Proceedings of the Society for Computation in Linguistics

Volume 5 Article 15

2-1-2022

Representing Multiple Dependencies in Prosodic Structures

Kristine M. Yu
University of Massachusetts - Amherst, krisyu@linguist.umass.edu

Recommended Citation

Yu, Kristine M. (2022) "Representing Multiple Dependencies in Prosodic Structures," *Proceedings of the Society for Computation in Linguistics*: Vol. 5 , Article 15.
DOI: https://doi.org/10.7275/hj4r-d776
Available at: https://scholarworks.umass.edu/scil/vol5/iss1/15

Representing multiple dependencies in prosodic structures

Kristine M. Yu
University of Massachusetts Amherst / Amherst, Massachusetts, USA
krisyu@linguist.umass.edu

Abstract

Association of tones to prosodic trees was introduced in Pierrehumbert and Beckman (1988). This included: (i) tonal association to higher-level prosodic nodes such as intonational phrases, and (ii) multiple association of a tone to a higher-level prosodic node in addition to a tone bearing unit such as a syllable. Since then, these concepts have been broadly assumed in intonational phonology without much comment, even though Pierrehumbert and Beckman (1988)'s stipulation that tones associated to higher-level prosodic nodes are peripherally realized does not fit all the empirical data. We show that peripherally-realized tones associated to prosodic nodes can be naturally represented with bottom-up tree transducers. Additionally, multi bottom-up tree transducers provide a way to represent non-peripheral boundary tones and multiple tonal association, as well as multiple dependencies in prosodic structures in general, including prosodically-conditioned segmental allophony.

1 Introduction

It is widely accepted that describing segmental and tonal distributions and processes over trees built with prosodic constituents (e.g., syllables (σ), feet (Ft), prosodic words (ω), accentual phrases (α), phonological phrases (φ), and intonational phrases (ι)) can help capture phonological generalizations. A classic example exemplifying this comes from Bengali (Hayes and Lahiri, 1991). As exemplified in Fig. 1, adapted from Khan (2008, p. 101)[1], rises in the pitch contour delineate phonological chunks in Bengali. In the example in Fig. 1, these chunks happen to be the size of a morphosyntactic word plus affixes, but chunk size can vary depending on speech rate. For example, Hayes and Lahiri (1991, (54)) provides the example in (1), where we indicate phonological chunks delineated by melodic rises using square brackets. In (1a), one melodic rise occurs per word as in Fig. 1. However, at faster speech rates, a speaker may utter the same sentence with the prosodic chunkings in (1b) or (1c), or at an even faster speech rate, as (1d).

(1) Variation in prosodic domains

a. [ɔmor] [t͡ʃador] [tara-ke] [diet͡ʃʰe]
Amor scarf Tara-obj gave
'Amor *gave* a scarf to Tara'

b. [ɔmot̠͡ʃ t͡ʃador] [tara-ke] [diet͡ʃʰe]

c. [ɔmor] [t͡ʃadot̠ tara-ke] [diet͡ʃʰe]

d. [ɔmot̠͡ʃ t͡ʃadot̠ tara-ke] [diet͡ʃʰe]

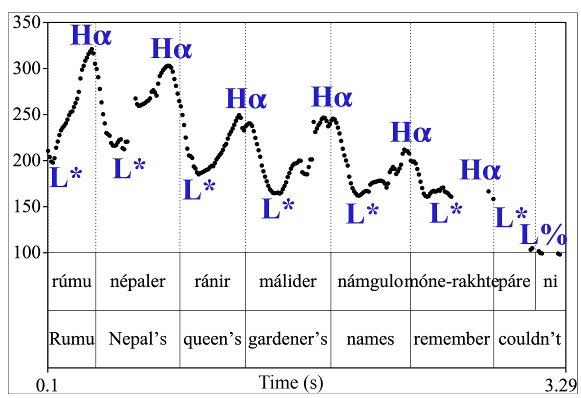


Figure 1: Melodic rises in Bengali analyzed as tonal sequences. Fundamental frequency (Hz) on y-axis, time (s) on x-axis. 'Rumu couldn't remember the names of the gardeners of the queen of Nepal.' Example from Khan (2008, p. 101).

It would be difficult to characterize all of these possible chunkings of the same sentence under the same information structural conditions as being

[1]The Bengali case study presented in this paper is based on Hayes and Lahiri (1991)'s analysis of a Kolkata variety, but we show a pitch track example from Khan (2008)'s analysis of a Bangladeshi variety since recordings from Khan (2008) are readily available.

Proceedings of the Society for Computation in Linguistics (SCiL) 2022, pages 171-183.
Held on-line February 7-9, 2022

morphosyntactically-conditioned. Moreover, the same chunks delineated by melodic rises also determine whether two other segmental processes occur (Hayes and Lahiri, 1991, §§9.1, 9.2): (i) total assimilation of /r/ to an immediately following coronal consonant, and (ii) voicing assimilation of a stop to an immediately following stop. These two segmental processes occur when both the segment that gets changed as well as its conditioning environment occur within the same chunk, as exemplified for the final [r]s in [ɔmor] and [t͡ʃador] in (1), which are underlined when they assimilate to [t͡ʃ] and [t], respectively. (Note: Hayes and Lahiri (1991) calls these prosodic constituents phonological phrases, while Khan (2008) calls them accentual phrases; here we use 'accentual phrase').

The generalization that melodic patterns delineate the edges of prosodic constituents also motivated one of the foundational assumptions of Autosegmental-Metrical (AM) Theory (Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988; Ladd, 1996; Arvaniti and Fletcher, 2020), a theory that dominates work on intonational phonology: the assumption that tones can associate not only to tone-bearing units (TBUs) "at the bottom of the tree" (i.e., non-terminal nodes that immediately dominate terminals) such as moras (μ) and syllables (σ), but also to any higher-level node in the prosodic tree, e.g., the accentual phrase or the intonational phrase (Pierrehumbert and Beckman, 1988, p. 21). While the concept of tones associating to TBUs was carried over from Autosegmental Theory (Goldsmith, 1976), the concept of tones associating to prosodic constituents in general was an innovation of AM theory, as well as the notion that tones can be multiply associated—both to a higher-level prosodic node as well as a TBU (Pierrehumbert and Beckman, 1988).

In Fig. 1, each melodic rise is analyzed as the phonetic realization of a sequence of two discrete tones: a low pitch accent (L*), and a high accentual phrase tone (H_α). The '*' diacritic indicates a pitch accent; the 'α' diacritic indicates an accentual phrase tone. The entire sentence comprises an intonational phrase, with a low intonational phrase tone, L%, at the right edge (the '%' diacritic indicates an intonational phrase tone). In AM Theory, a pitch accent like L* is a tone whose appearance and temporal location are determined by accented TBUs, i.e., TBUs with "an abstract phonological location indicator of tone" (Gussenhoven, To appear, §1.2) and is represented as being associated to an accented TBU. An edge tone like H_α or L% is a tone whose appearance and location is determined by prosodic constituent edges and is represented as being associated to a prosodic node at a higher-level node than the TBU.

The L* appears at the left edge of an accentual phrase, while the H_α appears at the right edge of an accentual phrase. So why is the L defined as a pitch accent rather than an edge tone? In Bengali, accented TBUs are syllables that receive stress, and Bengali has word-initial stress—thus, the L tones are always word-initial in Fig. 1. However, Hayes and Lahiri (1991, p. 56) shows that when a word is preceded by a clitic, the L tone is not phrase-initial and appears instead on the initial syllable of the word, after the clitic—thus tracking the accented TBU rather than the left edge of accentual phrases.

The Bengali example in Fig. 1 exemplifies the distinction between pitch accents and edge tones, but what about the concept of the association of a single tone to both a TBU as well as a higher-level prosodic node? Multiple association of this kind was first motivated by Pierrehumbert and Beckman (1988) for Tokyo Japanese due to differences in the phonetic realization of L_α tones systematically conditioned by the position of lexical accent. In Japanese, accented syllables are lexically specified and receive a bitonal H*+L tone, cf. *hasi* 'edge' vs. *hási* 'chopsticks' vs. *hasí* 'bridge' (Gussenhoven, 2004, p. 186), where accent is indicated with an acute accent mark. The comparison between unaccented *hasi* and initially-accented *hási* is represented in (2) using association of tones to labeled brackets for singly associated accentual phrase and intonational phrase tones, following notational conventions popularized by Hayes and Lahiri (1991). The analysis shown follows Gussenhoven (2004, 2014).

(2) Tonal associations for *hasi* vs. *hási*

$$[_\iota\, [_\alpha\, \text{h a s i}\,]_\alpha\,]_\iota \qquad [_\iota\, [_\alpha\, \text{h á s i}\,]_\alpha\,]_\iota$$

$$L_\alpha\ H_\alpha\ L\% \qquad L_\alpha\ H^*\ L\ \ L\%$$

In words like *hási* where a tone occupies the first TBU, i.e., the first mora, the word-initial L_α is pronounced with a mid pitch, but in words like *hasi* where no lexical accent occupies the first TBU, the L_α is pronounced fully low, see Pierrehumbert and Beckman (1988, §5.5); Gussenhoven (2004, p. 189). This difference is attributed to a difference in association: in *hasi*, the first TBU is available for the L_α to associate and so it associates not only to the α node but also this TBU; in *hási*, the first TBU is unavailable so the L_α is associated only to the α node. Similarly, phonetic evidence shows that the L of the lexical accent associates to an unoccupied TBU immediately following the accented TBU (Gussenhoven, 2014, §2), as shown for *hási* in (2). There is also an H_α following the peripheral (i.e., at the left edge) L_α. In *hasi*, the second TBU is available for the H_α to associate to, but since the second TBU of *hási* is occupied by the L of the lexical accent, the H_α is deleted. Non-peripheral, unassociated tones are deleted in Japanese (Gussenhoven, 2014, §2).

The concepts of association of tones to higher-level prosodic nodes and multiple tonal association introduced in Pierrehumbert and Beckman (1988) have been broadly assumed in intonational phonology without much comment (but see, e.g., Prieto et al. (2005); Gussenhoven (2018) for exceptions). However, while the computational properties of association of tones to TBUs have received much attention, e.g., Chandlee and Jardine (2021) and references therein, the computational properties of tones associating to prosodic trees, i.e., tones as terminals participating in dominance relations in prosodic trees, as well as multiple tonal association to TBUs and prosodic nodes, have not. In fact, as noted in Pierrehumbert (2011, p. 5), prosodic trees with multiple tonal associations are technically not trees anymore, since terminal nodes can have more than one parent.

Moreover, the formal properties of tones associating to prosodic trees defined in Pierrehumbert and Beckman (1988, Ch. 6) have not been revisited, although Pierrehumbert and Beckman (1988, Ch. 6) stipulates the temporal location of tones associated to prosodic nodes (i.e., edge tones, or boundary tones) to be at the periphery of the constituent they are associated to. The stipulation is problematic because Gussenhoven (2000) provides examples from Roermond Dutch where edge tones are not peripherally realized, i.e., a lexical accent tone is sequenced to appear after a right-edge aligned intonational phrase boundary tone. Gussenhoven (2000)'s response to the problematic peripherality stipulation (see also Gussenhoven (2018, §4)) is to abandon the idea of tonal association to higher-level prosodic nodes altogether in favor of Align constraints between tones and prosodic constituents. But the theory of tonal association to higher-level prosodic nodes as proposed in Pierrehumbert and Beckman (1988, Ch. 6) has remained a fundamental assumption of Autosegmental-Metrical Theory (Arvaniti and Fletcher, 2020), despite its inability to allow for non-peripheral prosodic boundary tones.

This paper shows that standard tools from formal language theory can be used to formalize the notion of tonal association to prosodic trees and handle both multiple tonal association and non-peripheral boundary tones. To define tonal association in prosodic trees, we make use of finite state tree rewrite grammars, which can be recognized by bottom-up tree transducers (Baker, 1978; Comon et al., 2007), and in the paper, we use the notation of finite state tree transducers to define our tree grammars (Rounds, 1970). The bottom-up tree transductions provide a natural mechanism for prosodic boundary tones to be sequenced peripherally, without stipulation.

Moreover, we show that a standard extension of bottom-up tree transducers—multi bottom-up tree transducers (mbutts) (Lilin, 1978; Fülöp et al., 2004; Maletti, 2008), see Maletti (2008, §4) for a formal definition—can represent multiple tonal association and allow non-peripheral edge tones. String yields from trees that can be built with finite state bottom-up tree transducers are context-free, i.e., strings that can be derived with CFG grammars (Comon et al., 2007, §2.4). String yields from trees that can be built with multi finite state bottom-up tree transducers are strings that can be derived with multiple CFGs (Engelfriet et al., 2009), grammars that that are more expressive than CFGs, in which one constituent can enter into relationships with two of its ancestors, e.g., in syntactic movement, see Clark (2014).

While mbutts have been used to express syntactic relations (Kobele et al., 2007; Graf, 2012) and also syntax-prosody mapping (Dolatian et al., 2021), we show here—building on Yu (2021)—that mbutts are of interest as representations for phonological phenomena in general. Multiple tonal association is only one instance of multiple dependencies in prosodic trees, but we show that so are prosodically-conditioned segmental processes such as Bengali r-assimilation, and that mbutts can handle these processes as well. The next section, §2, introduces a first tree transduction for single tonal associations in a single word of Bengali. §3 introduces mbutts in tree transductions for tone association in Japanese for *hasi* and *hási* in (2), and §4 shows how mbutts can represent r-assimilation in Bengali, too. §5 discusses issues raised by using mbutt representations.

2 A first tree transduction

A finite state bottom-up tree transducer can be thought of as a generalization of a string finite state transducer that can process multiple branches rather than a single branch (a string). A string finite state transducer processes a string from left to right, one symbol at a time, and enters one of finitely many states after each step. A string transduction is recognized as well-formed if and only if the transducer enters a final state after processing the entire string. A finite state bottom-up tree transducer processes a tree from leaves towards the root, one subtree at a time, and enters one of finitely many states after each step. A tree transduction is recognized as well-formed if and only if the transducer enters a final state after processing the tree all the way up to the root. A tree transduction step can re-label nodes, delete subtrees, or insert new material. However, bottom-up tree transductions cannot change structures that have already been built.

As a first introduction to tree transductions, we show the grammar and steps to insert the pitch accent (L*) and accentual phrase tone (H_α) and assign stress in an accentual (α) phrase of a single two-syllable prosodic word (ω) in Bengali, e.g., /t͡ʃador/ or /ɔmor/ from (1). (An even simpler warm-up transduction that inserts just the H_α and ignores stress and pitch accent assignment is given in Table 5 and (7) in Appendix A.) For this first transduction, we make the simplification that the pitch accent insertion rule in Bengali is only ω-based, i.e., a pitch accent is assigned to the stressed syllable in each ω. A transduction that assigns an L* to the stressed syllable of only an α-initial ω is shown in Appendix B.

Since the segments play no role in these processes, we leave them out and only show tonal association to the syllabic TBUs (σ) and α node. The rules in (3) take the input tree shown as the leftmost tree in the derivation in Table 1 and returns the rightmost tree in Table 1 as the output tree (ignoring the green filled circle at the moment). We assume a lexicon of low and high tones and a placeholder symbol, ε, that indicates a location where a tone can be filled, $\{L, H, \varepsilon\}$, and we define q_α to be a final state. A green filled circle decorating a tree in Table 1 indicates which state the transducer enters after the application of the transition rule labeling the rewrite arrow to the left of the tree, and the output at each step is shown as the subtree under the state. By convention, a state is positioned as the mother node of the subtree that has just been processed, but isn't actually part of the tree—it's just an annotation like a "you are here" marker.

(3) Grammar fragment for tree transduction of single-ω accentual phrase; q_α final state

$$\begin{array}{ll} [B1] & \varepsilon() \rightarrow q_\varepsilon(\varepsilon()) \\ [B2] & \sigma(q_\varepsilon(t)) \rightarrow q_\zeta(t) \\ [B3] & \omega(q_\zeta(t_1), q_\zeta(t_2)) \rightarrow q_\omega(\omega(\mathrm{str}(\sigma(L)), \sigma(t_2))) \\ [B4] & \alpha(q_\omega(t)) \rightarrow q_\alpha(\alpha(t, H)) \end{array}$$

The left-hand side of a rule shows the structure required for the rule to be applied, and its format differs depending on whether the transducer is at a leaf or not. When the transducer is at a leaf, e.g., Rule [B1], the left-hand side of the rule is just the leaf, which by definition, has no daughters underneath—indicated by the empty parentheses following the leaf label, e.g., $\varepsilon()$ in Rule [B1]. If the transducer is at an ε leaf, then Rule [B1] can apply, as shown in the first step in Table 1. The right-hand side shows the state entered, as well as the output, shown in the immediately following parentheses. For example, when the transducer applies Rule [B1], it processes the leaf $\varepsilon()$, enters state q_ε, and returns the input leaf $\varepsilon()$ unaltered, as output. And the first step in Table 1 shows the transducer processing both ε leaves with Rule [B1] (which is shown as applying twice with the notation $[B1]^2$) to enter state q_ε on both the left and right branches and output back ε on each branch, which is shown as the daughter of the q_ε node. The green circles in the derivation move from the leaves towards the root over the course of the derivation since the tree is processed bottom-up. When the transducer is at a non-terminal node (all rules but Rule [B1]), the current node label and the state(s) that the transducer is in must match the left-hand side of a rule for the rule to apply.

Rule [B2] states that if the transducer is at a unary σ node with its single daughter (variable t) in state q_ε, then the transducer can enter q_ζ, deleting the σ node but leaving its daughter (t) unchanged. The second step in Table 1 shows the transducer applying this rule for both the left and right branches; here, t is ε. Rule [B3] is a merge rule that states that if the transducer is at a binary-branching ω node

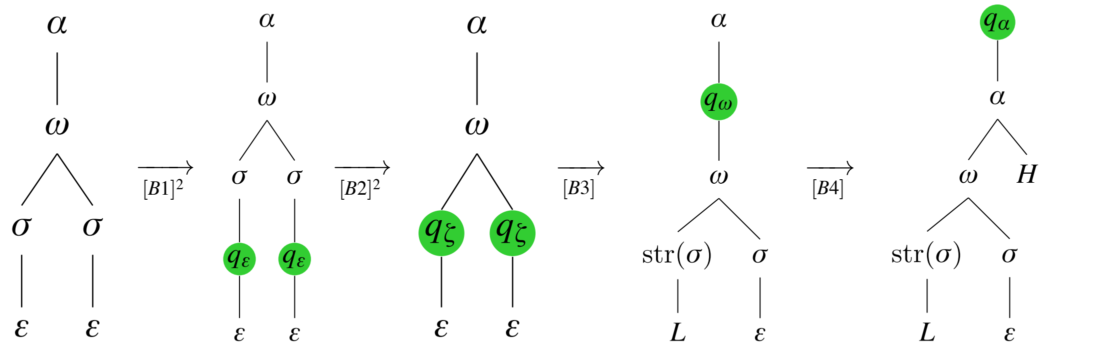


Table 1: Transduction of tone insertion and stress assignment in single-word accentual phrase using rules in (3)

with its left daughter (t_1) in state q_ζ and its right daughter (t_2) also in state q_ζ, then the transducer can enter q_ω and output back the ω subtree with σ nodes inserted above both daughters. Moreover, the σ dominating the first daughter (t_1) is assigned stress, $\mathrm{str}(\sigma)$,[2] following Bengali's ω-initial stress assignment rule, and its associated ε tone is replaced with a L tone, i.e., an L* pitch accent. Since the L pitch accent is already defined by where it is associated in the tree, there is no need to also add a '*' diacritic. The third step in Table 1 shows the transducer applying Rule [B3]. The replacement of ε with L in Rule [B3] is why Rule [B2] is defined to delete the σ node. The bottom-up transducer can modify the daughter tone of a σ node only if the σ node has not already been built. The L* associates to the stressed syllable, which is defined to be the ω-initial syllable. So stress assignment, and consequently pitch accent assignment, can only occur when the ω node is processed.

The transduction of the input tree can end successfully if the transducer completes processing the tree up to the root node and enters a final state—a state where the derivation can optionally terminate. Rule [B4] states that if the transducer is at an α node with a single daughter (t) in state q_ω, then the transducer can enter q_α and output back the α subtree with its daughter (t) unaltered and insert a new daughter H accentual phrase tone to the right. No tonal α diacritic is needed since the H tone is defined by where it associates in the tree. For the purposes of processing just an accentual phrase, we designate q_α as a final state.

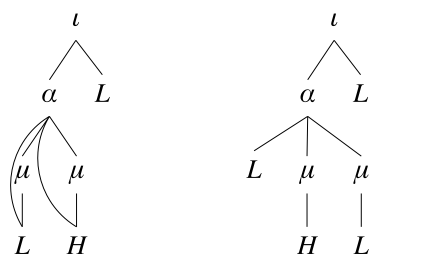
Upon the application of Rule [B4], the transducer has processed the entire tree up to the root, enters final state q_α (positioned as the mother node of the root node), and returns the output tree, which is shown as the daughter of q_α. Thus, the tree grammar in (3) recognizes that the transduction in Table 1 is well-formed. In fact, the transduction in Table 1 is the only transduction that (3) recognizes as well-formed. For instance, an output tree like the rightmost tree in Table 1, but with the second syllable under ω stressed rather than the first, is not well-formed under (3).

The rule to insert an accentual phrase tone in Table 1, Rule [B4], exemplifies how peripherality of tones associated to higher-level prosodic nodes is a natural consequence of the definition of bottom-up tree transducers. Since a bottom-up tree transducer cannot make changes to a subtree that has already been built, no rule in place of Rule [B4] can be defined to insert a tone inside the already-built ω-subtree. Rather, Rule [B4] inserts a tonal daughter of α that is a sister to the right of the already-built ω-subtree. Another possible rule in place of Rule [B4] could insert an H tone to the left of the ω-subtree, e.g., by replacing the right-hand side of Rule [B4], $q_\alpha(\alpha(t, H))$, with $q_\alpha(\alpha(H, t))$. The possible placements of an inserted tone are confined to the periphery of the accentual phrase. Non-peripheral boundary tones can be defined with multi bottom-up tree transductions, which we introduce in the next section.

3 Multiple dependencies in tonal associations

While the association of L* pitch accent in Table 1 is determined at the word-level, it does not have a multiple dependency because the L is inserted only at the step when the ω is processed (Rule [B3])—the L is not carried up the derivation over multi-

[2]We indicate a stressed syllable with $\mathrm{str}(\sigma)$ rather than a diacritic $\acute{\sigma}$ to make it explicit that the σ in $\mathrm{str}(\sigma)$ is copied and that a stressed σ isn't just another symbol with arbitrary relation to σ.



(a) Unaccented *hasi* (b) Accented *hási*

Figure 2: Derived trees showing tonal associations for *hasi* and *hási* described in (2).

ple steps. Multiple dependencies do, however, occur in the tonal associations of two-syllable words in Japanese such as *hasi* and *hási* (2), and mbutts give us a way to express them, as we show in this section. Since the transductions here involve only unary (and no bimoraic) syllables and tonal insertion is not conditioned on the prosodic word, we omit those constituents in the prosodic trees to conserve space. The final output trees to be derived, following (2), are shown in Figure 2. As is done in AM Theory as well as similar syntactic derivations, we explicitly indicate the multiple associations of the L and H tones in *hasi* by showing them each as having two mothers—the mora and the accentual phrase node. These kind of structures, where a single terminal node has two parents—can be interpreted as multidominance structures (Gärtner, 2002). (The multiple dependencies of the tones in accented *hási* are not explicitly represented in the same way in Figure 2 because they occur only in the course of the derivation and not also in the final output derived tree like for *hasi*.)

A transduction for tone assignment in an intonational phrase consisting of an unaccented 2σ word, e.g., *hasi*, is given in Table 2 and (4), and a transduction for tone assignment in *hási* is given in Table 3 and (5). State labels in common with those for the Bengali in §2 shouldn't be taken to identify shared states between the transductions.

The transduction for /hasi/ (Table 2, using the rules in (4)) must define L and H accentual phrase tones, meaning that L and H tones are to be inserted only once the α node is processed, as daughters of α. But these tones are also to be daughters of μ's, which themselves are daughters of α, and there cannot be any modifications to subtrees already built under the α node. Thus, much like in Table 1, after the ε leaves are processed without change (Rule [J1a]), the μ nodes are processed and deleted (Rule [J3a]).

(4) Grammar fragment for /hasi/ transduction; q_ι final state

$$\begin{array}{ll}
[J1a] & \varepsilon() \rightarrow q_\varepsilon(\varepsilon()) \\
[J3a] & \mu(q_\varepsilon(t)) \rightarrow q_E(t) \\
[J4a] & \alpha(q_E(t_1), q_E(t_2)) \rightarrow q_\delta(L, H) \\
[J5a] & q_\delta(t_1, t_2) \rightarrow \\
 & \quad q_\alpha(\alpha(t_1, t_2, \mu(t_1), \mu(t_2))) \\
[J6] & \iota(q_\alpha(t)) \rightarrow q_\iota(\iota(t, L))
\end{array}$$

When the α node is processed in Rule [J4a], the L and H tones are inserted. We could change the right-hand side of Rule [J4a] to $q_\alpha(\alpha(L, H, \mu(L), \mu(H)))$, skip Rule [J5a] entirely, and still generate the output of Rule [J5a]. But then the L tone that is daughter to the α node would have no specified relation to the L tone that is daughter to the left μ node; nor would the two H tones have a specified relation. Having instead the intermediate step of Rule [J4a] as written in (4) first inserts the L and H tones as lexical items and then carries them up the derivation to the next step as separate subtrees, without merging them at the α node.

Rule [J4a], which transitions the transducer to state q_δ, is our first example of a "multi" step—a step that carries multiple subtrees up the derivation rather than just one. The output of Rule [J4a] has a q_δ green circle that is not positioned as a mother node to a constituent because L and H have not been merged to build a constituent. The two daughter subtrees under q_δ in the input to Rule [J5a] also make that rule a "multi" rule. With L and H carried up as separate subtrees, Rule [J5a] builds μ constituents and an α constituent, associates the left daughter under q_δ (t_1) as both the leftmost daughter of the new α node and the daughter to the new leftmost μ node, and associates the right daughter under q_δ (t_2) as both the peninitial daughter of the new α node and the daughter to the new rightmost μ node. These multiple associations are represented with a multidominance structure, as discussed for Figure 2a. To end the derivation, Rule [J6] processes the ι node and adds an L sister to the right of the α subtree under ι.

"Multi" rules appear in the transduction for tonal assignment in *hasi* because the L and H tones each have multiple (two) dependencies in the course of the derivation. They each enter in Rule [J4a], when the α node is processed, but then they also enter relations with mother α and μ nodes in Rule [J5a]. While the output in Figure 2a derived by the *hasi* transduction shows multiple tonal associations, it

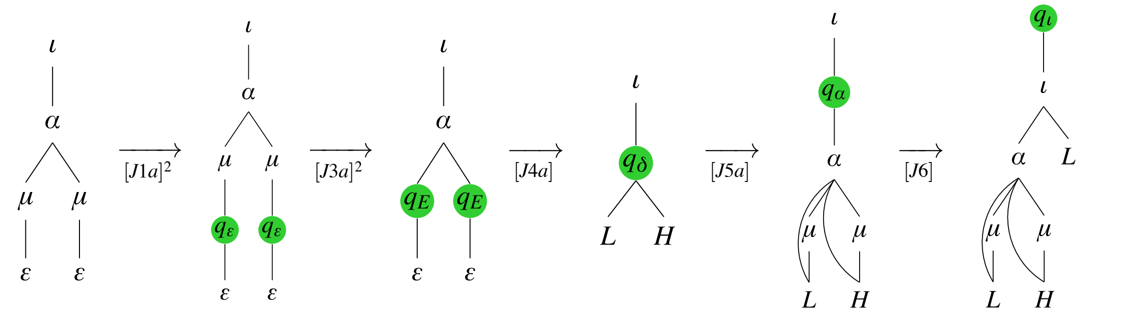


Table 2: Transduction for tonal association in unaccented /hasi/ using rules in (4)

is the multiple dependencies in the derivation steps that we are defining by including "multi" steps. The next transduction we show, which derives the output in Figure 2b, also has "multi" steps.

A transduction for a 2σ word with initial accent, e.g., *hási*, is given in Table 3—following rules already given in (4) and the additional rules in (5). As proposed in Pierrehumbert and Beckman (1988, p. 124-5), a T "tone" constituent is introduced. It keeps the two tones of the H*+L lexical accent as separate leaves so that the tones can dock onto separate TBUs. We assume here that the T node is deleted when the two tones dock onto separate TBUs, although alternative assumptions could be explored as well, see, e.g., Grice (1995).

(5) Grammar fragment for /hási/ transduction, not including rules in (4); q_ι final state

$$
\begin{array}{ll}
[J1b] & H() \rightarrow q_H(H()) \\
[J1c] & L() \rightarrow q_L(L()) \\
[J2] & \mathrm{T}(q_H(t_1), q_L(t_2)) \rightarrow q_A(t_1, t_2) \\
[J3b] & \mu(q_A(t_1, t_2)) \rightarrow q_B((t_1, t_2)) \\
[J4b] & \alpha(q_B(t_1, t_2), q_E(t_3)) \rightarrow \\
& \quad q_\alpha(\alpha(L, \mu(t_1), \mu(t_2)))
\end{array}
$$

Since accent in Japanese is lexical, unlike the Bengali pitch accent, the input tree to the transduction already has the first mora associated to a *T* subtree with *H* and *L* daughters, i.e., a lexical accent. The tonal leaves enter via Rules [J1a,b,c]. "Multi" Rule [J2] processes the *T* node and deletes it to expose the tonal daughters for tonal re-association, carrying the *H* and *L* leaves separately up the derivation. "Multi" Rule [J3b] processes the left μ node, deletes it, and continues to carry the *H* and *L* leaves up the derivation. Rule [J3a] processes the right μ node and deletes it. The *H* and *L* leaves have been carried up separately via the "multi" rules up to this point so that they can associate to separate moras in the next step. Rule [J4b] then shifts the *L* to the right branch to replace the placeholder ε, rebuilds the moras, and inserts an *L* accentual phrase tone as the leftmost sister to the μ's. Finally, Rule [J6] processes the ι node and inserts an *L* to the right of the α subtree as a daughter of ι, just like in the transduction for *hasi*.

The final output tree from Table 3 has no multiple tonal associations. Nevertheless, the transducer defined in (5) is an mbutt because "multi" steps arise from multiple dependencies in the derivation steps. Each tone of the lexical accent has two dependencies: (i) to the *T* node, where it enters as a daughter leaf, and (ii) to the μ node, where it re-associates as as daughter leaf. Note that the last "multi" step, Rule [J4b], can be easily modified to demonstrate how "multi" steps can accommodate non-peripheral temporal sequencing of edge tones. For example, the right-hand side could be changed to $q_\alpha(\alpha(\mu(t_1), L, \mu(t_2)))$ to insert the L_α between the two lexical accent tones.

4 Multiple dependencies in segmental associations

Multiple dependencies don't occur only with tonal association in prosodic trees, but also with prosodically-conditioned segmental processes. A segment enters as a leaf (location one in a prosodic tree) but then it cannot be merged into the prosodic tree until the prosodic constituent that conditions its realization is processed (location two). We illustrate this for the transduction of /r/-assimilation in the Bengali accentual phrase /t͡ʃador/, shown in Table 4, following (6). Recall from §1 that /r/ undergoes total assimilation to an immediately following coronal consonant within the same α. Therefore, the realization of any /r/ in Bengali can't be determined until α is processed. Moreover, any coronal consonant must also be carried up all the way to the α node, in case it may be immediately preceded by an /r/ within the same α.

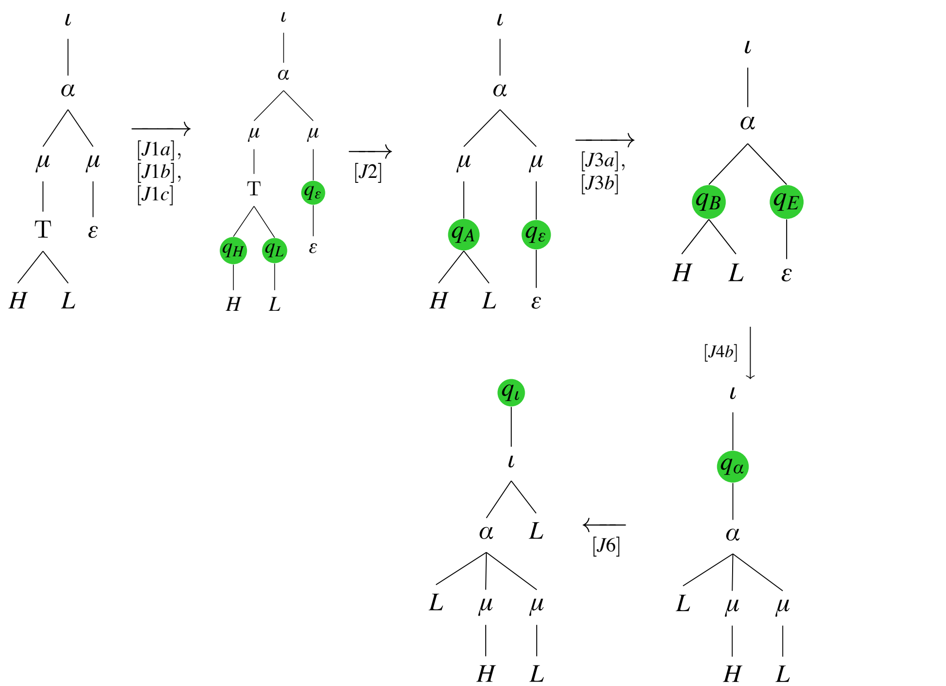


Table 3: Transduction for tonal association in accented /hási/ using rules in (5)

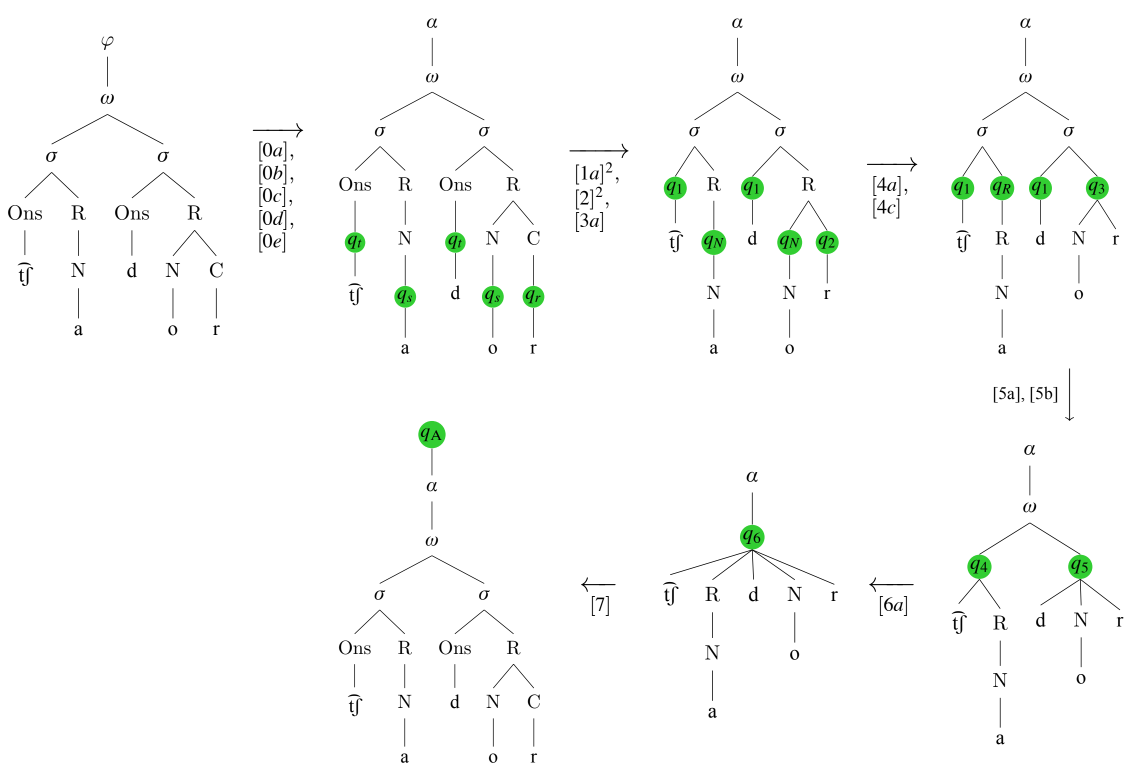


Table 4: Transduction for /r/-assimilation in /t͡ʃador/ using rules in (6)

(6) Grammar fragment for /t͡ʃador/; q_{A} final state

$$
\begin{array}{llll}
[0a] & \widehat{\mathrm{t\int}}() \rightarrow q_t(\widehat{\mathrm{t\int}}()) & [0b] & \mathrm{a}() \rightarrow q_s(\mathrm{a}()) \\
[0c] & \mathrm{d}() \rightarrow q_t(\mathrm{d}()) & [0d] & \mathrm{o}() \rightarrow q_s(\mathrm{o}()) \\
[0e] & \mathrm{r}() \rightarrow q_r(\mathrm{r}()) & [1a] & \mathrm{Ons}(q_t(t)) \rightarrow q_1(t) \\
[2] & \mathrm{N}(q_s(t)) \rightarrow q_N(\mathrm{N}(t)) & [3a] & \mathrm{C}(q_r(t)) \rightarrow q_2(t)
\end{array}
$$

$$
\begin{array}{ll}
[4a] & \mathrm{R}(q_N(t)) \rightarrow q_R(\mathrm{R}(t)) \\
[4c] & \mathrm{R}(q_N(t_1), q_2(t_2)) \rightarrow q_3(t_1, t_2) \\
[5a] & \sigma(q_1(t_1), q_R(t_2)) \rightarrow q_4(t_1, t_2) \\
[5b] & \sigma(q_1(t_1), q_3(t_2, t_3)) \rightarrow q_5(t_1, t_2, t_3) \\
[6a] & \omega(q_4(t_1, t_2), q_5(t_3, t_4, t_5)) \rightarrow q_6(t_1, t_2, t_3, t_4, t_5) \\
[7] & \alpha(q_6(t_1, t_2, t_3, t_4, t_5)) \rightarrow \\
 & q_{\mathrm{A}}(\alpha(\omega(\sigma(O(t_1), t_2), \sigma(O(t_3), \mathrm{R}(t_4, \mathrm{C}(t_5))))))
\end{array}
$$

The transduction in Table 4 begins by processing each segmental leaf via Rules [0a-e] to enter one of three states on each branch: q_t (coronals /t͡ʃ, d/), q_r (/r/), or q_s (vowels /a, o/). Then, each of the two branches with a nucleus (N) node with a daughter in state q_s can be processed to fix the realization of the segment and finish building the nucleus to enter state q_N (Rule [2]). However, any branch with a coronal or /r/ is processed to delete the onset (Ons) or coda (C) node and carry up the segment via Rules [$1a, 3a$]. A unary rime (R) can then be built (Rule [4a]), since it doesn't have coronals or /r/. But the /r/ must continue to be passed up, so we delete its *R* mother node and then hold the /r/ together with a nucleus subtree without merging to enter state q_3 ("multi" Rule [4c]).

The coronals and /r/ (and already-built nuclei and rime) continue to be passed up as the σ nodes are processed and deleted ("multi" Rules [5a,b]) to reach states q_4 on the left branch (carrying up 2 subtrees) and q_5 on the right (carrying up 3 subtrees). Similarly, the ω node is then processed and deleted and the coronals and /r/ (and already-built nuclei and rime) are passed up again to enter q_6 with 5 subtree daughters ("multi" Rule [6a]). Finally, we are ready to process the α node and can stop passing up the coronals and /r/. Rule [7] (with Ons abbreviated as O) processes the α node and outputs an α tree with the remaining prosodic structure built in a single step, including no change to the final /r/, since the /r/ has no coronal sister to the immediate right under state q_6. Although /r/-assimilation does not apply when /t͡ʃador/ is in its own α, the transduction we just stepped through underscores that even if a coronal does not immediately follows an /r/ within the same α and even if an /r/ does not immediately precede a coronal within the same α, the dependency to the α node for these types of segments is always there. Appendix §C shows the tree transduction for /ɔmor t͡ʃador/ → [ɔmot͡ʃ t͡ʃador] when /ɔmor t͡ʃador/ is within the same α as in (1b,d).

5 Conclusion

We've shown that tree grammars defined via bottom-up tree transductions—standard and well-studied tools from formal language theory—provide a way to represent tonal association to higher-level nodes in prosodic trees. The peripherality of prosodic boundary tones follows without further stipulation (unlike Pierrehumbert and Beckman (1988)), since bottom-up tree transductions cannot change structures that have already been created. Extension to mbutts provides a mechanism to define non-peripheral boundary tones, which cannot be handled by Pierrehumbert and Beckman (1988). Since non-peripheral boundary tones such as Gussenhoven (2000)'s case in Roermond Dutch appear to be typologically rare, it seems desirable that non-peripheral boundary tones come in with the additional expressivity of "multi" steps in the grammar. More generally, mbutts can represent the pervasive multiple dependencies in prosodic structures including those arising in tonal association and from prosodically-conditioned segmental allophony. They offer a way to precisely state and probe proposals in phonological analyses of tone and intonation, at a time when the fundamental assumptions of AM theory are being revisited (Grice, 2021).

(M)butts are a good starting point since their computational properties are relatively well-understood, but the sample transductions shown here already reveal issues with using them for phonology. For one, the transductions exemplified here only define bounded structures, e.g., two-syllable prosodic words. We can introduce recursion into the grammar to build words and phrases of arbitrary length (Yu, 2021), but it remains to be seen how resulting self-embedded structures fit with phonological patterns. Another issue is that the restriction that (m)butts cannot modify already-built structure—while potentially desirable for making non-peripheral boundary tones possible but exceptional—results in mass deletion of structure in the derivation followed by rebuilding this structure in a single step. Much more work is needed to refine, restrict, and adapt (m)butts to capture and identify generalizations about prosodic structure.

References

Amalia Arvaniti and Janet Fletcher. 2020. The Autosegmental-Metrical theory of intonational phonology. In Carlos Gussenhoven and Aoju Chen, editors, *The Oxford Handbook of Language Prosody*, chapter 6, pages 78–95. Oxford University Press.

Brenda Baker. 1978. Tree transducers and tree languages. *Information and control*, 37:241–266.

Jane Chandlee and Adam Jardine. 2021. Input and output locality and representation. *Glossa*, 6(1):43.

Alexander Clark. 2014. An introduction to multiple context-free grammars for linguists. Available at: `http://www.cs.rhul.ac.uk/home/alexc/lot2012/mcfgsforlinguists.pdf`.

H. Comon, M. Dauchet, R. Gilleron, C. Löding, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. 2007. Tree automata techniques and applications. Available on: `http://www.grappa.univ-lille3.fr/tata`. Release October, 12th 2007.

Hossep Dolatian, Aniello De Santo, and Thomas Graf. 2021. Recursive prosody is not finite-state. In *Proceedings of SIGMORPHON 2021*.

Joost Engelfriet, Eric Lilin, and Andreas Maletti. 2009. Extended multi bottom–up tree transducers: Composition and decomposition. *Acta Informatica*, 46(8):561–590.

Zoltán Fülöp, Armin Kühnemann, and Heiko Vogler. 2004. A bottom-up characterization of deterministic top-down tree transducers with regular look-ahead. *Information Processing Letters*, 91:57–67.

Hans-Martin Gärtner. 2002. *Generalized Transformations and Beyond*. Akademie Verlag, Berlin.

John Anton Goldsmith. 1976. *Autosegmental phonology*. Ph.D. thesis, Massachusetts Institute of Technology.

Thomas Graf. 2012. Movement-generalized Minimalist Grammars. In *LACL 2012*, volume LNCS 7351, pages 58–73.

Martine Grice. 1995. Leading tones and downstep in English. *Phonology*, 12(2):183–233.

Martine Grice. 2021. Commentary: The autosegmental-metrical model of intonational phonology. In Stefanie Shattuck-Hufnagel and Jonathan A. Barnes, editors, *Prosodic Theory and Practice*. MIT Press, Cambridge, MA.

Carlos Gussenhoven. 2000. The boundary tones are coming: on the nonperipheral realization of boundary tones. In Michael B. Broe and Janet B. Pierrehumbert, editors, *Papers in Laboratory Phonology V: Acquisition and the lexicon*, pages 132–151. Cambridge University Press, Cambridge, UK.

Carlos Gussenhoven. 2004. *The phonology of tone and intonation*. Cambridge University Press, Cambridge, UK.

Carlos Gussenhoven. 2014. On the privileged status of boundary tones: evidence from Japanes, French, and Cantonese English. In Eugene Buckley, Thera Crane, and Jeff Good, editors, *Revealing structure*, pages 1–13. CSLI Publications.

Carlos Gussenhoven. 2018. Prosodic typology meets phonological representations. In Larry M. Hyman and Frans Plank, editors, *Phonological typology*, pages 389–418. Walter de Gruyter GmbH, Berlin/Boston.

Carlos Gussenhoven. To appear. Just how metrical is the Autosegmental-Metrical model? Evidence from pitch accents in Nubi, Persian, and English. In Haruo Kubozono, Junko Ito, and Armin Mester, editors, *Prosody and Prosodic Interfaces*. Oxford University Press.

Bruce Hayes and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language & Linguistic Theory*, 9:47–96.

Sameer ud Dowla Khan. 2008. *Intonational phonology and focus prosody of Bengali*. Ph.D. thesis, University of California Los Angeles.

Gregory M. Kobele, Christian Retoré, and Sylvain Salvati. 2007. An automata-theoretic approach to minimalism. In J. Rogers and S. Kepser, editors, *Model Theoretic Syntax at 10*, pages 71–80.

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

Eric Lilin. 1978. *Une generalisation des transducteurs d'etats finis d'arbres: les S-transducteurs*. Thése 3éme cycle, Université de Lille.

Andreas Maletti. 2008. Compositions of extended top-down tree transducers. *Information and computation*, 206:1187–1196.

Janet Pierrehumbert. 2011. Representation in phonology. Presented at 50 Years of Linguistics at MIT.

Janet Pierrehumbert and Mary Beckman. 1988. *Japanese Tone Structure*. The MIT Press.

J.B. Pierrehumbert. 1980. *The phonology and phonetics of English intonation*. Ph.D. thesis, Massachusetts Institute of Technology.

Pilar Prieto, Mariapaola D'Imperio, and Barbara Gili Fivela. 2005. Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech*, 48:359–396.

William C. Rounds. 1970. Mappings and grammars on trees. *Mathematical Systems Theory*, 4(3):257–287.

Kristine M. Yu. 2021. Computational perspectives on phonological constituency and recursion. *Catalan Jouurnal of Linguistics*, 20:77–114.

A Warm-up tree transduction for only accentual phrase tone insertion

This "warm-up" bottom-up tree transduction inserts only an accentual phrase tone in a single word-accentual phrase in Bengali, while ignoring stress assignment and pitch accent assignment. We include it to show an example of a bottom-up tree transduction that only inserts material in building the output tree. In contrast, the transduction given in Table 1 and (3) includes the step of Rule [B2] which deletes the σ node.

(7) Grammar fragment for tree transduction of single-ω accentual phrase without pitch accent; q_α final state

$$\begin{array}{lr} [B1] & \varepsilon() \rightarrow q_\varepsilon(\varepsilon()) \\ [B2a] & \sigma(q_\varepsilon(t) \rightarrow q_\sigma(\sigma(t)) \\ [B3a] & \omega(q_\sigma(t_1), q_\sigma(t_2)) \rightarrow q_\omega(\omega(t_1, t_2)) \\ [B4] & \alpha(q_\omega(t)) \rightarrow q_\alpha(\alpha(t, H)) \end{array}$$

Rules [B1, B4] are already discussed in §2 and we do not repeat discussion of them here. Rule [B2a] states that if the transducer is at a unary σ node with its single daughter (the variable t) in state q_ε, then the transducer can process the σ node and enter state q_σ, leaving its daughter (the variable t, and in this case, ε) unchanged. The second step in Table 5 shows the transducer applying Rule [B2a] for both the left and right branches. Rule [B3a] is a merge rule that states that if the transducer is at a binary-branching ω node with its left daughter (t_1) in state q_σ and its right daughter (t_2) also in state q_σ, then the transducer can process the ω node to enter q_ω and output back the ω subtree without any change to daughters t_1, t_2.

B Tree transduction for tonal association in two-ω accentual phrase

Table 1 and (3) in §2 made the simplification that the pitch accent insertion rule in Bengali is only ω-based, i.e., a pitch accent is assigned to the stressed syllable in each ω. But in fact, pitch accent assignment is both ω- and accentual-phrase based, i.e., a pitch accent is only assigned to the stressed syllable of an α-initial ω. The rules in (8) and the tree transduction in Table 6 show one way this can be done. The leftmost ω that is pitch-accented is built with steps in Table 1, but the rightmost ω, which is unaccented, uses another rule, Rule [B3b]. Note that since Rules [B3] and [B3b] share the same left hand side, there is non-determinism in the grammar and either of the rules could apply when an ω node is processed in the second step of the transduction. However, Rule [B4b] restricts well-formed two-ω accentual phrases to being initially accented.

(8) Grammar fragment for tree transduction of two-ω accentual phrase, repeating rules already in (3); q_α final state

$$\begin{array}{lr} [B1] & \varepsilon() \rightarrow q_\varepsilon(\varepsilon()) \\ [B2] & \sigma(q_\varepsilon(t)) \rightarrow q_\zeta(t) \\ [B3] & \omega(q_\zeta(t_1), q_\zeta(t_2)) \rightarrow q_\omega(\omega(\mathrm{str}(\sigma(L)), \sigma(t_2))) \\ [B3b] & \omega(q_\zeta(t_1), q_\zeta(t_2)) \rightarrow q_\Omega(\omega(\mathrm{str}(\sigma(t_1)), \sigma(t_2))) \\ [B4b] & \alpha(q_\omega(t_1), q_\Omega(t_2)) \rightarrow q_\alpha(\alpha(t_1, t_2, H)) \end{array}$$

C Tree transduction for /r/-assimilation in /ɔmor t͡ʃador/ → [ɔmot͡ʃ t͡ʃador]

Some rules below are repeated from the rules for the /t͡ʃador/ transduction in (6).

First, we show the transduction for a single-word accentual phrase for /ɔmor/ in Table 7, using the rules in (9). Like /t͡ʃador/, /ɔmor/ has a /r/ that needs to be passed up to the α node.

(9) Grammar fragment for /ɔmor/ transduction; q_α final state;

$$\begin{array}{lr} [0d] & \mathrm{o}() \rightarrow q_s(\mathrm{o}()) \\ [0e] & \mathrm{r}() \rightarrow q_r(\mathrm{r}()) \\ [0h] & \text{ɔ}() \rightarrow q_s(\text{ɔ}()) \\ [0i] & \mathrm{m}() \rightarrow q_s(\mathrm{m}()) \\ [1] & \mathrm{Ons}(q_s(t)) \rightarrow q_O(\mathrm{Ons}(t)) \\ [2] & \mathrm{N}(q_s(t)) \rightarrow q_N(\mathrm{N}(t)) \\ [3a] & \mathrm{C}(q_r(t)) \rightarrow q_2(t) \\ [4a] & \mathrm{R}(q_N(t)) \rightarrow q_R(\mathrm{R}(t)) \\ [4c] & \mathrm{R}(q_N(t_1), q_2(t_2)) \rightarrow q_3(t_1, t_2) \\ [5c] & \sigma(q_O(t_1), q_3(t_2, t_3)) \rightarrow q_7(t_1, t_2, t_3) \\ [5d] & \sigma(q_R(t)) \rightarrow q_\sigma(\sigma(t)) \\ [6c] & \omega(q_\sigma(t_1), q_7(t_2, t_3, t_4)) \rightarrow q_9(t_1, t_2, t_3, t_4) \\ [7b] & \alpha(q_9(t_1, t_2, t_3, t_4)) \rightarrow \\ & q_\alpha(\alpha(\omega(\mathrm{str}(t_1), \sigma(\mathrm{Ons}(t_2), \mathrm{R}(t_3, \mathrm{C}(t_4)))))) \end{array}$$

Putting together the tranductions of /ɔmor/ and /t͡ʃador/ in Tables 7 and 4 up through the penultimate steps, we can define the transduction of /r/-assimilation in the single accentual phrase /ɔmor t͡ʃador/ in Table 8 with the additional rule in (10).

(10) Grammar fragment for transduction of /r/-assimilation in /ɔmor t͡ʃador/; q_α final state

$$\begin{array}{l} [7c]\ (q_9(t_1, t_2, t_3, t_4), q_6(t_5, t_6, t_7, t_8, t_9)) \rightarrow \\ \qquad q_\alpha(\alpha(\omega(t_1, \sigma(t_2, \mathrm{R}(t_3, \mathrm{C}(t_5)))), \\ \qquad \omega(\sigma(\mathrm{Ons}(t_5), t_6), \sigma(\mathrm{Ons}(t_7), \mathrm{R}(t_8, \mathrm{C}(t_9)))))) \end{array}$$

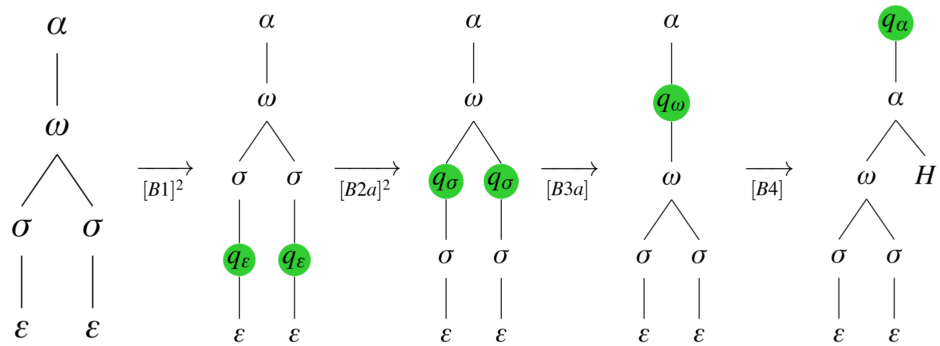


Table 5: Transduction of accentual phrase tone insertion in single-word accentual phrase using rules in (7)

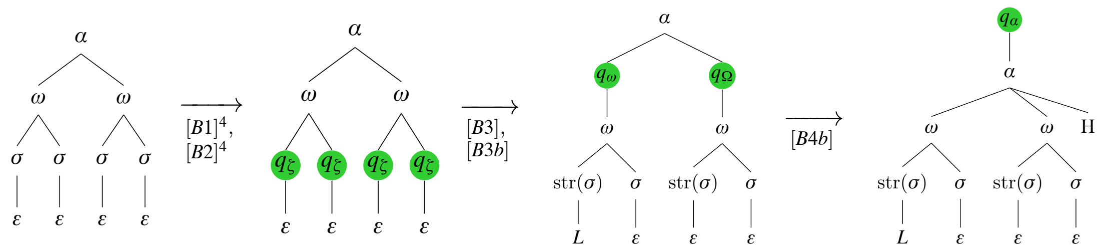


Table 6: Transduction of tone insertion and stress assignment in two-word accentual phrase using rules in (8)

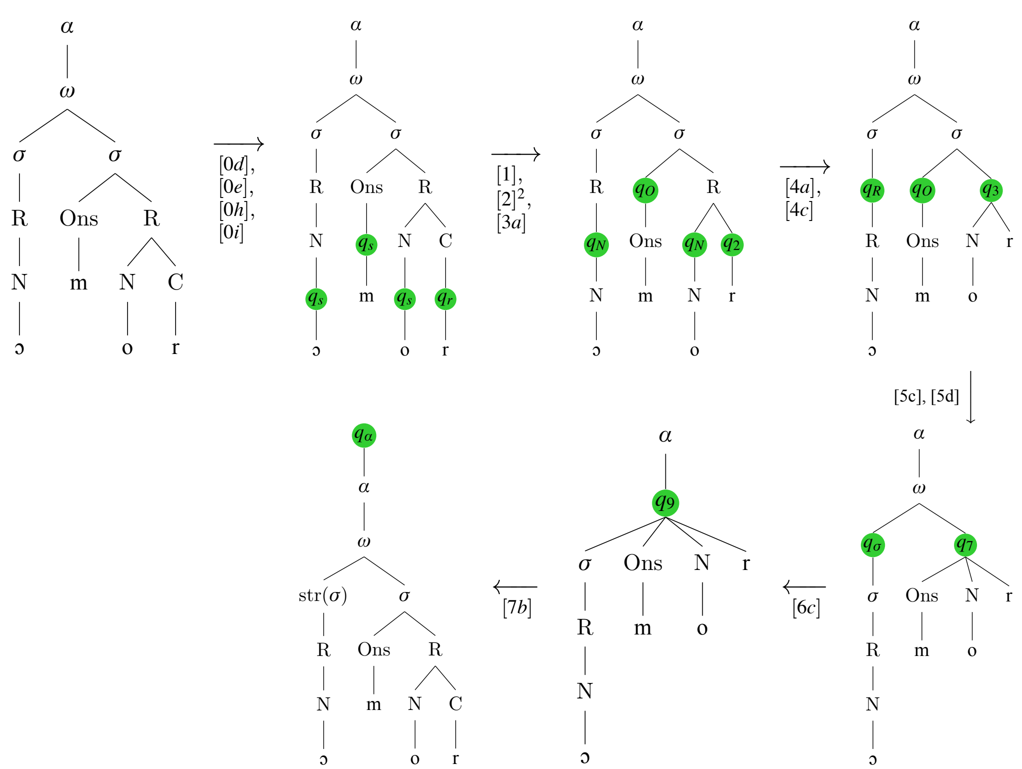


Table 7: Transduction for /r/-assimilation in /ɔmor/ using rules in (9)

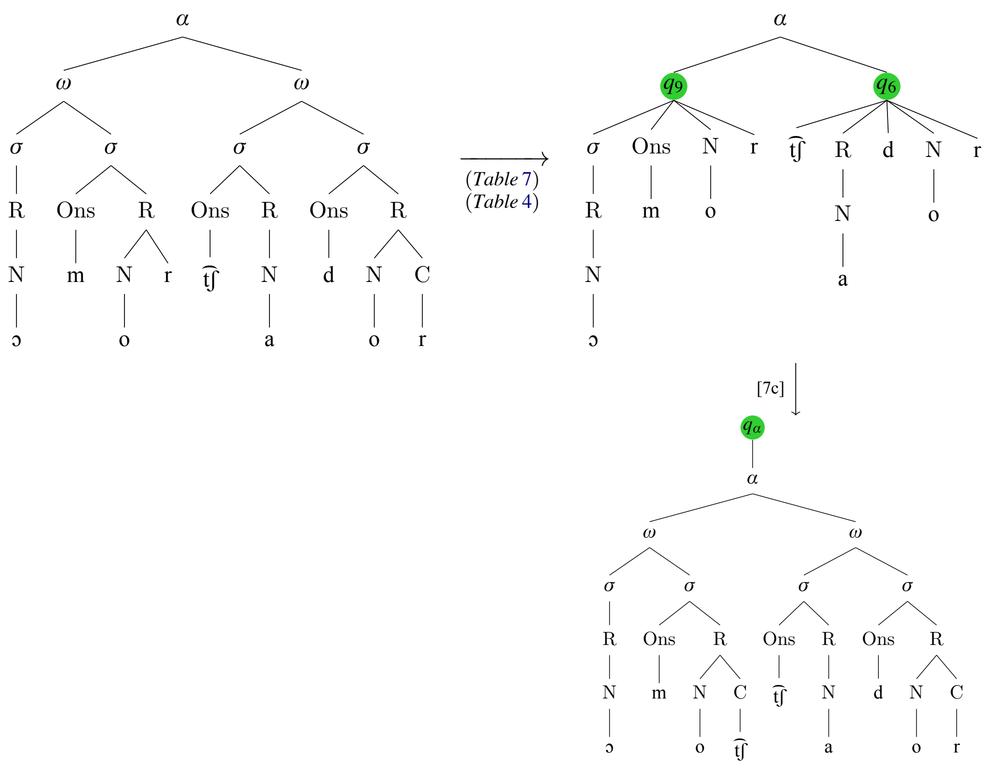


Table 8: Transduction for /r/-assimilation in /ɔmor t͡ʃador/ using rules in (10)