The acquisition of CCV branching onsets in Brazilian Portuguese: Revisiting the roles of variation and phonological density

This work aims to discuss the acquisition of CCV (Consonant$_1$ + Consonant$_2$ + Vowel) branching onsets in Brazilian Portuguese (BP) by casting light on the roles of linguistic variation and phonological density into the syllabic development. The goal of the study is to better understand why the CCV acquisition process in BP takes so long compared to other Latin languages like European Portuguese, French or Spanish. In BP, children prove to be able to fully articulate the branching onsets by the age of 4;0 years old, and yet CCVs are variably replaced by CVs (like in /**pɾe**to/ → [ˈ**pe**.tʊ] 'black') or C/ɾ/Vs are replaced by C[l]Vs (like in /p**ɾ**eto/ → [ˈp**l**e.tʊ]) until 5;0 to 6;0 years old. We defend that this variability is tied to the low phonological density and the contextual variability present in the input – a hypothesis that was not yet discussed by the BP phonological development literature. The research works with experimental data (picture-naming task and mispronunciation detection task) and a corpora study comparing adult speech, child directed speech and child speech. The theoretical framework used to model the roles of linguistic variation and phonological density into the syllabic development is the Tolerance Principle (YANG, 2016).

Brazilian Portuguese phonological system bears CCV syllables formed by /p, b, t, d, k, g/ plosives or /f, v/ fricatives plus the liquids /l, ɾ/, like in /pɾato/ 'plate', /bluza/ 'shirt'. Previous studies determined that the acquisition of CCV syllables in BP starts before the age of 2;0 but is completed only by the age 5;0-6;0. Throughout the BP acquisitional literature, a longstanding debate discusses whether the late acquisition of this syllable type is due to their complex articulatory properties or to their phonological properties, both regarding the acquisition of /l, ɾ/ segments and the development of the branching structure (RIBAS, 2002, also discussed by ROSE, 2000 to French). On the segmental note, the debate also goes around on which of the consonantal sequences are first acquired: some studies defend an initial C/l/V stabilization, others defend an initial C/ɾ/V stabilization, and some other yet defend that there is no predominant order of acquisition between C/ɾ/V and C/l/V (RIBAS, 2002). Therefore, both the CCV syllable structure and its segmental tier hold debates in BP acquisitional literature.

Our study points out that both the segmental and the structural debate should be tied together, as the distributional characteristics of the target CCV syllable in BP might be the central player of the debate – not the articulatory complexity or the phonological properties *per se* of /l, ɾ/ segments and the branching structure. According to our corpora research, the CCV syllable type shows low type-token frequency in the input (both in adult speech and child directed speech), is highly concentrated into C/ɾ/V segmental patterns (in despite of C/l/V), and present low-density phonological neighborhoods (both when comparing CCV versus CV structures, like in the minimal pair /**pɾa**to/-/**pa**to/ 'plate'-'duck', or when comparing C/ɾ/V versus C/l/V sequences, like in /**bɾin**daR/-/**blin**daR/ 'to toast'-'to shield'). However, when comparing the same /l, ɾ/ segments in simple CV onset position, like in /ɛ**ɾa**/-/ɛ**la**/ 'it was'-'she', the input shows high phonological density and equally distributed segmental frequencies between /l/ and /ɾ/. We argue that this discrepancy could lead to an incorrect hypothesis of /l/-/ɾ/ contrast neutralization in CCV syllables by the child developing phonological system. Contrast neutralization is also observed in a structural process: unstressed CCV syllables are susceptible to optional CCV→CV reduction in adult speech (as in /o**tɾo**/→[ˈo.**tʊ**] 'another', /**pɾe**siza/→[**pɪ**ˈsi.zɐ] 'it needs'). Paired with the low density observed between CCV and CV structures, we argue that this process could also lead to an incorrect hypothesis of CCV-CV neutralization by the child developing grammar. Finally, in child directed speech (CDS), reducible CCV contexts are highly early and frequent, but the first CCV-CV minimal pair is found only between the 400 most frequent words. Also, only 6% of the 50 most frequent CDS words represent the C/l/V pattern. No minimal pairs between C/ɾ/V and C/l/V were found in CDS, whereas /ɾ/V-/l/V pairs were found between the 300 most frequent words.

Such input properties may lead to (i) an incorrect structural neutralization, taking CV as a free alternating form of CCV; (ii) an incorrect segmental neutralization, taking /l/ and /ɾ/ as not contrastive in CCV contexts. That is, children are overgeneralizing contrast neutralization rules based on the distributional properties of their early stages of vocabulary. Therefore, we argue that the establishment of the phonological contrast of CCV-CV and of C/ɾ/V-C/l/V are the causes for the long-lasting phonological development of CCV syllables in BP – but not in other Latin languages, since their distributional patterns and processes are not the same as BP's.

To model the productivity of our overgeneralization hypothesis, we use the Tolerance Principle (YANG, 2016), which deals with how the contrastive value between two varying structures, like the CCV/CV production, are built in the linguistic system. The Tolerance Principle assumes that there is a tipping point of exposure that leads to the generalization of a variation. This tipping point is ruled by an equation and is measured by quantifying how many varying items (CCV-CV neutralization) exist in the input's most frequent words (Total N). Chart 1 shows that the high concentration of neutralizable CCVs in CDS initially goes over the tolerable tipping point (θN). Therefore, only with bigger vocabularies children should be able to realize the contrast between CCV and CV.

Chart 1: CCV-CV reduction (over)generalization model

TYPES	TOTAL N	Stressed CCV	Unstressed CCV	CCV-CV neutralization	θ_N	Is the neutralization tolerable?
50	1	0	1	1	-	-
100	5	1	4	3	3,1	?
200	11	3	8	7	4,6	NO
500	50	22	28	13	12,8	NO
750	84	38	46	16	19,0	YES
1500	172	81	91	25	33,4	YES

To test the hypothesis of structural and segmental overneutralization, a mispronunciation detection task was carried out with 71 children between 2;0-5;0 years old. Results point out that CCV→CV mispronunciations (/pɾeto/ 'black' →*[ˈpe.tʊ] - non existing word) are not detected even though CV→CCV mispronunciations (/dente/ 'tooth'→ d[ɾ]ente) are detected by children who simplify clusters in their own production. Higher detection rates can be found, however, when CCV→CV stimuli have phonological neighbors: /pɾato/ 'plate' → [ˈpa.tʊ] 'duck' is more detected than /pɾeto/ 'black' →*[ˈpe.tʊ] - non existing word. At the segmental level, results show a high detection rate of /l, ɾ/ substitutions in CV syllables, like in /ga**l**iɲa/ 'chicken' → [gaˈ**ɾ**i.ɲɐ] and /ko**ɾ**uʒa/ 'owl' → [koˈ**l**u.ʒɐ], but not in CCV syllables, by the same children. In this context, the lowest rates of detection are the C/l/V→C[ɾ]V condition (like in /b**l**uza/ 'shirt' → [ˈb**ɾ**uzɐ]), which follows the most frequent pattern in the input, than in the C/ɾ/V→C[l]V condition (like in /p**ɾ**ato/ 'plate' → [ˈp**l**a.tʊ]).

The experimental results show that contrastive relations are a key point in the development of branching onsets. We argue that there is a moment in child development when simple onsets are taken as an alternative form of branching onsets (but never the opposite), and the contrastivity between /l, ɾ/ depends on their syllable context. In sum, the study suggests that CCV acquisition in BP goes through a moment of incorrect structural and segmental contrast neutralization caused by an overgeneralization of input distributional patterns, therefore why CCV production takes so long to stabilize in child speech in this language.

RIBAS, L. P. *Aquisição do onset complexo no português brasileiro*, Master's dissertation, 2002.
ROSE, Y. Internal and External Influences on Child Language Productions. In: *Approaches to Phonological Complexity*, BerlIn Mouton de Gruyter, 2009, p. 329-351.
YANG, C. *The price of linguistic productivity: How children learn to break the rules of language*. MIT Press, 2016.