

Acquisition of Socio-Phonetic Variation: A Case Study of Coda /r/ Variation in Dominican Spanish-Speaking Children and Their Caregivers

Emily Herman, Karen Miller, and Matthew T. Carlson

1. Introduction

As children acquire the phonemic inventory of their first language, they also must learn the probabilistic distributions constraining phonological alternations. Socio-phonetic variation complicates the learning problem because children must learn the distribution of socio-indexical conditioning as well as the linguistic conditioning of speech-sound realizations. In the case of Dominican Spanish, children's production of syllable coda /r/ necessitates the ability to produce the three sounds [l, r, ɾ] as well as sensitivity to the sociolinguistic conditioning present in their community. In Dominican Spanish (DS), syllable coda /r/ probabilistically alternates with [l, r, ɾ, ø] (e.g. *verde* 'green' has four representations: [βer.de], [βel.de], [βeɾ.de], [βe.de]) in adult speech (Bullock & Toribio, 2009). As children are exposed to conditioned sociolinguistic variation by both their caregivers and their community at large, children's own production of this variation may reflect the patterns observed in their input (Smith & Durham, 2007; Miller, 2013). While this type of variation is distinctly related to sociolinguistic variation in the child's community, it is also important to note that young children show variability in their production of both variable and categorical (nonvariable) forms as they acquire the target language.

There is considerable variation in children's own production of phonemes due to lack of articulatory control, which makes some sounds harder for young children to reliably produce than others. In a crosslinguistic study of children's phoneme acquisition across 27 languages, McLeod and Crowe (2018) find that /l/ is reliably produced with 90-100% accuracy between 3;0 – 3;11, /r/ is produced between 4;0 – 4;11, and /ɾ/ is produced between 5;0 – 5;11. This means that prior to the acquisition of each phoneme individually, variation in the production of /l, r, ɾ/ may be affected by the lack of the precise articulatory control needed to produce each phoneme, leading to inevitable substitutions that may or may not be representative of the socio-phonetic variation in the child's community.

Goldstein (2005) highlights the complexity of teasing apart developmental and sociolinguistic variation. Goldstein (2005) finds that, in an experimental task, typically developing Spanish L1 children from Puerto Rico aged 3;2 – 4;11 substituted /r/ with [l, ɾ, n, s, w, b, ð, h] in 12.84% of trials; these same children substituted /l/ with [p, b, h, n, j] in 2.94% of trials. Based on the age of the children

in this study, it is unsurprising that /l/ has a lower substitution rate compared to /r/. However, across many Caribbean dialects of Spanish, including both Puerto Rican and Dominican Spanish, syllable coda /r/ probabilistically lateralizes to [l] among adult speakers (Lipski, 2012). In this study, children's production of both syllable initial and syllable final consonants were analyzed in aggregate, making it difficult to interpret whether the higher rate of substitution of /r/ is due to coda /r/ variation found in the community or whether this is primarily variation due to on-going phonological development.

It is important to consider, then, that children's production of [l, r, ɾ] cannot be overgeneralized as a reflection of sociophonetic variation before their production of these phonemes is reliably produced in invariable contexts. To avoid the confound of developmental constraints on the production of these allophones, the present study analyzes two child-caregiver dyads (children aged 6;6 and 6;4) from Santo Domingo, Dominican Republic. The present study aims to address the following questions:

1. After acquiring the ability to produce all relevant liquids, do children produce coda /r/ variation?
2. Do caregivers produce the coda /r/ variants [l, r, ɾ] in child-directed speech?
3. Are children sensitive to the constraints of coda /r/ variation as represented in their input (via caregiver speech)?

1.1. Background: Coda liquid variation in DR Spanish

Early observations of Dominican Spanish noted impressionistically that syllable coda liquids /l, r/ alternate with [l, r, h, ø, j] as well as gemination of the following phoneme (Henríquez Ureña, 1940; Alba, 1985). More detailed analysis of regional variation has distinguished the typical substitutions into geographic zones in which a specific variant is most likely to be used. Alba (2004, p. 85) notes that the lateralization of /r/ to [l] is most frequent in the Capital and Eastern dialects of Dominican Spanish, while the rhotacization of /l/ to [r] is more frequent in the Southwest and vocalization of /l, r/ to [j] is most frequent in the Cibao region, as summarized in the figure below (Willis & Díaz-Campos, 2021).

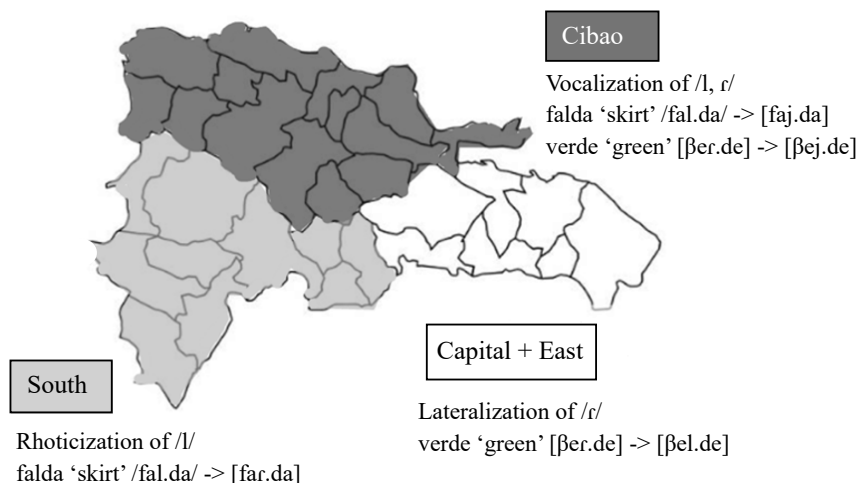


Figure 1. Most frequent coda liquid variant substitutions by region. Adapted from Willis & Díaz-Campos, 2021 (page 36).

It is important to note that while these are the most frequent variants associated with each region, the production of [l, ɾ, h, ø, j] and gemination can be found probabilistically throughout the DR, with the exception that vocalization to the [j] in coda liquids appears to be limited to the Cibao region. Additionally, both Ureña (1940) and Willis and Ronquest (2021) have noted an additional 'merged' variant, which is a type of neutralized sound between an [l] and [ɾ]. In an acoustic analysis, Willis and Ronquest (2021) characterize this 'merged' variant as having merged F2-F3 formants in the segment. As this is consistent with Wilson's (2008) characterization of a rhotic [ɽ], this study will hereby refer to the 'merged' variant as a 'rhotic'.

Both internal and external linguistic factors condition the probabilistic distribution of the variants [l, ɾ, ɽ, h, ø, j] in adult speech. Infinitival endings, which always end in a coda /ɾ/ in Spanish (example 1), are most frequently produced with vocalization or elision, though production of the [ɾ] and [l] are also attested infrequently with infinitival endings (Alba, 2004; Bullock & Toribio, 2009). Syllable stress has also been found to condition this variation, with tonic syllables being the most likely to be produced with a [ɾ], while atonic syllables are more likely to be produced with the non-prescriptive [l, ɽ, h, ø, j] variants (Willis & Ronquest, 2021).

(1) Possible productions of *estudiar* 'to study' /es.tu.ði.ar/:

- a. elision [es.tu.ði.a]
- b. vocalization [es.tu.ði.aj]
- c. conservation [es.tu.ði.ar]
- d. lateralization [es.tu.ði.al]

Socio-economic status has also been found to condition the production of coda liquids in DS. In Santo Domingo, working class speakers most frequently elide or vocalize the coda /r/, whereas middle and high-SES speakers tend to produce the [r] most frequently (Alba, 2004; Willis & Ronquest, 2021). Female speakers also tend to use more of the canonical [r] compared to other variants (Willis & Ronquest, 2021).

Although the patterns of variation have been analyzed in adult-directed speech, Bullock and Toribio (2009) provide the only attestation for children's production. In their analysis of coda liquids in the Cibao region, Bullock and Toribio (2009) find that children under the age of 11 produce the variants [l, r, h, ø, j], and find that the Cibaeño children nearly categorically produce the non-syllabic [j] with infinitival endings. The present study aims to establish the use and patterns of coda [r] variation in both child-directed speech and child-produced speech at the ages of 6;4 and 6;6.

2. Methods

2.1. Participants

Each child-caregiver dyad was recorded in Santo Domingo, Dominican Republic during the summers of 2014 and 2015. The data presented in this paper are from the Dominican Child-Caregiver Corpus (Miller & Cárdenas, 2019). The recordings were made in a playroom lab space in a small apartment in a working-class neighborhood. The lab play space was set up to be an informal speech setting, as children and their caregivers were left alone for each 60 minute session. All participants were compensated for their participation.

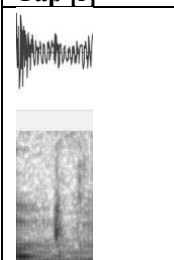
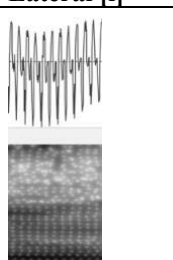
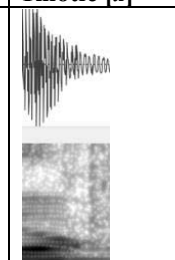
The present study analyzes one hour of speech from two child-caregiver dyads from this corpus. Sofia, aged 6;6, interacts with her father, while Samuel, aged 6;4, interacts with his mother. These children were chosen because they have both reached the age where they are physically able to produce all available variants [l, r, ɹ] and demonstrate this competence in their speech by producing [l] and [r] in syllable-initial (invariable) contexts. There were 442 total tokens (237 child, 205 caregiver) after exclusions. 181 tokens were excluded for unintelligibility due to mumbling, overlapping speech, or complete elision of the syllable containing the candidate /r/.

2.2. Coding

Annotated TextGrids in PRAAT (Boersma, 2025) were used to identify candidate /r/ tokens. For each utterance that included a syllable coda /r/, the entire utterance was extracted and annotated to a separate TextGrid. Each token was segmented and coded for the candidate /r/ production, the previous vowel, and the following phonological segment as applicable. Since all candidate /r/s occur in syllable coda position, there is always a vowel prior to the candidate segment. The candidate segments were coded as [l, r, ɹ, ø]. Following segment identification by Wilson (2008) and Willis and Ronquest (2021), laterals were

identified as segments with distinct F2-F3s and even waveforms on the spectrogram, while rhotics [ɹ] were identified as segments with a waveform that tapered. Taps [ɾ] were identified by a break in the spectrogram. Nulls [ø] were identified as a vowel with no evidence of a tap, lateral, or rhotic at the end of the segment. Although the Cibao region includes the vocalization variant, we did not find evidence of either children or their caregivers using the [j] in this dataset. Additionally, aspiration was not observed, but devoicing of the vowel was found in the data set and counted as elision for the purposes of analysis. Table 1 shows examples of what each segment looks like in context.

Table 1. Examples of annotated segments.

Tap [ɾ]	Lateral [l]	Rhotic [ɹ]	Null [ø]
			
<i>Ar<u>mar</u> las piezas.</i> 'To assemble the pieces'	<i>Casi estoy <u>ter</u>minando.</i> 'Almost, I'm finishing'	<i>Mami, acor<u>ta</u> tus tarjetas.</i> 'Mom, take your cards'	<i>Bueno, este voy a hab<u>lar</u>.</i> 'Well, this I'm going to talk'

After token extraction and annotation, tokens were further coded for word class, vowel length, and word position. Word class included the following syntactic categories: noun, proper noun, verb, preposition, conjunction, question word, adjective, and adverb. Vowel length was measured in milliseconds. Word position recorded whether the syllable containing the candidate tap was word-internal or word-final; for example, the token *armar* 'to assemble' includes two candidate /ɾ/s, with the former being word-internal and the latter being word-final.

3. Results and discussion

Both children and caregivers produced each of the [l, ɾ, ɹ, ø] variants in the coda /ɾ/ position.

Figure 2 shows the proportional production of each variant (adding up to 100%) for each participant. Sofia (6;6) and Samuel (6;4) produce more nulls (42% and 47% respectively) compared to the tap, rhotic, and lateral. Sofia’s dad and Samuel’s mom do not show the same apparent preference for the null variant; instead, the adults produce more taps (Sofia’s dad 57%, Samuel’s mom 35%) relative to the null, rhotic, and lateral variants.

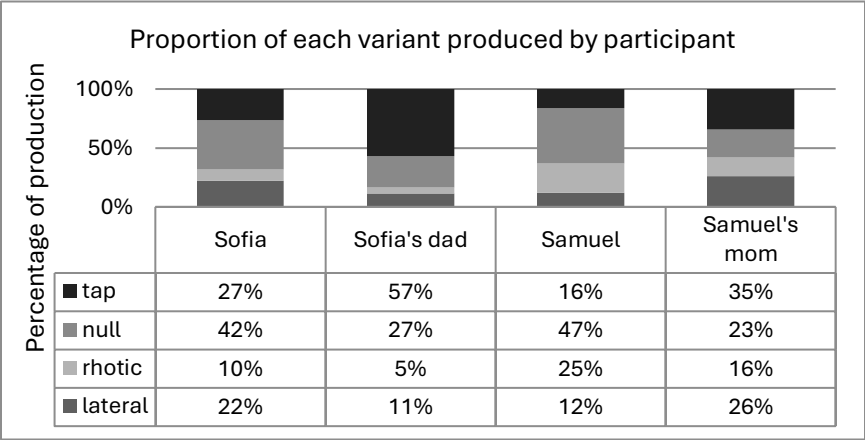


Figure 2. Proportional production of each variant by participant.

Although both children and caregivers use all available variants, it is also important to consider how the children compare to their caregivers in their usage of each variant. As the position in a word has been found to condition coda /t/ variation in the DR (Alba, 2004), the role of word position is compared between each child-caregiver dyad, as represented in Figures 3 and 4.

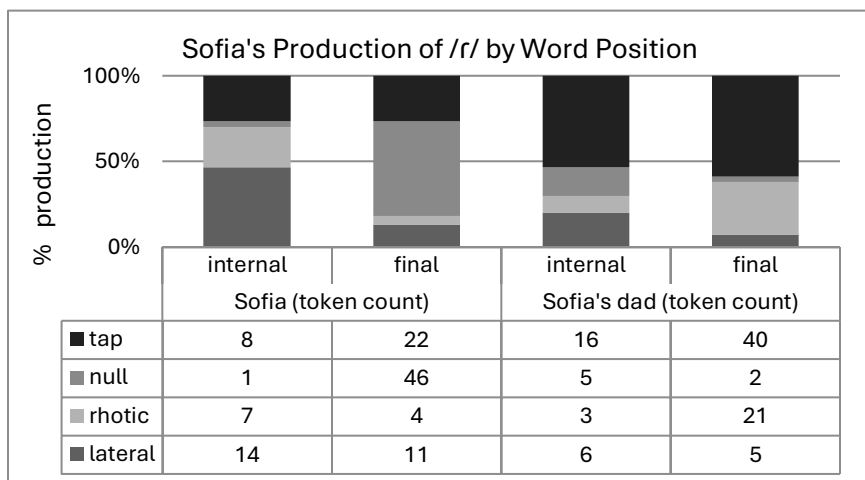


Figure 3. Comparison of Sofia and her dad's production of variants by word position within word.

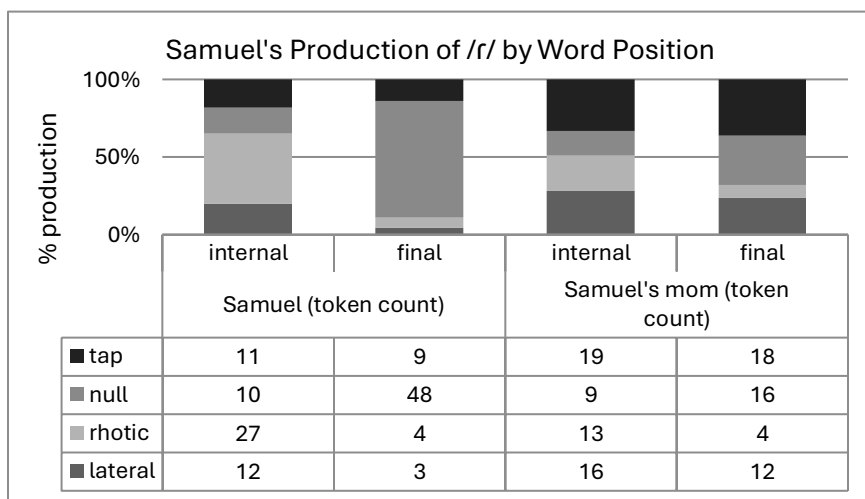


Figure 4. Comparison of Samuel and his mom's production of variants by word position within word.

Sofia produces proportionally more null tokens in the word-final position compared to the other variants (e.g. *estudiar* 'to study' [es.tu.ði.a]). With word-internal syllable coda /r/ she produces more laterals (e.g. *verde* 'green' [βel.de]). In contrast, Sofia's dad produces more taps [ɾ] in both the word-internal and word-final conditions, and the least frequent in his production is the null variant. This shows that while both Sofia and her dad produce all of the coda [ɾ] variants, their

usage of each variant does not appear to be equally conditioned by syllable position in a word.

Similarly, Samuel and his mother's production of each variant also differ from each other. While Samuel produces more of the null variant with word-final syllable coda /r/, he uses more rhotic [ɹ] in the word-internal position. Similar to Sofia's dad, Samuel's mom also appears to favor the tap [ɾ] in both word-internal and word-final positions, though her production is more variable overall. The second most-frequent variant in the word final position for Samuel's mom is the null variant, with a difference in production of 4%. This means that although Samuel's mom does not produce the overwhelming majority of null that Samuel produces, her distribution of production is much more similar to Samuel's than Sofia's is to that of her father.

4. Discussion and Conclusion

As expected, Samuel's mother produces more taps proportional to the null, rhotic, and lateral variants. Sofia's dad surprisingly produces nearly double the rate of taps compared to Samuel's mom. While we expected that the female participant would use proportionally more of the prescriptive [ɹ] compared to the male participant, there are several reasons we may have found the opposite trend. One reason could be the small dataset, as there are only two child-caregiver dyads and they are only compared to each other.

Another possible explanation is that Sofia's dad may reflect gender differences in parent-child interactions, as fathers tend to speak less during parent-child interactions compared to mothers (Davidson & Snow, 1996, and references therein). Additionally, both mothers and fathers tend to use more complex language when engaging with daughters compared to sons (Lovas, 2011), which may extend to the use of more prescriptive variants in these interactions. Research on socio-phonetic variation has also found that parents differentiate the use of socio-phonetic variants according to the gender identity of the child, with mothers of boys using more local variants as they are represented in the adult male community (Foulkes, Dougherty, & Watt, 2005). In the present data set, Samuel's mom may be using relatively less of the prescriptive [ɹ] with Samuel due to the tendency of Dominican male speakers to use less of [ɹ] compared to the other variants (Willis & Ronquest, 2021).

Bi-directional effects on the use of the [ɹ] may also play a role (Lovas, 2011). Both Sofia and Sofia's dad have proportionally higher rates of [ɹ] compared to Samuel and his mom, respectively. Sofia's rate of [ɹ] is more similar to Samuel's mom than to Samuel or her own father, which may closer reflect the overall rate of [ɹ] in the female population. An alternative explanation is that Sofia's dad boosts Sofia's production of the [ɹ] by priming her to use the more prescriptive variant by over-representing the [ɹ] in his speech.

Overall, the present data confirm that all available coda /r/ variants [l, ɹ, ɾ, ø] are attested in both child-directed and child-produced speech. However, more data is needed to evaluate whether parents are differentiating their use of the

prescriptive [r] with sons compared to daughters and how this usage compares to adult-directed speech patterns within the community.

Acknowledgements

Thank you to Kay Rolland for her work and support in this project. Funding Support: National Science Foundation BCS-1322763.

References

- Alba, Orlando. (1979). *Análisis Fonológico de las Líquidas en un Dialecto Rural Dominicano*.
- Alba, Orlando. (2004). *Cómo hablamos los dominicanos: Un enfoque sociolingüístico*. Dominican Republic: Grupo León Jimenes.
- Boersma, Paul & Weenink, David (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.32. retrieved 14 May 2025 from <https://praat.org>
- Bullock, Barbara E., Toribio, Almeida Jacqueline. (2009). Reconsidering Dominican Spanish: Data from the rural Cibao. *Revista Internacional de Lingüística Iberoamericana*, 2009, Vol. 7, No. 2 (14), El español del Caribe/Spanish in the Caribbean (2009), pp. 49-73
- Davidson, R. G., & Snow, C. E. (1996). Five-year-olds' interactions with fathers versus mothers. *First Language*, 16, 223–242.
- Foulkes, Paul, Docherty, Gerard. J., & Watt, David. (2005). Phonological Variation in Child-Directed Speech. *Language*, 81(1), 177–206. <https://doi.org/10.1353/lan.2005.0018>
- Goldstein, Brian. A. (2005). Substitution patterns in the phonology of Spanish-speaking children. *Journal of Multilingual Communication Disorders*, 3(3), 153–168. <https://doi.org/10.1080/14769670500096724>
- Henríquez Ureña, Pedro. (1940). *El español en Santo Domingo*. Biblioteca de dialectología hispanoamericana, 5.
- Hualde, J. I. (2013). *Los sonidos del español: Spanish Language edition*. Cambridge: Cambridge University Press.
- Lipski, John M. (2012). Geographical and social varieties of spanish: An overview. In J. I. Hualde, A. Olarrea & E. O'Rourke (Eds.), *The handbook of hispanic linguistics* (pp. 1-26). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118228098.ch1>
- Lovas, Gretchen S. (2011). Gender and patterns of language development in mother-toddler and father-toddler dyads. *First Language*, 31(1), 83–108. <https://doi.org/10.1177/0142723709359241>
- Núñez-Cedeño, Rafael, Acosta, Junice. (2009). En torno al contexto real de la vocalización cibaëña: un nuevo replanteamiento prosódico. Selected Proceedings of the 13th Hispanic Linguistics Symposium.
- McLeod, Sharynne, & Crowe, Kathryn. (2018). Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review. *American Journal of Speech-Language Pathology*, 27(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100
- Miller, Karen. (2013). Variable Input: What Sarah Reveals About Nonagreeing *Don't* and Theories of Root Infinitives. *Language Acquisition*, 20(4), 305–324. <https://doi.org/10.1080/10489223.2013.828061>

- Smith, Jennifer, Durham, Mercedes, & Fortune, Liane. (2007). "Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, 19(01). <https://doi.org/10.1017/S0954394507070044>
- Willis, Erik., & Ronquest, Rebecca. (2021). A sociophonetic exploration of coda liquids and vocalization in Cibao Dominican Spanish. In M. Díaz-Campos, M. Lacorte, & J. Muñoz-Basols (Eds.), *The Routledge Handbook of Variationist Approaches to Spanish* (1st ed., pp. 207–230). Routledge. <https://doi.org/10.4324/9780429200267>
- Willis, Erik. W., & Díaz-Campos, Manuel. (2021). The perception of coda liquids in Dominican Spanish. In L. A. Ortiz-López & E.-M. S. Büdenbender, *Topics in Spanish Linguistic Perceptions* (1st ed., pp. 35–53). Routledge. <https://doi.org/10.4324/9781003054979-4>

Proceedings of the 50th annual Boston University Conference on Language Development

edited by Romi Hill, Xinhan Jiang,
Danutham Worapipat, and Aditya Yedetore

Cascadilla Press Somerville, MA 2026

Copyright information

Proceedings of the 50th annual Boston University Conference on Language Development
© 2026 Cascadilla Press. All rights reserved

Copyright notices are located at the bottom of the first page of each paper.
Reprints for course packs can be authorized by Cascadilla Press.

ISSN 1080-692X
ISBN 978-1-57473-047-0 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com