Binarity, branchingness and size effects

Jennifer Bellik and Nick Kalivoda
University of California, Santa Cruz

Constraints on Binarity are commonly used to capture **size effects**: the tendency for longer strings to be parsed into more prosodic constituents (e.g., Inkelas & Zec 1990, Ito & Mester 1992, Sandalo & Truckenbrodt 2002, Prieto 2007, Selkirk 2011, Elfner 2012). Implementations of Binarity come in two major flavors, which we refer to as **branch-counting** (1) and **leaf-counting** (2):

(1) **Bin-Br(K)**: Assign a violation for every node of category K with more than two branches (immediate children).

(2) **Bin-Lv(K, L)**: Assign a violation for every node of category K that dominates more than two nodes of category L at any level, where L<K.

We show that **only branch-counting binarity (1) captures size effects as intended**.

Branch-counting motivates size-driven recursion. Branch-counting binarity (1) assigns a violation to nodes that branches into more than two immediate children. It motivates the selection of candidates that have more prosodic structure. For example, glottal accent ('stød') diagnoses the right edge of a prosodic word in Danish, revealing length-driven differences in compound phrasing: $_{\omega}$[toːg $_{\omega}$[passageːˀr]] 'train passenger' but $_{\omega}$[$_{\omega}$[passageːˀr] $_{\omega}$[toːˀg]] 'passenger train' (Ito & Mester 2015). We capture this with the ranking in (3). NonRecursivity (NonRec) penalizes recursive ωs and outranks Match, ruling out the perfectly isomorphic candidate (3b) and favoring flat structures (3c). High-ranked Bin-Br compels the building of a recursive ω (3a) when a flat structure would create a prosodic word with more than two branches (3c). Bin-Lv does not have the same effect, since the maximal ω still dominates three feet even with recursive structure.

(3) $_{X0}$[toːˀg] $_{X0}$[passageːˀr]	Bin-Br	NonRec	Match	Bin-Lv(Ft)
a. ☞ $_{\omega}$[(toːg) $_{\omega}$[(passa)(geːˀr)]]		*	*	*
b. $_{\omega}$[$_{\omega}$[(toːˀg)] $_{\omega}$[(passa)(geːˀr)]]		**!		*
c. $_{\omega}$[(toːg)(passa)(geːˀr)]	*!		**	*

A similar interaction occurs in Irish phrasing (Elfner 2012). StrongStart (Selkirk 2011) penalizes $_{\varphi}$(ω φ ...) structures, and outranks Match. But in a five-word sentence (4), a StrongStart violation is tolerated (4a) in order to avoid a Bin-Br violation (4c). Bin-LvBinarity must be assessed by counting branches for φ_1 to avoid a binarity violation, since it dominates five leaves (ωs).

(4) $_{\Sigma P}$[V$_{TP}$[$_{DP}$[N Adj] $_{VP}$[$_{DP}$[N Adj]]]]	Bin-Br	StrongStart	Match	Bin-Lv(ω)
a. ☞ $_{\varphi 1}$(V $_{\varphi 2}$($_{\varphi 3}$(N Adj) $_{\varphi 4}$(N Adj)))		*	*	**
b. $_{\varphi 1}$(V $_{\varphi 2}$(N $_{\varphi 3}$(Adj) $_{\varphi 4}$(N Adj)))		**!		**
c. $_{\varphi 1}$($_{\varphi 2}$(V N Adj) $_{\varphi 3}$(N Adj))	*!		**	**

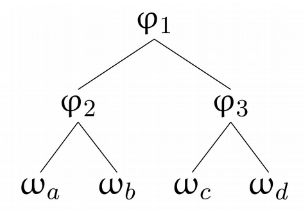
Similar examples are found in Kimatuumbi (Kalivoda 2018) and Mandarin (Shih 2017). Branch-counting (but not leaf-counting) motivates a closer syntax-prosody match.

Leaf-counting motivates size-driven category change. Leaf-counting binarity counts dominated nodes of a lower prosodic category (cf. Dresher & van der Hulst 1998). When Strict Layering is observed (5a-b), counting branches (1) and counting children of the next lower prosodic category (2) are equivalent. In effect, branch-counting binarity is already employed in analyses with Strict Layering (e.g. Shih 2017, Prieto 2007, Sandalo & Truckenbrodt 2002). But when recursion (5c) or level skipping (5d) are permitted, branch-counting and leaf-counting pull apart. In (6)=(5c), φ_1 has only two branches (to φ_2 and φ_3) and satisfies Bin-Br, but violates Bin-Lv since it dominates four leaves ($\omega_{a,b,c,d}$). Since recursive structure does not alleviate Bin-Lv, leaf-counting binarity cannot motivate the size-driven recursion seen Danish (3) and Irish (4).

(5) Comparing Binarity implementations

$[_{xp}w\ [_{xp}w\ [_{xp}w\ w]]]$	Bin-ω (2)	Bin-Br (1)
a. $(_{\iota}(_{\varphi}\ \omega\ \omega\ \omega\ \omega))$	*	*
b. $(_{\iota}(_{\varphi}\ \omega\ \omega)\ (_{\varphi}\ \omega\ \omega))$		
c. $(_{\varphi}\ (_{\varphi}\ \omega\ \omega)\ (_{\varphi}\ \omega\ \omega))$	*	
d. $(_{\iota}\ \omega\ \omega\ \omega\ \omega)$		

(6)



However, leaf-counting can motivate a size-driven change of prosodic category. In (5b,d), the violation of Bin-Lv(φ,ω) is avoided by punting the binarity violation up to the level of ι, so that (5b,d) outperform (5a,c). Bin-Br does not distinguish (5b,d) from (5c). Such a category-promotion can be seen in Japanese compound phrasing (Ito & Mester 2013). Bin-Lv(ω, [Ft, σ]) causes a compound of the form [[Ft][σ Ft]] to be rooted in φ, rather than ω as for smaller compounds. Since this is the only use of Bin-Lv we have found that is not reducible to Bin-Br plus other prosodic constraints, we suggest **restricting Bin-Lv to counting rhythmic categories (σ, Ft)**, such that only Bin-Br counts interface categories (ω, φ, ι). This hypothesis is motivated by the following disadvantages of leaf-counting binarity compared to branch-counting binarity:

- *Leaf-counting is more computationally complex*. While branch-counting only examines a node's immediate children, leaf-counting involves searching through theoretically unbounded levels of recursion until all nodes of a lower prosodic category are found.
- *Leaf-counting predicts larger typologies*. In a case study of Kinyambo phrasing, leaf-counting systems predicted typologies that were, on average, 80% larger than branch-counting typologies, with an average of 10.67 more languages (Bellik & Kalivoda 2016). Moreover, if Bin-leaves were top-ranked, syntaxes like [a[b[c d]], [[ab][cd]] and [[[a]b]c]d] would all be optimally parsed as $_{\iota}(\ _{\varphi}(a\ b)\ _{\varphi}(c\ d))$. To our knowledge, this is unattested. In contrast, Bin-Br cannot compel syntax-prosody mismatches since the syntactic input is already binary-branching.
- *Leaf-counting binarity is redundant*. Call phrasings that violate Bin-Lv but not Bin-Br **leaf-violators**. In an examination of all phrasings of three-word strings, all leaf-violators incur additional penalties from other well-established prosodic well-formedness constraints, compared to all non-leaf-violating possible optima.

Selected References. **Elfner, E.** 2012. *Syntax-Prosody Interactions in Irish*. University of Massachusetts - Amherst dissertation. **Ito, J. & A. Mester.** 1992. Weak Layering and Word Binarity. Linguistic Research Center, LRC-92-09, University of California, Santa Cruz. **Ito, J. & A. Mester** 2013. Prosodic subcategories in Japanese. *Lingua* 124, pp. 20-40. **Ito, J. & A. Mester.** 2015. The perfect prosodic word in Danish. *Nordic Journal of Linguistics*, Vol. 38, No. 1, pp. 5-36. **Kalivoda, N.** 2018. *Syntax-Prosody Mismatches in Optimality Theory*. Ph.D. Thesis, UC Santa Cruz. **Selkirk, E.** 2011. The syntax-phonology interface. In J. Goldsmith, J. Riggle & A. Yu (eds.) *The Handbook of Phonological Theory*, 2nd edition.