Journal of Portuguese Linguistics

Toni A. (2024). Measuring the roles of variation and phonological density into the development of branching CCV onsets in Brazilian Portuguese. *Journal of Portuguese Linguistics*, 23, pp. 1–6. DOI: https://doi.org/10.16995/jpl.10954

Open Library of Humanities

Measuring the roles of variation and phonological density into the development of branching CCV onsets in Brazilian Portuguese

Andressa Toni, UNICENTRO, BR, andressa.toni@alumni.usp.br

This study focuses on the acquisition of CCV syllables (Consonant1+Consonant2+Vowel) in Brazilian Portuguese. Our aim is to cast light on the roles of linguistic variation and phonological density into the phonotactic development. We used adult and child directed speech corpora to quantify the optional process of CCV→CV reduction in unstressed contexts in the input, as well as the frequency, segmental pattern, and phonological density of CCV syllables. Based in their distributional properties, we argue that CCV acquisition in BP goes through a moment of incorrect structural and segmental contrast neutralization, modelled by the Tolerance Principle (Yang 2006). The neutralization is caused by an overgeneralization of CCV variation, combined with CCV-CV low density; and an overgeneralization of the C/ɾ/V frequent segmental pattern, combined with the low density of C/ɾ/V-C/l/V but high density and equal frequency of /ɾ/V-/l/V. Data from a mispronunciation detection task confirms that the phonological density influences the development: children who do not articulate CCV syllables detect more CCV→CV stimuli when there are phonological neighbors (/prato/→['pa.tʊ], but not /preto/→*['pe.tʊ]) and high phonological density ([l]V↔[ɾ]V are detected, but not C[l]V↔C[ɾ]V, with the lowest rates of detection in the C/l/V→C[ɾ]V direction).

Dissertation abstract

This dissertation focuses on the development and processing of CCV syllables (Consonant1 + Consonant2 + Vowel) in Brazilian Portuguese (BP) in order to better understand why their acquisition path is so variable, unstable, and late, as described by previous studies (Mezzomo et al., 2008; Miranda & Cristófaro-Silva, 2011; Ribas 2002; Toni, 2016). Based on experimental and corpus data modeled by the Tolerance Principle (Yang, 2016) and the Contrastive Hierarchy Theory (Dresher 2003), we propose that linguistic variation and phonological density play an important role in the phonotactic development path.

BP's phonological system bears CCV syllables formed by /p, b, t, d, k, g, f, v/ plus the liquids /l, ɾ/, as in /**pɾa**to/ 'dish' and /**blu**za/ 'shirt'. Previous studies determined that CCV acquisition starts before age 2;0 but is completed only by 5;0–6;0 (Miranda & Cristófaro-Silva, 2011; Lamprecht, 1993). However, children prove to be able to fully articulate branching onsets by age 4;0 (Ribas, 2002; Toni, 2016). Yet, CCVs are variably reduced to CVs (/**pɾe**to/ → *[ˈ**pe**.tʊ] 'black') or have their liquid quality altered (/**pɾe**to/ → [ˈ**ple**.tʊ] 'black', /**blu**za/ → [ˈ**bɾu**.zɐ] 'shirt') until 6;0 years old (Mezzomo, Ribas, 2004; Toni, 2016).

Throughout the acquisitional literature, a longstanding debate discusses whether the late acquisition of this syllable type is due to its complex articulatory properties or its phonological properties, both regarding the /l, ɾ/ consonants and the branching structure. On the segmental note, the debate also revolves around which of the consonantal sequences are first acquired: some studies found an initial C/l/V stabilization (Lamprecht, 1993; Teixeira, 1988), others found an initial C/ɾ/V stabilization (Queiroga et al., 2011; Wertzner, 2003), and others yet did not find a predominant order of acquisition (Ribas, 2002; Staudt, 2008). Therefore, both the CCV structure and its segmental tier hold debates in BP acquisitional literature.

Our study points out that both the segmental and the structural debates should be tied together, since the distributional characteristics of CCV might be an important player in the question. Thus, aside from influential factors already listed by the literature like the articulatory complexity (Berti, Ferreira-Gonçalves, 2012), the phonological properties of the segments, the branching structure complexity (Ribas, 2008), the absence of morphological and phonological roles for CCV in the system (Santos, 1998), and the low CCV frequency (Miranda & Cristófaro-Silva, 2011), we argue that the low phonological density and contextual variability concentrated in the child's early input are also factors to be considered – a hypothesis that was not yet discussed by the BP development literature.

To describe WHAT constitutes the child's linguistic target and its input, we conducted a corpus study comparing the distributional properties of adult speech (AS) (Corpus ABG, Benevides, Guide, 2016), child-directed speech (CDS) (Corpus FDC, Santos, Toni, 2021) and children's target words (IS) (Corpus FI, Santos, Toni, 2021), as well as a quantification of the CCV → CV reduction found

in São Paulo's adult speech (as in /ow**tɾo**/ → [ˈo.**tʊ**]) (Corpus Projeto SP2010, Mendes, 2013). According to our data, CCV syllables show low frequency (<10% tokens in AS, CDS, IS), are highly concentrated into C/ɾ/V segmental patterns (>90% in AS, CDS, IS), and present low-density phonological neighborhoods, both when comparing CCV-CV, like /**pɾa**to/ 'dish' vs. /**pa**to/ 'duck' (AS: 288 pairs; CDS: 61; IS: 26), and when comparing C/ɾ/V-C/l/V, like /b**ɾ**indar/ 'to toast' vs. /b**l**indar/ 'to shield' (AS: 9 pairs; CDS, IS: 0). However, when comparing the same /l, ɾ/ segments in simple CV onsets, like in /ɛ**ɾ**a/ 'it was' vs. /ɛ**l**a/ 'she', the input shows high phonological density (1,334 pairs in AS, Agostinho, Soares & Mendes, 2020), and equally distributed segmental frequencies between /l, ɾ/. Regarding unstressed CCV reduction, it applies in 20% of AS and CDS CCV tokens. Crucially, those reducible CCV contexts are highly frequent and early in CDS and IS, as well as /ɾ/V-/l/V pairs, while the first CCV-CV minimal pairs and C/l/V patterns are late, and C/ɾ/V-C/l/V pairs are absent. In sight of these skewed distributional properties, we conducted an experimental task with adults to check the perception, productivity, and acceptability of CCV in BP. Despite the low frequency, phonological density and high variability, our experimental results revealed that both C/ɾ/V and C/l/V are productive in the target language, and CCV reductions or segmental substitutions are not accepted in stressed positions.

To explore WHEN and HOW the structural and segmental properties of CCV are acquired and processed, we conducted an experimental study comparing results from production and mispronunciation detection tasks of 70 monolingual children from São Paulo, with no history of phonological, articulatory, or hearing disorders. Both tasks tested the same words containing C/l, ɾ/V (with and without CV minimal pairs) and /l, ɾ/ in CV as controls. Children's age ranged between 2;0–6;0 years old, with 15–20 children per age group. However, data was analyzed not by age, but by performance measures, based on the results of the production test (cf. Toni & Santos, 2022, for discussion). Children were grouped according to their predominant CCV production/repair patterns (*C0Vgroup*: regular C2 deletion; *C?Vgroup*: regular C/ɾ/V reduction and correct C/l/V production; *C_Vgroup*: regular C/ɾ/V→C[l]V and C/l/V→C[ɾ]V substitution; *Controlgroup*: regular CCV correct production). This categorization pointed that there is a moment when children can articulate the branching onsets but not with the correct liquid, even when in simple CV the liquids are fully acquired. The branching structure is thus acquired before the CCV segmental tier, and syllable context influences the segmental production.

The same groupings took the mispronunciation detection task, which tested two conditions:

(a) structural mispronunciations: CV→CCV (/**den**te/ 'tooth' → *[**dɾẽʲ**.tʃɪ]), CCV→CV with minimal pairs (/**pɾa**to/ 'dish' → [ˈ**pa**.tʊ] 'duck') and CCV→CV without minimal pairs (/**pɾe**to/ 'black' → *[ˈ**pe**.tʊ] (nonword));

(b) segmental mispronunciations: C[ɾ]V↔C[l]V (/pɾato/ 'dish' → *[ˈ**pl**a.tʊ], /b**l**uza/ 'shirt' → *[ˈb**ɾ**uzɐ]), [ɾ]V↔[l]V (/ga**l**iɲa/ 'chicken' → *[gaˈ**ɾ**i.ɲɐ], /ko**ɾ**uʒa/ 'owl' → *[koˈ**l**u.ʒɐ]).

Results showed that CV→CCV mispronunciations are detected even by children who categorically simplify clusters in their own production, which points that CV and CCV are processed as different structures. *Controlgroup* detected the mispronunciations as expected. However, CCV→CV mispronunciations were not detected by *C0Vgroup*, which indicates that those detection patterns are caused by the undergoing phonological development (and not by other causes like input variation). Intermediate *C?Vgroup* and *C_Vgroup* presented higher detection rates when CCV→CV stimuli have phonological neighbors: /**pɾa**to/ 'dish' → [ˈ**pa**.tʊ] 'duck' was more detected than /**pɾe**to/ 'black' → *[ˈ**pe**.tʊ]. Without phonological neighbors, only *C?Vgroup* and *Controlgroup* had high detection rates. At the segmental level, accordingly to the production results, there was a high detection rate of /l, ɾ/ mispronunciations in CV than in CCV by *C?Vgroup*. Additionally, the lowest rates of detection are the C/l/V→C[ɾ]V condition, which follows the most frequent pattern in the input. *C_Vgroup* detected at chance the liquid substitutions in CCV and failed to detect /ɾ/→[l] in CV. Like the structural condition, *C0Vgroup* could not detect any mispronunciations, and *Controlgroup* detected them as expected.

To answer WHY CCV is processed and acquired as described by our experimental and corpus results, we defend that the distributional input properties are leading to (i) an incorrect structural neutralization, taking CV as a free alternating form of CCV; (ii) an incorrect segmental neutralization, taking /l, ɾ/ as not contrastive in CCV contexts. The productivity of (i) is modeled by the Tolerance Principle, an equation which quantifies how many exceptions are too many for a system to tolerate. The productivity of the incorrect contrast neutralization between CCV and CV structures stems from the high concentration of reducible CCVs in the child's initial vocabulary, paired with the low phonological density of CCV-CV. As for (ii), the failure to contrast /l, ɾ/ in CCV but not in CV can be modeled by Contrastive Hierarchy Theory, based on the discrepancy of contrast cues between the phonotactic systems. In sum, 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. These incorrect overgeneralizations about CCV can only be overcome with more dense phonological neighborhoods, which asks for bigger vocabularies; hence, why CCV production takes so long to stabilize in child speech.

Acknowledgements

I would like to thank Raquel Santana Santos and Charles Yang for their valuable comments and insights. Thanks also to an anonymous reviewer for their comments on content and format, which significantly improved my text. Any mistakes and oversights contained herein are entirely my own. I am also grateful to the Brazilian funding agencies CNPq and CAPES for providing me with grants for the completion of my doctoral degree at the University of São Paulo, Brazil, and my Doctoral Exchange Program at the University of Pennsylvania, United States of America.

Competing Interests

The author has no competing interests to declare.

References

Benevides, A., & Guide, B. (2016). *Corpus ABG* [dataset]. FFLCH-USP. https://github.com/SauronGuide/corpusABG

Berti, L. C., & Ferreira-Gonçalves, G. (2012). A aquisição do contraste/t/ e /k/ sob a ótica dinâmica [/t/ and /k/ segmental contrast acquisition under the dynamic perspective]. *Revista da ABRALIN*, *11*(1), 139–196. DOI: https://doi.org/10.5380/rabl.v11i1.32465

Dresher, E. B. (2003). On the Acquisition of Phonological Contrasts. Paper presented at *GALA*, Utrecht, 2003.

Lamprecht, R. R. (1993). A aquisição da fonologia do Português na faixa etária dos 2;9–5;5 [Phonological acquisition of Portuguese between ages 2;9–5;5]. *Letras de Hoje, 28*(2), 99–106.

Mendes, R. B. (2013) *Projeto SP2010: Amostra da fala paulistana* [SP2010 Project: São Paulo's speech sample]. [dataset]. FFLCH-USP. http://projetosp2010.fflch.usp.br.

Mezzomo, C., & Ribas, L. (2004). Sobre a aquisição das líquidas [On the liquids acquisition]. In R. R. Lamprecht (Ed.), *Aquisição fonológica do português: Perfil de desenvolvimento e subsídios para terapia* (pp. 95–112). Artmed Editora.

Mezzomo, C. L., Mota, H. B., Dias, R. F., & Giacchini, V. (2008). O uso da estratégia de alongamento compensatório em crianças com desenvolvimento fonológico normal e desviante [Compensatory lenghthening as a repair strategy for typical and atypical children]. *Letras de Hoje*, *43*(1), 35–41.

Miranda, I., & Cristófaro-Silva, T. (2011). Aquisição de encontros consonantais tautossilábicos: Uma abordagem multirrepresentacional [Branching onsets acquisition: A multirepresentational approach]. *Revista Linguíſtica*, *7*(1), 1–20.

Queiroga, B., Alves, J., Cordeiro, A., Montenegro, A., & Asfora, R. (2011). Aquisição dos encontros consonantais por crianças falantes do português não padrão da região metropolitana do Recife [Branching onsets acquisition by children from the non-default dialect of Recife]. *Revista CEFAC, 13*(2), 214–226. DOI: https://doi.org/10.1590/S1516-18462010005000139

Ribas, L. P. (2002). Aquisição do onset complexo no português brasileiro [Branching onsets acquisition in Brazilian Portuguese]. [Unpublished Master's thesis]. Catholic University of Rio Grande do Sul.

Ribas, L. P. (2008). Aquisição das líquidas por crianças com desvio fonológico: Aquisição silábica ou segmental [Liquids acquisition by children with phonological disorders: Segmental or syllabic acquisition]. *Letras*, *36*(1), 129–150. https://doi.org/10.5902/2176148511970

Santos, R. R., & Toni, A. (2021). *Corpus FDC (Fala Dirigida à Criança)* [FDC Corpus: Child Directed Speech] [dataset]. FFLCH-USP. http://leal.fflch.usp.br/Corpora-FI-FDC

Santos, R. R., & Toni, A. (2021). *Corpus FI (Fala Infantil)* [FI Corpus: Child Speech] [dataset]. FFLCH-USP. http://leal.fflch.usp.br/Corpora-FI-FDC

Santos, R. S. (1998). A aquisição da estrutura silábica [Syllable structure acquisition]. *Letras de Hoje*, *33*(2), 91–98.

Staudt, L. B (2008). Aquisição de onsets complexos por crianças de dois a cinco anos: um estudo longitudinal com base na teoria da otimidade [The acquisition of branching onsets by children aged 2–5 years old: a longitudinal study based on Optimality Theory]. [Doctoral dissertation, University of Vale do Rio dos Sinos]. Unisinos Campus Repository: http://repositorio.jesuita.org.br/handle/UNISINOS/3510?locale-attribute=en.

Teixeira, E. R. (1988). Processos de simplificação fonológica como parâmetros maturacionais em português [Phonological simplification processes as developmental marks in Portuguese]. *Cadernos de estudos linguísticos*, *14*(1), 53–63.

Toni, A. (2016). Representação subjacente do ataque ramificado CCV na aquisição fonológica [Branching onsets underlying representation in language acquisition]. [Doctoral dissertation, University of São Paulo]. USP Campus Repository. https://www.teses.usp.br/teses/disponiveis/8/8139/tde-03102016-134317/pt-br.php

Toni, A., & Santos, R. S. (2022). A idade reflete o domínio linguístico? Efeito das medidas de desempenho na análise de dados em aquisição fonológica [Does age reflect linguistic mastery? Effect of performance measures on data analysis in phonological acquisition]. *Revista de Estudos da Linguagem*, *30*(1), 376–424. DOI: https://doi.org/10.17851/2237-2083.30.1.376-424

Wertzner, H. F. (2003) Distúrbio Fonológico [Phonological disorders]. In C. R. F. Andrade & E. Marcondes (Eds.), *Fonoaudiologia em Pediatria* (pp. 70–78). Sarvier.

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language* (1st ed). MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262035323.001.0001