for this blockage can be the H* of an initially accented syllable, or the phrasal H that spreads from the second to the first mora of a long syllable. The representations of these tonal associations are given in (7).[5]

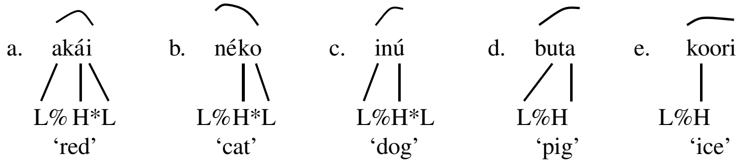
(7) Representations of phrasal contours:



The structure in (7a) shows how L% is linked to the initial mora when the phrase-initial syllable is short and unaccented. In (7b), the initial H* prevents the L% from linking to the phrase-initial syllable. In (7c), the phrasal H, which is assigned to the second mora of the phrase, reassociates with the initial mora and holds the L% in check. Based on experimental evidence, Pierrehumbert and Beckman (1988) argue that an unlinked L% still has a weak "allophonic" phonetic realisation with shorter and higher F_0 value, producing a brief and slight rise in the initial syllable. As illustrated by (4b) and (4e), this type of rise is different from that seen in (4a) and (4c–d) in that it does not extend to the following syllable.

Abbreviated autosegmental representations of the contours in (4) and (5) are given in (8) and (9) below.

(8) Autosegmental analysis of Tokyo Japanese words in citation forms:



[5]The symbol "α" stands for "accentual phrase" (a lower phonological phrase), which assigns the phrasal H. One of Pierrehumbert and Beckman's (1988) important claims is that the initial L% is licensed by the preceding accentual phrase. The dashed line indicates linking and "=" delinking. The half-broken line indicates that intermediate levels of prosodic structure between the accentual phrase and the syllable (e.g., prosodic word, feet) are omitted.

(9) Autosegmental analysis phrases in Tokyo Japanese:

a. kono néko-ga	b. kono inú-ga	c. kono buta-ga	d. kono koori-ga
/ | | |	/ | | |	/ |	/ |
L% H H*L	L%H H* L	L%H	L%H
this cat-NOM	this dog-NOM	this pig-NOM	this ice-NOM

Note the unlinked L% boundary tone in (8b) and (8e), which fails to give *néko* and *koori* the phrase-initial cross-syllable rise. Note as well the phrasal origin of the contour manifested in the citation form of *buta* in (8d). The rise here is due to the L% boundary tone and the phrasal H tone, neither of which is associated with the same word in a phrasal context (9c), hence the lack of rising contour on *buta* in the latter case. The H* tone is usually higher in F_0 value than the H tone in the unaccented word (Poser 1984; Pierrehumbert and Beckman 1988). This explains the shallower slope of the rise in (8d) in comparison to (8c), and also the step-up observed on the accented syllable in (9a–b) where the pitch accent H* follows the phrasal H. Thus, although *inú* and *buta* only show a subtle phonetic difference in citation form, they have different tonal structures, as revealed in the context of a longer phrase.

3.2. Predictions for the Japanese developmental data

Based on the above description and analysis of the lexical pitch and phrasal intonation in Tokyo Japanese, we can now lay out the specific predictions of the autosegmental acquisition model for pitch phonology development in Japanese. First, we predict that the contours outlined in (8) and (9) are acquired as components made up of L%, H, H*, and L tones, and therefore do not necessarily develop uniformly. Second, the H*L lexical tones are acquired before the L% and H phrasal tones, which means the falling contour of the lexical accent emerges before the phrase-initial rising contour. This is because the falling contour that surfaces from H*L is consistently observed between the accented mora and its subsequent sonorant segment, whether the word occurs in a single-word context or a phrasal context.[6] The L% boundary tone, on the other hand, does not have a constant phonetic realisation within its domain. As shown in (7a), it manifests itself in the form of a cross-syllable rise only when the initial syllable is short and unaccented. Deciphering the tone-domain association of L% is contingent upon acquiring the association of H*L because, apart from the reassociation of phrasal H, the learner must know that an initial accent blocks the linking of L%. The third prediction is that there will be stages of contour development demarcated by the emergence of tonal targets. Specifically, we predict the presence of an initial stage

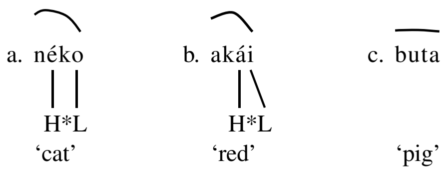
[6]There is an exception to this generalization. An accent assigned to a word-final CV syllable (e.g., *inú*) is not accompanied by a falling contour in a single-word context. However, even for these words, the high pitch of the accented syllable is consistently realised, so at least the assignment of H* is not hampered by this difference.

with no clear F_0 elbows (no tones), followed by the emergence of a falling contour whose F_0 peak is aligned with the first mora of the accented syllable (addition of H*L), and finally the lowering of the short initial syllable in unaccented words (addition of L%).

Findings relevant to these predictions are reported in a study by Hallé, de Boysson-Bardies, and Vihman (1991), who examined the global pitch contours of disyllabic words produced by four Japanese-speaking children at the 25-word session (i.e., when they produced at least 25 word types in a 30-minute recording session). The biological ages of the children ranged from 1;3.11 to 1;11.2. The results showed that the majority of words with a falling contour in the adult target form were produced with a global fall (83%), while less than half of the words with a rising contour in the adult target form were produced with a global rise (43%). These disyllabic words were in one-word utterances, so their targets are the citation forms shown in (8). The "falling" category consisted of *néko*-type disyllabic lexical items with initial pitch accent (as in (8b)). The "rising" category consisted of *inú*-type disyllabic words with the accent on the second syllable (as in (8c)) and also of *buta*-type words with no lexical accent (as in (8d)). It is not clear whether the analysis included *akái*-type words that manifest both an initial rise and a fall on the accented syllable.

The fact that many of the "rising" targets were produced with a falling contour is consistent with the interpretation that the lexical falling contour is acquired before the phrasal rising contour.[7] If H* and L come into place before L%H, there should be a stage where *néko*-type words ("falling" targets) show a falling contour, but *buta*-type words ("rising" targets) lack a phonologically meaningful pitch movement. This state of affairs is illustrated in (10).

(10) Predicted contours of citation forms in an intermediate stage:



At this stage, the contour for unaccented words, such as *buta* (10c), should be essentially flat, but because of the phonetic variability in the data, some will have a slight rise and some a slight fall. This may be the pattern underlying the data reported in Hallé et al. (1991). To test if this is the correct interpretation, however, we need to know the difference in the degree of slopes, not only the proportion of rising or falling contours produced for each type of targets.

Also, by examining only the global contours of the child's words, Hallé et al. may have overlooked some local movements in pitch that provide evidence

[7] This is not the interpretation given by Hallé et al. (1991), who link this pattern to the predominance of falling contours contained in the ambient language, which may have been overgeneralised to the rising targets.

for tonal acquisition. The child's production of words such as *akâi* (10b) may have an initial rise, but can still be classified as having a global falling contour if the difference between the initial F_0 level and the peak is smaller than the difference between the peak and the end-point of the fall. In order to resolve these methodological issues, I have conducted a more detailed analysis of the pitch contours produced by Japanese-speaking children of the same age group. The results are reported below.

3.3. The data

The data analysed in this study consisted of one-word utterances spontaneously produced by five children, each at their 25-word session. Speech samples of two children were taken from the same corpus studied by Hallé et al. (1991), originally collected by de Boysson-Bardies and Vihman (1991). Other children from the corpus were not included in the analysis as their parents spoke variants of Japanese other than the Tokyo dialect.[8] The data of the three other children were taken from Ota (1999, 2003). The age and gender of the children are given in Table 1.

Table 1: Profile of the Japanese subjects

Child	Haruo[a]	Maiko[a]	Takeru[b]	Hiromi[b]	Kenta[b]
Age at 25-word session	1;6.7	1;6.18	1;8.13	1;9.12	2;1.25
Sex	Male	Female	Male	Female	Male

Data sources: [a]de Boysson-Bardies and Vihman (1991)
[b]Ota (1999, 2003)

Samples with background noise and unanalysable vocalisation due to whispering, crying, shouting, and creaky voice were discarded. F_0 contours were extracted from the remaining samples using the autocorrelation function in Praat (Boersma 1993), and then smoothed (bandwidth = 10Hz) to remove perturbation effects. The total number of samples analysed for each child is: 50 (Haruo), 24 (Aiko), 21 (Takeru), 35 (Hiromi), and 29 (Kenta).

In addition, as a sample of the input pattern the children were exposed to, I extracted the pitch contours of their mothers' utterances that included the same words produced by the children. Because the mother was not the main interlocutor for the recording sessions that featured Hiromi and Kenta, the data consisted only of the mothers' utterances directed to the three other children. A total of 83 samples from mothers were analysed.

[8]The pitch phonology of other dialects of Japanese can differ from the Tokyo dialect in lexical and/or phrasal aspects. For example, Osaka Japanese lacks initial lowering but has two types of phrasal tones, H and L.

3.4. Global contours of early Japanese word production

In the first analysis, the steepness of the global contours was measured in the productions of disyllabic targets. The excursion of F_0 was calculated for each single-word utterance using the formula in (11). Falls were marked with a negative value, rises with a positive one. Figure 2 illustrates the method of measurement.

(11) F_0 excursion:

F_0 excursion = F_0 range / F_0 onset value x 100

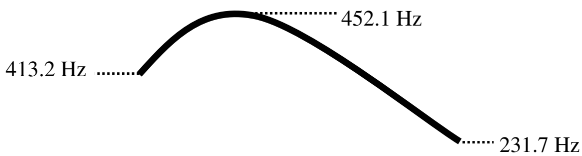


Figure 2: Measurement of global pitch excursion. In this hypothetical example, the F_0 excursion is $(231.7 - 452.1) \div 413.2 \times 100 = -37.5$ (%).

As with the case in the Hallé et al. (1991) study, target words with an initially accented syllable (e.g., *néko*) were considered to have a global fall ("falling targets"). Only target words with no lexical accent (e.g., *buta*), but not those with a final accent (e.g., *inú*), were included in the category of global rise ("rising targets") because there were too few cases of the latter in the data. Target words with both an initial rise and a fall (e.g., *akái*) and those with a long initial syllable (e.g., *koori*) were excluded from this analysis.

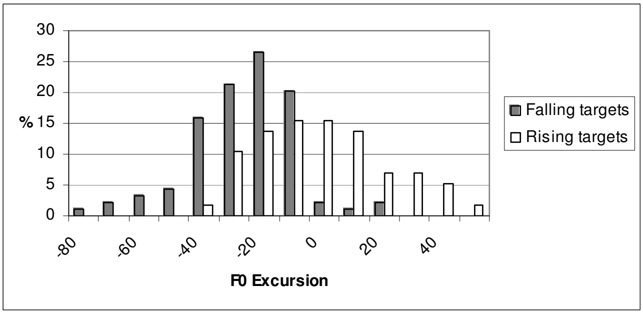


Figure 3: F_0 excursions of children's disyllabic production (pooled data). The x-axis shows the ranges of excursion. The y-axis indicates the proportion of the excursion values that fell into each range. For example, the highest dark bar shows that 27% of all the child productions for falling targets had an excursion that ranged from −10% to −20%.

Figure 3 shows the distribution of excursion values pooled for all subjects. The results confirm Hallé et al.'s findings that the majority of disyllabic falling target words are produced with a global fall, while a good portion of the disyllabic rising target words fail to receive a global rise. However, there was a significant difference between the mean production contours for falling and rising targets [-30.66 vs. 5.85; $t = 11.20$, $df = 146$, $p < 0.001$]. Crucially, many of the contours for rising adult words were very shallow in slope, as one would expect if the data capture a stage where the lexical H*L tones have been acquired but the tonal association of the phrasal L%H has not been mastered.

Not all children followed such a pattern, however. Table 2 summarises the mean excursions for falling and rising targets produced by each child. The results of the *t*-test show that one child, Maiko, does not show a contrast between rising and falling words, both of which are produced with a small negative excursion (i.e., slight fall).

Table 2: Mean F_0 excursions of children's disyllabic production (%)

Child	Falling targets	Rising targets	Comparison of means
Maiko	−16.2	−15.3	$t = -0.17$, $df = 17$, n.s.
Hiromi	−28.0	−2.1	$t = -5.11$, $df = 28$, $p < 0.001$
Kenta	−25.8	4.9	$t = -6.22$, $df = 31$, $p < 0.001$
Takeru	−35.7	19.3	$t = -9.70$, $df = 29$, $p < 0.001$
Haruo	−38.0	21.6	$t = -6.50$, $df = 33$, $p < 0.001$

Figure 4 shows the range of excursions produced by each child. There are two children who exhibit the tendency consistent with the pooled data: Hiromi and Kenta. For both children, the excursion rate is lower for the falling targets than for rising targets. However, while they reliably produce a falling contour for falling targets, most of their productions for words with a global rising contour have excursions that centre around the 0% line. Two other children, Takeru and Haruo, not only have a lower excursion rate for falling targets than for rising targets, but also have a positive excursion for the majority of their productions of rising targets, and a negative excursion for all of their productions of falling targets. The fifth child, Maiko, produces falls for both falling and rising targets, and the excursions of the fall are much smaller than those of the other children.

The results, therefore, support the predicted asymmetric development of the falling lexical pitch accent and the rising phrase-initial intonation pattern. While three of the children exhibit both or neither components of pitch contours in their disyllabic words, two others (Hiromi and Kenta) show productions of falling contours for initially accented target words without exhibiting clear rising contours for unaccented targets. These two fit the profile of a learner who has acquired the lexical H*L but not the phrasal L%H. None of the subjects has the opposite pattern of development – consistent production of rising contours for rising targets and flat contours for falling targets – which would have contradicted the prediction.

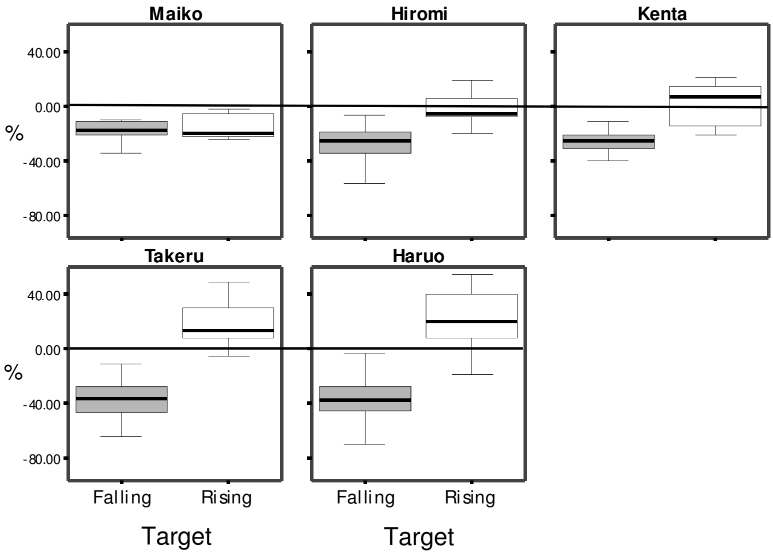
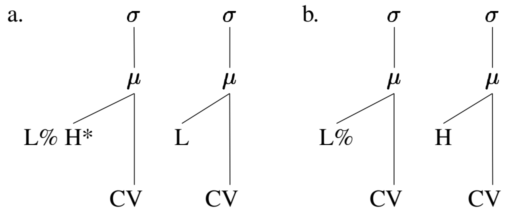


Figure 4: F_0 excursions of children's disyllabic production (individual data). The boxplots represent the interquartile range of excursions produced by each child for falling targets (shaded box) and rising targets (white box). The handles show the highest and lowest values, and the thick line in the box, the median value. A line is drawn at the zero excursion point as a reference point.

3.5. Local pitch movements in early Japanese word production

The analysis above examined the contours at the whole-word/utterance level. But the hypothesis that the tonal features of the lexical falling contour emerge before those of the phrase-initial rise predicts pitch movements at a more local level. If H*L is acquired, there should be a pitch downfall between the accented mora and the following mora (see (6)). If L%H is acquired, there should be a pitch rise between the initial mora and the following mora, as long as the first mora is part of a short unaccented syllable (see (7a)). The best place to examine this contrast is the first two syllables of a single-word utterance where the first syllable is short (see (12)). If both tonal components are acquired, the pitch from the first to second syllable should fall when the first syllable is accented as in (12a), but rise when the first syllable is not accented as in (12b). The developmental prediction is that children undergo a stage during which only the pitch movement associated with (12a) is produced.

(12) Representations of phrase-initial /CV.CV. . . / words:



In measuring the pitch between the two syllables, it is important to take into account the potential effects of the unlinked L% in (12a), which may lower the pitch of the left edge of the initial syllable. Therefore, measurements of the F_0 change were made from the offset of the vowel in the initial syllable and the midpoint of the first vowel in the second syllable. A schematic illustration of the method is shown in Figure 5.

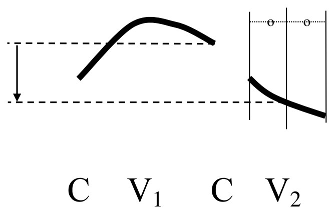


Figure 5: Measurement of initial local contours in Japanese words

The targets analysed were initially accented words, and unaccented words in phrase-initial position. Words with a non-initial accent were not included in this quantitative analysis, again because too few examples were available in the data (however, a qualitative analysis of these target words is given in the following section).

Table 3: Mean F_0 difference between CV_1 and CV_2 (Hz)

Child	#H*L. . . targets	#LH . . . targets	Comparison of means
Maiko	−21.0	−12.6	$t = -0.58$, $df = 13$, n.s.
Hiromi	−41.2	−4.7	$t = -4.36$, $df = 23$, $p < 0.001$
Kenta	−43.0	5.4	$t = -4.27$, $df = 31$, $p < 0.001$
Takeru	−90.5	41.0	$t = -7.26$, $df = 28$, $p < 0.001$
Haruo	−41.2	16.6	$t = -3.90$, $df = 24$, $p < 0.001$

The results are shown in Table 3 and Figure 6. Table 3 shows the mean F_0 difference between the first and second syllable for initially accented words, and unaccented words. As the *t*-test results comparing the mean values of the two categories indicate, Maiko does not show a difference between initially accented words and unaccented words. Both word types have a slight drop in pitch between

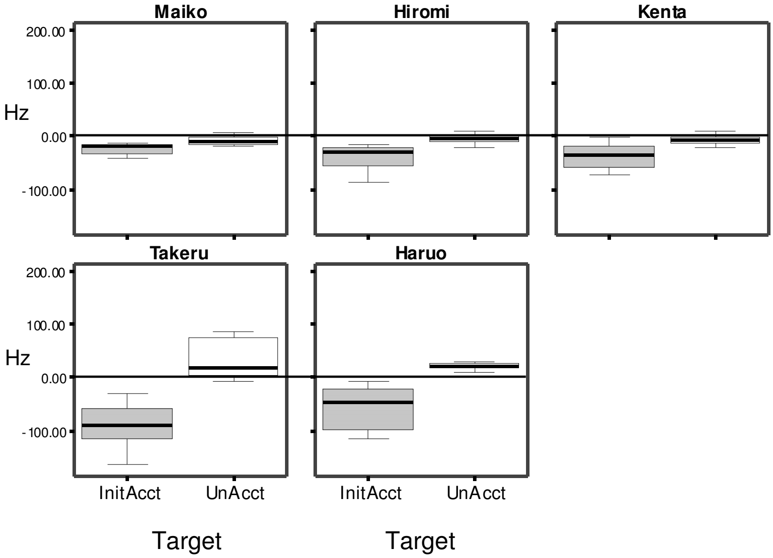


Figure 6: F_0 difference between CV_1 and CV_2 (Hz). Each boxplot shows the child's interquartile range of the pitch difference between the first and second syllable measured in Hz. The handles show the full range, and the thick line the median point. A reference line is drawn where the difference is zero.

the two syllables. The four other children make a distinction between the two words types.

Figure 6 shows the range of F_0 differences found between the initial short syllable and second syllable. Hiromi, Kenta, Takeru, and Haruo all exhibit a local fall for initially accented targets. However, only Takeru and Haruo produce a local rise for unaccented targets. Hiromi and Kenta, on the other hand, show almost no pitch movement for unaccented targets. In other words, for these two children, the transition from the first to second syllable is flat when the syllables are in a phrase-initial unaccented word — a pattern that is predicted if H*L, but not L%H, is acquired. The results of this analysis corroborate those from the previous section. In support of the prediction that lexical accent H*L tones are acquired before phrasal L%H tones, some children display only the falling contour associated with the former, while none shows only the rising contour associated with the latter.

3.6. Developmental stages in Japanese pitch contours

In section 3.2, we predicted that the pitch data of Japanese pitch phonology should reveal demarcated changes as tonal targets are assigned to the contours step by step, beginning with a stage with no identifiable F_0 turning points, followed by the emergence of an aligned pitch peak due to H*, and then by the lowering of a phrase-initial short syllable to reflect the addition of L%. Although a thorough examination of these predictions requires longitudinal data, a preliminary analysis can be conducted using the cross-sectional data of this study.

The analyses in sections 3.4 and 3.5 imply that the five children are in different stages of such a progression, with Maiko showing signs of the "no tonal structure" stage, Hiromi and Kenta, of the "lexical pitch accent only" stage, and Takeru and Haruo, of the "emergence of phrasal pitch" stage. If these stages are defined in terms of tonal assignments, we expect to find the exact contour characteristics described above in the three groups of children. Below I draw examples of pitch contours from four children to demonstrate that this is the case.

Figure 7 displays three samples of pitch curves from Maiko's production. What is striking about these contours is their similarity and the lack of discernible F_0 elbows.

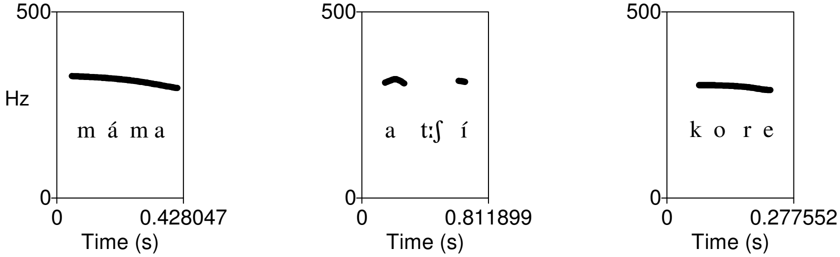


Figure 7: Maiko's pitch contours

The first target word *máma* 'mama' has a pitch accent on the first syllable, which should result in a sharp fall into the second syllable, but instead we find a slight slope that extends over the whole word/utterance, which is reminiscent of the declination we find in the adult production of non-phrase-initial unaccented words (cf. *buta* in (9c)). More or less the same shape is found in the contours of the other two examples, except that the line for *attʃí* 'that way' is interrupted by the affricate geminate. Both the final accented *attʃí* and unaccented *kore* 'this' have a cross-syllable rise in the adult targets, but neither of Maiko's productions shows such a pitch movement. The evidence here points to the lack of underlying tonal structures in these contours.

Figure 8 shows the contours of similar words produced by Hiromi. Here we see the adult-like drop of pitch between the two syllables in the initially accented target *pápa* 'papa', which suggests the emergence of the lexical tones H*L. A question that arises is whether *pápa* really has a tonal structure that contains H* and L instead of simply being associated with a falling pitch movement. If in fact distinctive tonal features are behind the falling contour of accented words, the most reliable correlate must be the alignment of the pitch peak; that is, the phonetic realisation of H* associated with the first mora of the accented syllable,

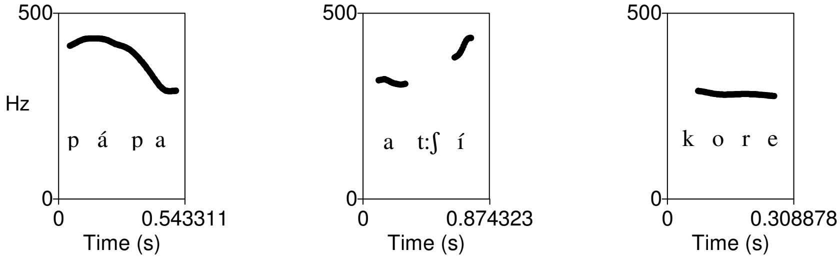


Figure 8: Hiromi's pitch contours

rather than the shape and degree of the slope of the falling contour. In order to verify this, we need to look at words with slightly different structures.

The contours in Figure 9 show two more phrases produced by Hiromi, *máma* 'mama' and *nénne yo* (sleep EMPHATIC.PARTICLE) 'sleep!', both consisting of only sonorant segments. For this reason we see every segment bearing some level of pitch. Both *máma* and *nénne* are initially accented, and we see an overall decline of pitch in the two contours.

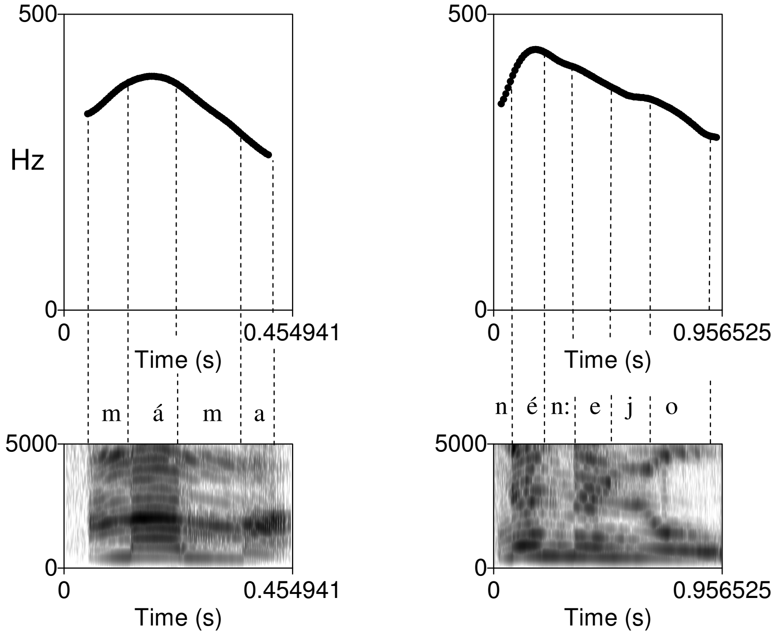


Figure 9: Alignment of H* with the first mora of the accented syllable (Hiromi's production)

However, in both cases, there is also a pitch rise at the beginning, which is more pronounced than the slight rise that can be observed in *pápa* (Figure 7). The invariable property in all of these cases is the F_0 turning point, which always occurs within the vowel of the accented syllable. The sonorant onsets /m/ and /n/ in *máma* and *nénne* show a rise because the pitch must peak on the nucleus which follows them. This epiphenomenal pitch rise is almost invisible in the *pápa* case because the onset, which is a voiceless stop, cannot carry pitch, and the peak appears much sooner in the contour. Once peaked on the nucleus of the accented syllable, the pitch turns around, as can be seen in the decline on the onset of the second syllable in *máma* and the geminate nasal of the second syllable in *nénne*. These contours are therefore best defined in terms of the association of H* with the accented syllable, rather than as a constant configuration of pitch slopes.

Returning to Figure 8, the contrast between *attʃí* 'that way' and *kore* 'this' is interesting in that only the former exhibits an adult-like rise between the syllables. This difference corresponds to the different tonal structures behind these two forms; while both have rising contours in the adult form, only *attʃí* has a lexical accent (on the second syllable). Thus, if L%H is missing, but H* is already in place, we expect *attʃí* and *kore* to have the tonal structures shown in (13); these are consistent with the surface forms we see in Figure 8 if we assume that the initial syllable in (13b), which is underspecified for tonal features, receives a neutral pitch value that is lower than the phonetic target of H*.

(13) The tonal structures of *máma*, *attʃí*, and *kore* in Hiromi's pitch system:

a.	máma	b.	attʃí	c.	kore
	| |		|		
	H*L		H*		

Further support for this analysis is found in Kenta's pitch contours, which are given in Figure 10.

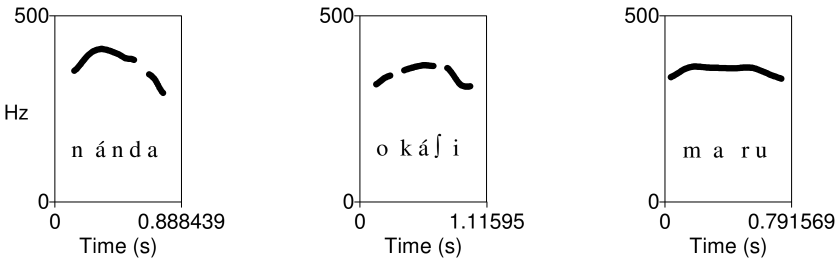


Figure 10: Kenta's pitch contours

As in Hiromi's case, there is a clear pitch peak aligned with the vowel of the accented syllable in *nánda* 'what's (that)'. There is no sign of cross-syllable rising in the unaccented disyllabic target *maru* 'circle', again like Hiromi's *kore*. These indicate the presence of H*L and the absence of L%H. If the analysis in (13) is correct, and Kenta is in the same stage of development as Hiromi, then we predict his production of a word such as *okási* 'candy' to exhibit a rise from the underspecified initial syllable toward the H*-associated second syllable, followed by a fall from the second to third syllable due to the lexical L. As Figure 10 shows, this is indeed the case.

Figure 11 shows three word/utterance forms produced by Takeru. The initially accented disyllabic word *kákka* 'mom' has the peak aligned with the vowel of the accented syllable. The lack of within-syllable initial rise is attributable to the voiceless stop onset of the initial syllable, as in Hiromi's production of *pápa*. However, a cross-syllable rise appears in both the second-accented *itái* '(it) hurts' and the unaccented *buta* 'pig'. Since the rise in unaccented words is not due to H*, the contour on *buta* must be a product of L%H. Therefore, the acquisition of L% and H is what makes Takeru's contour for unaccented words different from that of Hiromi and Kenta.

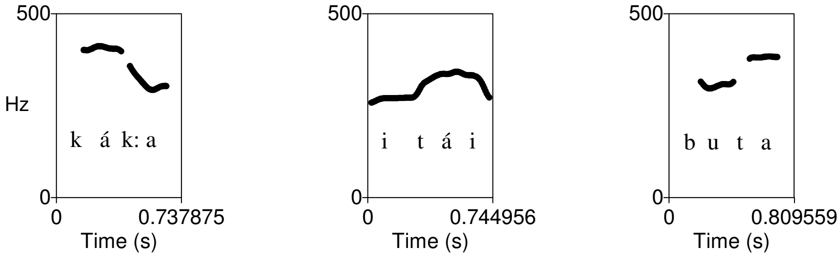


Figure 11: Takeru's pitch contours

Comparing these pitch contour samples, we see that the differences among the children reflect developmental stages with different sets of tonal features active in the system. The general lack of pitch turning points in Maiko's data indicates that no tonal autosegments are assigned to the contours. The pattern in Hiromi's and Kenta's data has the properties of a system in which the H* and L tones are associated with the segment that carries lexical pitch, but the L%H tones are still not linked to their phrasal positions. Crucially, the presence of H* and the absence of L%H explain the asymmetry between non-initially accented words (e.g., *attʃí*, *okáʃí*) and unaccented words (e.g., *kore*, *maru*), only the former of which show a rising contour, even though both have a rise in the adult system. The invariant property of the shape of the contours on accented words is the pitch peak, which is consistently aligned with the nucleus of the accented syllable. This supports the analysis that the pitch movements of lexical accent arise from the association of H* to the accented syllable. Finally, Takeru's data present a pattern that is consistent with the mature system. The progression suggested by these stages of development is characterised by discontinuous accumulation of local features rather than global and gradual approximation of the adult contour slopes.

3.7. Variability of phrase-initial rise

The previous sections have demonstrated that the falling contour of the lexical pitch accent develops earlier than the phrase-initial rising contour. In this section, I briefly comment on the source of this staggered development of the two aspects of Japanese pitch phonology. According to the autosegmental model of pitch phonology acquisition, such a pattern of development follows from the idea that the acquisition of pitch contour involves learning how to associate tonal features with the rest of the phonological structure. What makes the learning of the phrase-initial rise more difficult than that of the lexical pitch accent is the inconsistent tone realisation due to the allophony of the L% tone. The cross-syllable rise is not observed when the phrase-initial syllable is accented or long, because the linking of L% to the initial syllable is blocked by a H tone associated with the initial syllable.

Children are in fact faced with this type of variability in the phonetic realisation of phrase-initial rise. This can be demonstrated by examining child-directed speech using the type of local pitch movement analysis given in section 3.5. The results are shown in Figure 12, which displays the F_0 changes between the first two syllables of initially accented words (shaded boxes) and unaccented words (white boxes) as produced by the mothers. The same method described in Figure 5 is used to measure the F_0 difference between the vowel offset of the short initial syllable and the midpoint of the vowel in the second syllable. Measurements were made separately for words produced phrase-initially (the boxes on the left) and phrase medially (the boxes on the right).

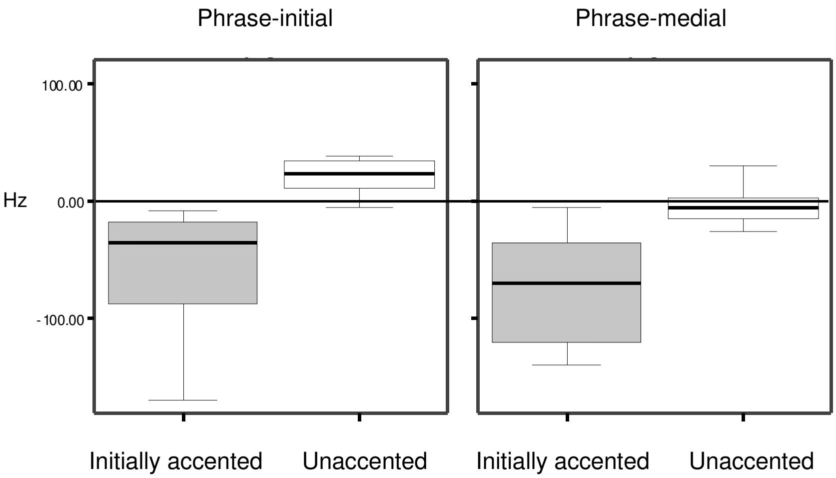


Figure 12: F_0 change (Hz) between CV_1 and CV_2 in Mothers' productions of /CV_1CV_2. . . /words: phrase-initial and phrase-medial positions. The plot on the left shows the initial pitch movement in initially accented words, and in unaccented words in the phrase-initial position. The plot on the right shows the initial pitch movement of the same words produced in the phrase-medial position.

By comparing the two shaded box plots, we see that initially accented words in child-directed speech has a robust falling contour, whether the word appears at the beginning (the left shaded box) or in the middle (the right shaded box) of the prosodic phrase. On the other hand, the comparison of the two box plots on the left shows that the pitch rises between the first two short syllables when a phrase-initial word is unaccented (the left white box) but not when it is initially accented (the left shaded box). In other words, the child is given consistent phonetic evidence for lexical pitch accent (i.e., H*L), but not for phrase-initial rise (i.e., L%H(*)), which is dependent on the absence of initial accent (and also on the initial syllable being short).

The reliability of phonetic cues for the phrase-initial rise may be further reduced by the amount of word types in child-directed speech that fail to exhibit the pitch pattern. In Table 4, all initial words that appeared in 791 utterances produced by the mothers were classified into those with an initial accent, a non-initial accent, and no accent. Names of the children and interjections were not included in the count. The non-initially accented words and the unaccented words were further divided into two sub-categories depending on the length of the initial syllable. Of the five resulting categories, those that do not show an initial cross-syllable rise in the adult production are words with an initial accent or a long initial syllable. Together, these account for 54.7% of the words children hear in phrase-initial position. Therefore, telltale evidence for phrase-initial rise is available only in about half the words that appear at the beginning of the phrase.

Table 4: Distribution of word types in mothers' utterance initial position

Accent	Initial	Non-initial		None	
Length of initial syllable	Short or Long	Short	Long	Short	Long
Initial rise	No	Yes	No	Yes	No
Percentage	48.5	24.2	3.0	20.2	3.2

Another potential reason for the lagging development of the phrase-initial rise is that children may find it more difficult to produce a rising contour than a falling contour because pitch rises take longer and require more physiological effort to implement (see Snow 1998). This, however, cannot explain the asymmetry described in Figures 8 and 10, where some children produce a rise for words with a non-initial accent, but still fail to produce the phrase-initial rise for unaccented words.[9] The overall evidence is in favour of the account that the delayed development of the phrase-initial rise is caused by the inconsistent surface realisation of the L% tone, which makes its tone-domain association more difficult to acquire.

[9]Similarly, the physiological account does not square with the Swedish data, in which the rising portion of the complex Accent II contour emerges before the falling portion.

4. Conclusion

Human languages display a wide range of pitch phonology phenomena, which seemingly poses a learnability problem for children who are not equipped with a top-down process to determine the type of system they are exposed to. This will not be a problem, however, if the learning process consists in associating a set of distinct tonal autosegments with the relevant phonological structures. Under this view, the task involved in learning typologically different systems, such as lexical tone, lexical pitch accent and/or intonation, is simply that of learning the organisation of the tones drawn from a common inventory.

This article has demonstrated the applicability of such a model to the development of lexical pitch accent systems. Predictions that follow from the autosegmental model match many of the facts we know about the acquisition of Swedish and hold up in detail against the production data from child Japanese. Specifically, in support of the view that contours in a language with lexical pitch accent and phrasal intonation are acquired in components, the falling and rising slopes of overall contours have been shown to develop separately. In Japanese, the late emergence of the phrase-initial rise is attributable to the allophonic surface realisation of the boundary L% tone. Moreover, transitions in contour shapes are characterised by punctuated appearance of F_0 turning points, providing further evidence that pitch contours are acquired as strings of tonal targets.

The last point has some important implications for the theory of pitch phonology acquisition in general. The traditional view on prosodic acquisition is that pitch is controlled early in development, and that prelinguistic children are already capable of encoding pragmatic information both in intonational languages such as English (Halliday 1979; Furrow 1984) and in tone languages such as Mandarin Chinese (Clumeck 1980). There is also some evidence that intonation patterns in babbling display language-specific characteristics (e.g., de Boysson-Bardies, Sagart, and Durand 1984). However, it is also clear that most of the crucial characteristics of the pitch contours of the target language are acquired during the one-word stage (Crystal 1986; Hallé et al. 1991; Vihman 1996). The view that pitch movements should be analysed as tonal autosegments linked to segmental strings can explain why there should be a qualitative transition in pitch phonology that coincides with the linguistic acquisition of segmental phonology, since the tonal structure is parasitic on the segmental structure.

Finally, the developmental data in this study offer a new perspective on an on-going debate in intonational phonology. As mentioned in section 1, the autosegmental framework differs fundamentally from the "overlay" approach which views F_0 patterns as overlaid intonational and lexical contours defined in terms of their shape and slope. For example, the Japanese contour is seen as the product of superimposing the curve generated by the accent component onto the global shape generated by the phrase component (e.g., Fujisaki 1983). While the developmental pattern presented here does not contradict the overlay approach, it corresponds more to the analysis that what is acquired is the structural association of segment-size tones rather than slopes defined over larger domains. Given the difficulty in finding decisive empirical evidence in adult data for one position or the other, further research in the development of pitch phonology may shed new light on this issue.

References

Andersson, Lennart. 1990. *Ett barns väg till accent I*. Göteborg: University of Göteborg.

Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Beckman, Mary, and Janet Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3:255–310.

Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* 17:97–110.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Lund: Gleerup.

Clumeck, Harold. 1980. The acquisition of tone. In *Child phonology* 1, ed. Grace Yeni-Komshian, James Kavanaugh, and Charles A. Ferguson, 257–275. New York: Academic Press.

Crystal, David. 1986. Prosodic development. In *Language acquisition: Studies in first language development*, ed. Paul Fletcher and Mike Garman, 174–197. Cambridge: Cambridge University Press.

de Boysson-Bardies, Bénédicte, and Marilyn M. Vihman. 1991. Adaptation to language: Evidence from babbling and early words in four languages. *Language* 61:297–319.

de Boysson-Bardies, Bénédicte, Laurent Sagart, and Catherine Durand. 1984. Discernible differences in the babbling of infants according to target language. *Journal of Child Language* 16:1–17.

Demuth, Katherine. 1992. The acquisition of Sesotho. In *The crosslinguistic study of language acquisition* 3, ed. Dan I. Slobin, 557–638. Hillsdale: Lawrence Erlbaum.

Demuth, Katherine. 1993. Issues in the acquisition of the Sesotho tonal system. *Journal of Child Language* 20:275–301.

Demuth, Katherine. 1995. The acquisition of tonal systems. In *Phonological acquisition and phonological theory*, ed. John Archibald, 111–134. Hillsdale: Lawrence Erlbaum.

Donahue, Mark. 1997. Tone systems in New Guinea. *Linguistic Typology* 1:347–386.

Engstrand, Olle, Karen Williams, and Sven Strömqvist. 1991. Acquisition of the Swedish tonal word accent contrast. *Actes du Congrès International de Sciences Phonétiques* 7:324–327.

Fujisaki, Hiroya. 1983. Dynamic characteristics of voice fundamental frequency in speech and singing. In *The pronunciation of speech*, ed. Peter F. MacNeilage, 39–55. Heidelberg: Springer.

Furrow, David. 1984. Young children's use of prosody. *Journal of Child Language* 11:203–213.

Gårding, Eva. 1983. A generative model of intonation. In *Prosody: Model and measurements*, ed. Anne Cutler and D. Robert Ladd, 11–25. Heidelberg: Springer.

Gårding, Eva. 1987. Speech act and tonal pattern in standard Chinese: Constancy and variation. *Phonetica* 44:13–29.

Goldsmith, John. 1976. Autosegmental phonology. Doctoral dissertation, Massachusetts Institute of Technology.

Hallé, Pierre, Bénédicte de Boysson-Bardies, and Marilyn M. Vihman. 1991. Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by French and Japanese infants. *Language and Speech* 34:299–318.

Halliday, M.A.K. 1979. One child's protolanguage. In *Before speech: The beginning of interpersonal communication*, ed. Margaret Bullowa, 171–190. Cambridge: Cambridge University Press.

Inkelas, Sharon, and Draga Zec. 1988. Serbo-Croatian pitch accent. *Language* 64:227–248.

Ladd, D. Robert. 1983. Phonological features of intonational peaks. *Language* 59: 721–759.

Ladd, D. Robert. 1988. Declination 'reset' and the hierarchical organization of utterances. *Journal of the Acoustic Society of America* 84:530–544.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Leben, William. 1973. Suprasegmental phonology. Doctoral dissertation, Massachusetts Institute of Technology.

Ota, Mitsuhiko. 1999. Phonological theory and the development of prosodic structure: Evidence from child Japanese. Doctoral dissertation, Georgetown University.

Ota, Mitsuhiko. 2003. *The development of prosodic structure in early words: Continuity, divergence and change*. Amsterdam: John Benjamins.

Peters, Ann. M., and Sven Strömqvist. 1996. The role of prosody in the acquisition of grammatical morphemes. In *Signals to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James L. Morgan and Katherine Demuth, 215–232. Mahwah: Lawrence Erlbaum.

Pierrehumbert, Janet B. 1980. The phonology and phonetics of English intonation. Doctoral dissertation, Massachusetts Institute of Technology.

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

Plunkett, Kim, and Sven Strömqvist. 1992. The acquisition of Scandinavian languages. In *The crosslinguistic study of language acquisition* 3, ed. Dan I. Slobin, 457–556. Hillsdale: Lawrence Erlbaum.

Poser, William. 1984. The phonetics and phonology of tone and intonation in Japanese. Doctoral dissertation, Massachusetts Institute of Technology.

Schmid, Beata. 1986. A comparative study of children's and adults' acquisition of tone accents in Swedish. *Language Learning* 36:185–210.

Snow, David. 1998. Children's imitations of intonation contours: Are rising tones more difficult than falling tones? *Journal of Speech, Language and Hearing Research* 41:576–587.

Vihman, Marilyn. 1996. *Phonological development*. Oxford: Blackwell.