Coding Manual

This coding manual is designed for analyzing liquid variation in Spanish, specifically the variable production of /ɾ/, /r/, and /l/, in words from interviews in the **Boston Spanish Corpus (BSC)**. Originally developed for Lee-Ann Vidal-Covas' dissertation within the Spanish in Boston Project, the manual provides detailed guidelines for data collection and analysis of the **BSC** interviews.

Data Preparation

To ensure systematic and reproducible analysis, follow the process detailed below. The study uses audio recordings of sociolinguistic interviews, which are transcribed orthographically in *Praat*. These transcriptions are paired with audio to identify and annotate lexical items containing liquid segments.

1. Identifying and Segmenting Liquid Tokens
   Lexical items with liquid tokens are identified and manually segmented in Praat text grids. Each host word containing a liquid segment should be segmented on a designated tier. A minimum of 250 tokens must be collected per speaker. While identifying tokens for taps and laterals is relatively straightforward, trill tokens require special attention due to their lower frequency.

2. Annotating Liquid Segments
   After segmenting the host words, create an additional text grid tier for annotating the liquid segments within each host word. The annotation process follows the segmentation protocol detailed in the Section.

3. Adding Predictor Variable Tiers
   Additional tiers, as described in the section, are added to systematically code predictor variables for all tokens. This structured annotation ensures the dataset's integrity and supports robust acoustic and variationist analysis in subsequent stages.

Textgrid Tiers

Below is a list and description of the interval tiers to be added to the **Praat** (Boersma & Weenink 2020) TextGrid for data collection. These tiers represent the dependent and independent variables used to study liquid variation. While most variables are coded manually during TextGrid preparation, some independent variables are automatically extracted into a CSV file using a *Praat Script* after the coding process is complete. Detailed instructions for identifying sites of variation, segmenting liquid sounds, and applying the coding schema are provided in the sections that follow.

List of Tiers

The following tiers are added to the TextGrid during annotation:

- Interviewer
- Orthographic
- Host Word
- Number of Tokens per Host Word
- Phonological Form
- Surface Form
- Number of Syllables
- Syllable Stress
- Syllable Type
- Position in Syllable
- Position in Word
- Word Class
- Previous Sound
- Following Sound

Added with Praat extraction script After the TextGrids are completed, the following measurements are generated automatically using the *Praat Script* `Liquids_Data_Extraction.praat`, which compiles data from all TextGrids into a clean CSV file:

- Duration of Host-Word
- Acoustic Analysis (Formants at 11 points)

Added in R Following the data extraction, the following variables are added during post-processing in **R**. The script `Create_liquids_master_df.Rmd` automates this step by calculating and incorporating these variables:

- Lexical Frequency
- Speech Rate

$$\text{Speech Rate (in ms per syllable)} = \frac{\text{dur(ms) of Host-Word}}{\text{\# of syllables}}$$

Tier Descriptions

This section provides detailed descriptions of each tier included in the TextGrid annotation process. These tiers are critical for ensuring consistent and accurate data coding and represent the dependent and independent variables used in the analysis. Each tier description outlines its purpose, how it should be coded, and any relevant considerations or notes. By following these guidelines, the annotation process becomes more systematic, reducing the likelihood of errors and enabling efficient post-processing.

Orthographic Orthographic transcription of the consultant's speech, which is used when segmenting the host words that include liquids. This tier will not need to be added, as the conversations have been transcribed by the time the liquids are to be extracted. It will also not appear in the data frame.

Interviewer Orthographic transcription of the interviewer's speech. This tier will also be completed previous to liquid data collection and will not appear in the data frame.

Host Word Interval of speech which includes the lexical item that hosts the liquid tokens. Interval boundaries are essential as the duration of *Host Word* will be used to compute the variable *Speech Rate*.

Number of Tokens per host word Total number of liquid tokens in the host word. Examples are provided Table 1.1.

Table 1.1: Tokens per Word Examples

HOST WORD	TRANSLATION	# OF TOKENS
puerto	port	1
latinoamérica	Latin America	2
otorrinolaringologo	otolaryngologist	4

Phonological Form This variable codes the presumed phonological identity of the token as one of three categories: /ɾ/, /r/, or /l/. We use the orthographic spelling of the host word, its presumed phonological form, and an expanded version of Hualde's (2014: 183) guide to Spanish rhotic distribution to differentiate between the two rhotics: /ɾ/ and /r/.

Table 1.2: Hualde's (2014) Spanish Rhotic Distribution Expanded

ENVIRONMENT	CODE	CONTEXT	EXAMPLE	TRANSLATION
Contrastive *(word-dependent)*				
V__V	a1 a2	Intervocalic & Word-internal	/ka.ɾo/ vs. /ka.ro/	expensive vs. car
/r/				
#__	b1	Word-initial (or root-initial)	/ro.ka/	rock
C.__	b2	After heterosyllabic consonant	/al.re.de.dor/	around
/ɾ/				
#C__V	c1	After tautosyllabic consonant; syllable onset; word-initial	/bɾo.ma/	joke
C.C__	c2	After tautosyllabic consonant; syllable onset; word-internal; consonant in previous syllable coda	/en.tɾe/	between
V.C__	c3	After tautosyllabic consonant; syllable onset; word-internal; vowel in previous syllable coda	/a.bɾu.mar/	to overwhelm
V__#V	c4	Word final followed by vowel	/seɾ#a.mi.gos/	to be friends
Variable *(usually /ɾ/)*				
V__C	d1	After vowel & Before consonant	/paɾ.te/	part
V__#C	d2	Word final followed by consonant	/seɾ#po.e.ta/	to be a poet
V__##	d3	Word final followed by pause	/seɾ/	to be

[1] # represents word boundary
[2] ## indicates a pause or phonological phrase boundary
[3] **.** represents a syllable boundary

The expanded version introduces specific codes (*e.g.*, a1 – d3) to facilitate working in Praat, streamlining the process of identifying and classifying instances of rhotic variation. Additionally, new contexts (*i.e.*, a1 – a2 & c1 – c3) and corresponding examples have been incorporated. These additions enable simultaneous coding of multiple independent variables when coding for Phonological Form, thereby saving time and enhancing accuracy. See Table 1.2 for the full coding schema.

Important Notes for Coding Phonological Form

- Please use the codes in the **Code** column when coding in Praat.
- When coding for the lateral /l/:
  - Add an "l" at the beginning of codes b–d
  - For the **Environment** V__V (codes a1 & a2), just write "la" and exclude the number.

Surface Form The *perceived* segmental identity of the liquid token. For this tier, the researcher will code what they hear, even if different from the reading in the spectrogram and waveform. This variable has the values:

[ɾ], [r], [r̥], [ɹ], [ɽ], [l], [ɫ], [x], [v], [ð], [i], [h], [ʔ], [∅]

Number of Syllables The syllable count of each host word need to be coded to calculate speech rate; see Table 1.3 for examples.

Table 1.3: Number of Syllables in Word

Host Word	Translation	# of Syllables
las	the (fem. plural)	1
puer.to	port	2
la.ti.no.a.mé.ri.ca	Latin America	7
o.to.rri.no.la.rin.gó.lo.go	otolaryngologist	9

Syllable Stress Whether the syllable to which the token belongs bears primary stress. The values are *stressed* or *unstressed*. Examples are provided in Table 1.4.

Table 1.4: Syllable Stress Example

Host Word	Syllabification	Stress
diferente	di.fe.ˈ**ren**.te	stressed
escuela	es.ˈcue.**la**	unstressed
gustaría	gus.ta.ˈ**rí**.a	stressed
otras	ˈo.**tras**	unstressed
problema	pro.ˈ**ble**.ma	stressed
quieres	ˈquie.**res**	unstressed

- Primary stress is marked with a stress diacritic (ˈ) placed before the stressed syllable.
- Syllables containing the liquid token are **bolded**.
- Monosyllabic words should be coded as *stressed*.
- Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Syllable Type Whether the token belongs to an *open* or *closed* syllable. Open syllables end in a vowel sound, whereas closed syllables end with a consonant sound. Examples are provided in Table 1.5.

1. open
2. closed

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Table 1.5: Syllable Type Examples

Host Word	Translation	First Syllable	Syllable Type	Second Syllable	Syllable Type
clase	class	cla	open	se	open
parte	part	par	closed	te	open

Position in Syllable Position of the liquid token within the syllable. Examples are listed in Table 1.6. The values are:

1. onset
2. coda

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Table 1.6: Position in Syllable Examples

Host Word	Translation	First Syllable	Position in Syllable	Second Syllable	Position in Syllable
real	royal / real	re	onset	al	coda
partir	to leave / to split	par	coda	tir	coda

Position in Word Position of the liquid token within the within the host word. Refer to Table 1.7 for examples. The values are:

1. initial
2. internal
3. final

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Word Class Word class to which the host word belongs. The same tier will also code for specific features in some classes to allow for post-hoc research, as seen in Table 1.8.

Table 1.7: Position in Word Examples

Host Word	Translation	First Syllable	Position in Word	Second Syllable	Position in Word
real	royal / real	re	initial	al	final
partir	to leave / to split	par	internal	tir	final

Table 1.8: Word Class Coding Schema

Class	Code	Function	Type	Pronoun +	Examples
Noun	1		proper (p) / common (c)		Puerto Rico, carro
Verb	2		infinitive (i) conjugated(c) past participle (t) gerund (g)	P	celebramos, celebrar, celebrarle Review of Spanish Participles
Determiner	3	modify nouns			el la los las un una unos unas todos algunos estos aquellos su sus numerals (primero, segundo)
Adjective	4	describe nouns			grande, largo, corto, hermoso, azul
Adverb	5	describe verbs			lentamente, rapidamente
Pronoun	6				yo, tú, él, ella
Preposition	7	explain relationships b/w nouns			a, ante, bajo, cabe, con, contra, de, desde, en, entre, hacia, hasta, para, por, según, so, sobre, tras
Conjunction	8	connect nouns, verbs, phrases, & sentences			y, o, pero, sin embargo, por lo tanto, asi que, aunque, porque, para que
Other	9	catch-all			Use this for anything that doesn't fall into any of the previous categories

Note: PRONOUN + will be coded for in the case of verbs with a suffixed pronoun, i.e. the verb *celebramos* would be coded as 2c while *celebrarle* would be coded as 2iP

Previous Sound The identity of the phonetic segment immediately preceding the liquid token.

- [a], [b], [β], [d], [ð], [e], [f], [g], [ɣ], [h], [i], [j], [k], [l], [m], [n], [ŋ], [o], [p], [k], [r], [ɾ], [ʀ], [ɽ], [ɹ], [s], [t], [u], [v], [x], [z]
- "CT" for when the sounds are unrecognizable
- "Pause" for when there is no sound (i.e. there is a pause)
- These will be recoded to create new variables in R:
  - if *vowel*: Vowel Height, Vowel Displacement
  - if *consonant*, Place & Manner of Articulation

Following Sound The identity of the phonetic segment immediately following the liquid token. (This variable has the same possible values as the variable PREVIOUS SOUND, above.)

Identifying the Tokens

Inclusions

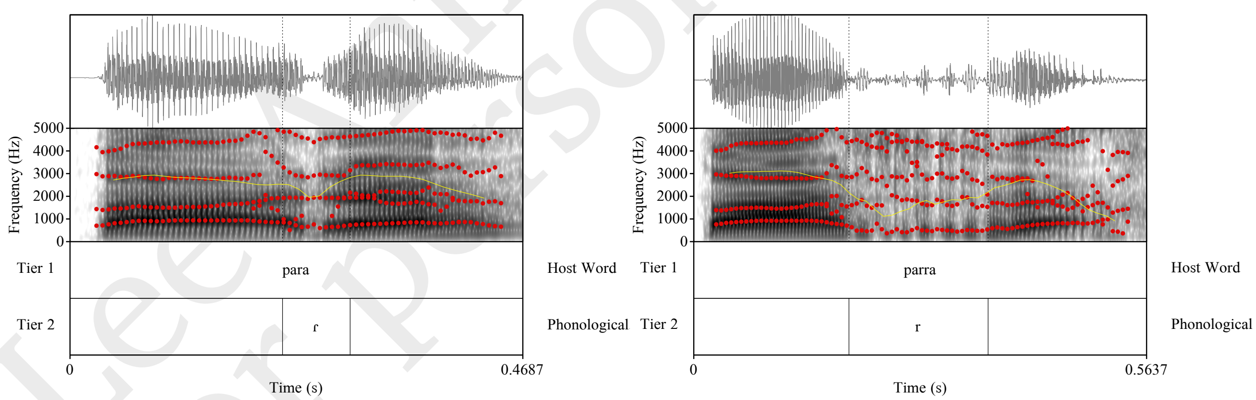
When identifying tokens, it is important to include all liquid segments present in a word. The extraction script is designed to accommodate multiple tokens from a single word, ensuring that no relevant data is overlooked. This comprehensive approach allows for the systematic analysis of liquid variation within and across words.

Segmenting liquids

Following the data preparation process, the segmentation of host words and individual liquid tokens is to be carried out in Praat. Each liquid token should be carefully segmented by closely examining both the waveform and the spectrogram to distinguish it from the surrounding sounds. In the segmentation process, two key objectives guide this protocol: (1) distinguishing the liquid from the surrounding sounds, and (2) classifying its identity based on its acoustic properties. These steps are essential for accurately analyzing the tokens, particularly in cases where the realization of liquids diverges from canonical expectations. The *Surface Form* variable, central to this study, captures such variation by providing a structured framework for classifying these sounds, even when they exhibit atypical acoustic characteristics.

The waveform, which represents frequency and amplitude over time, was analyzed to assess periodicity. Periodic speech sounds exhibit complex repeating waves, indicating regular vocal fold vibration, while aperiodic sounds display random, non-repetitive patterns. It should be noted that periodic and aperiodic components can co-occur within the same sound, as is often the case in speech. Spectrograms provide more detailed visual cues for identifying boundaries between sounds by displaying formant structure. For all liquids, changes in formant frequency—such as shifts in the first (F1), second (F2), or third (F3) formants—provide critical information for distinguishing between the preceding and following sounds. These changes often reflect the articulatory transitions between the liquid and adjacent sounds, such as the movement of the tongue or the opening and closing of the vocal tract, which can cause abrupt or gradual shifts in the formant structure. Additionally, intensity changes, such as a dip in amplitude when transitioning from a vowel to a liquid, are key cues for segmentation. The next section shows some examples to illustrate the segmentation and coding procedure for both rhotics and laterals.

Rhotics Consider the two rhotics in Figure 1.1. Figure 1.1a includes a flap in the word [paɾa], while Figure 1.1b includes a trill in the word [para].



(a) Pa[ɾ]a example

(b) Pa[r]a example

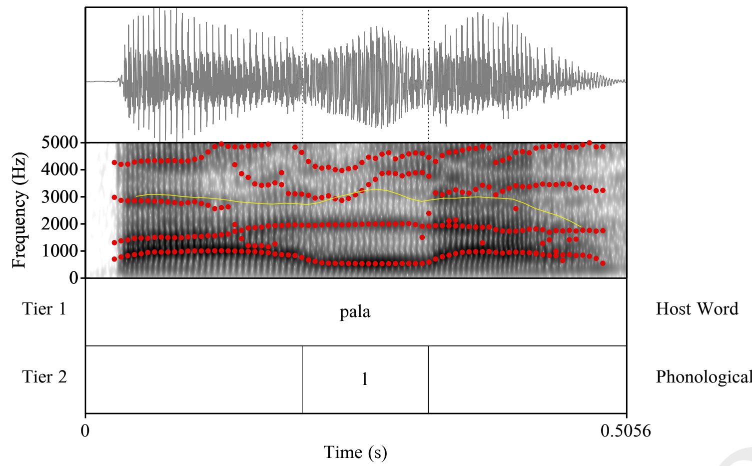
Figure 1.1: Waveform, spectrogram, and annotation of pa[ɾ]a and pa[r]a

- Some possible cues to segment rhotics are:
  - There is a-periodicity in the waveform.

- A steep fall of the third formant is a defining acoustic signature of rhotic sounds, particularly in retroflex or bunched articulations.
- Duration helps in distinguishing rhotics from adjacent vowels. Rhotics, especially flaps, tend to have shorter durations.
- To differentiate from adjacent vowels, a dip in intensity can help (yellow line in Praat picture).
- It should be noted that a trill is much easier to differentiate from adjacent vowels than a flap because trills typically consist of multiple rapid closures of the tongue against the alveolar ridge, creating distinct bursts of energy visible in both the waveform and spectrogram. In contrast, flaps are characterized by a single, brief contact of the tongue, resulting in a shorter and less acoustically distinct interruption, which makes them more challenging to segment from adjacent vowels.
- A radical drop in amplitude can be a cue to where rhotics begin, as vowels have a higher amplitude than rhotics.
- F4-F5 range: Subtle acoustic differences in F4 and F5 provide additional information when coding the *Surface Form* of rhotics, particularly for distinguishing bunched and retroflex variants (Zhou et al. 2008: 4467).

Lateral Laterals are approximants, which means they are among the most sonorous consonants. As their classification hints, they are somewhere between a vowel and a consonant, a fact that can make it hard to differentiate between the sounds. Figure 1.2 includes the lateral /l/ in the word [pala].

- The following cues can be used to find and segment the lateral in order to differentiate between the previous and following sounds:
  - The lateral's waveform is clearly periodic, a characteristic that can help differentiate it from flanking consonants. However, periodicity alone does not distinguish laterals from vowels, as both exhibit periodic waveforms. Additional acoustic cues, such as differences in intensity, formant transitions, and amplitude changes, are critical for segmentation.



(a) *Source:* SIB 0049PR

Figure 1.2: Waveform, spectrogram, and annotation of the word [pala]

- For a lateral, there is no major dip in amplitude or intensity *(yellow line in Praat picture)* in contrast to rhotics, which tend to have lower intensity relative to vowels. This distinction is particularly relevant in contexts where liquids may exhibit overlapping acoustic properties, such as when a purportedly phonological rhotic sound is realized with lateral-like qualities.
- When differentiating between a lateral and a vowel, the lateral typically shows slightly lower intensity due to the inherently higher sonority of vowels. This difference often appears as a subtle dip in intensity at the boundary between the two sounds. Additionally, while both vowels and laterals exhibit formant transitions, those of laterals tend to be more gradual and less pronounced.
- In post-vocalic positions, laterals often exhibit abrupt shifts in formant structure, providing a key cue for segmentation in these contexts.

It should be noted that the goal of the study is to examine variation in spontaneous speech. As such, “idealized” tokens, like the ones shown in Figures 1.1 and 1.2, will rarely be encountered. The ultimate goal is to make the best segmentation possible, within reason.

Data Extraction and Cleaning

The data processing workflow consists of two main steps: extraction and cleaning. These steps ensure a systematic approach to preparing the data for analysis, minimizing human error and automating repetitive tasks.

Step 1: Data Extraction with Praat Script

The first step involves extracting annotated data from Praat TextGrids using a custom *Praat Script*, `Liquids_Data_Extraction.praat`. This script extracts all relevant tiers and acoustic measurements, compiling them into a comprehensive CSV file. Additionally, it converts coded annotations from the coding manual into human-readable labels and performs the following tasks:

- Extracts tier labels, acoustic measurements (e.g., formants, segment durations), and contextual information.
- Converts Spanish diacritics (e.g., accents) and IPA codes for seamless compatibility with downstream analysis in **R**.
- Handles multiple tokens within a single word and calculates additional metrics, such as segment and word durations.
- Generates a structured CSV file containing fields for speaker metadata, phonological forms, surface forms, contextual sounds, and acoustic measurements at 11 equidistant time points (F1–F5).

The extraction script reduces manual errors, streamlines the data preparation process, and ensures consistency across all files. Detailed comments embedded in the script outline its functionality and modifications.

Step 2: Data Cleaning with R Script

Once the data has been extracted into a CSV file, the second step is data cleaning and transformation in **R**, using the script `Create_liquids_master_df.Rmd`. This script further refines the data and performs the following tasks:

Data Tidying - Removes unnecessary whitespace, fills undefined values with `NA`, and corrects inconsistent labels.

Token Categorization - Separates word classes (e.g., `verb`, `noun`), creates phonological and surface form types (e.g., `lateral`, `rhotic`), and revalues contextual sound categories (e.g., `vowel`, `consonant`).

Feature Engineering - Adds derived variables such as:

- Speech Rate - Calculates milliseconds per syllable based on word duration and syllable count.
- Lexical Frequency - Counts occurrences of each word in the dataset.
- Phonetic-Phonological Agreement - Indicates whether surface and phonological forms align.
- Contextual Features - Extracts manner and place of articulation for preceding and following sounds, as well as vowel height and displacement.

Data Integration - Merges linguistic data with speaker metadata (e.g., sociolinguistic background) for richer analysis.

Export - Saves the cleaned and aggregated data to a new CSV file for downstream analysis.

This two-step process ensures that the final dataset is both comprehensive and ready for variationist and acoustic analysis.