To Fill or Not to Fill?
Filled and Silent Pauses in North Carolina Spanish

Allyson Gee
Department of World Languages and Cultures

NC STATE UNIVERSITY

Filled and Silent Pauses

- Pauses, both filled and silent, are natural and abundant features that serve discourse functions in the **production** and the **perception** of speech (Erker & Bruso, 2017).
- Filled pauses (**FP**s) can be lexical or phonological units.
  - English: [ə(m)], [a(m)], *so*, *like*
  - Spanish: [e(m)], [a(m)], *este*, *bueno*, *osea*, *entonces, y*, *como*.
- Silent Pauses (**SP**s) are common in everyday speech but excessively long occurrences disrupt comprehensibility and often mark ***disfluency*** (García-Amaya, 2022).

Relevance and Hypothesis

- Due to their prevalence FPs are among the **first borrowed** elements in an environment of language contact (Erker & Bruso, 2017).
- Previous research related to borrowed FPs and discourse markers centers on **established bilingual communities** (see Erker & Bruso, 2017; Lipski, 2005).
  - These previous studies find that more exposure to English contact correlates with higher use of centralized vowels, [a] and [ə], and more English discourse markers such as *so*.
- NC is a region of new growth of Hispanic populations which affords the documentation of linguistic contact phenomenon in the **early and intermediate stages** (Michnowicz et al., 2023).

1) Speakers with higher exposure to English, G2 women, should produce more English-like FPs.
2) More English-dominant bilingual speakers will produce longer SPs.

Purpose of the study

To document possible **contact induced variation** of filled and silent pauses in NC Spanish speech as speakers incorporate themselves into the **dominant English culture**.

Research Questions

1) What is the distribution of FP and SP variants in NC Spanish?

2) How do FPs pattern according to origin and generation?

3) Do SP durations reflect the level of language dominance and/or integration into the majority speech community?

Methodology

- 15 sociolinguistic interviews from *El Corpus del Español de Raleigh-Durham* – includes **G1** and **G2** speakers from **Mexico** and **Central America.**
- Coded in **Praat**: 10 minutes starting at 2 minute mark.
  - **SPs** of 100 ms or more within an utterance.
  - Lexical FPs coded according to the lexical unit: *bueno*, *como*, *este*, *osea*, *so*, *y*, and *entonces*.
  - Phonological FPs coded impressionistically according to vowel and nasal – [e(m)], [a(m)] or [ə(m)].
- **2695 tokens** were pulled using a PRAAT script and then analyzed with the phonR, lme4, partykit packages in **R**.

Results

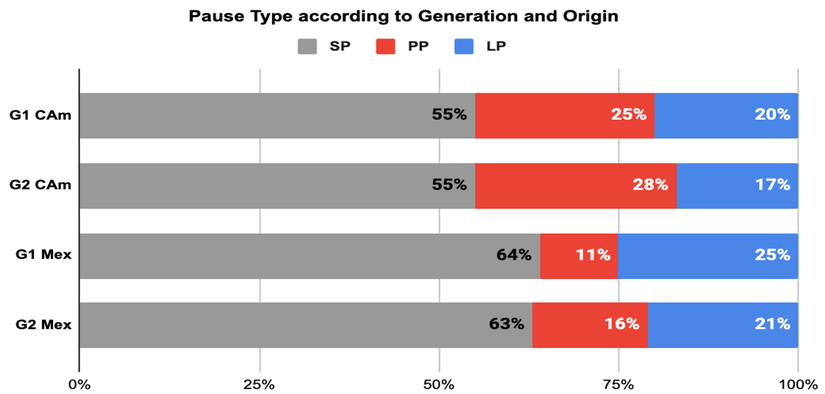

Figure 1. Mex speakers produce more SPs and CAm speakers prefer **PPs**. G2 speakers produce more **PPs** and less **LPs** regardless of origin.

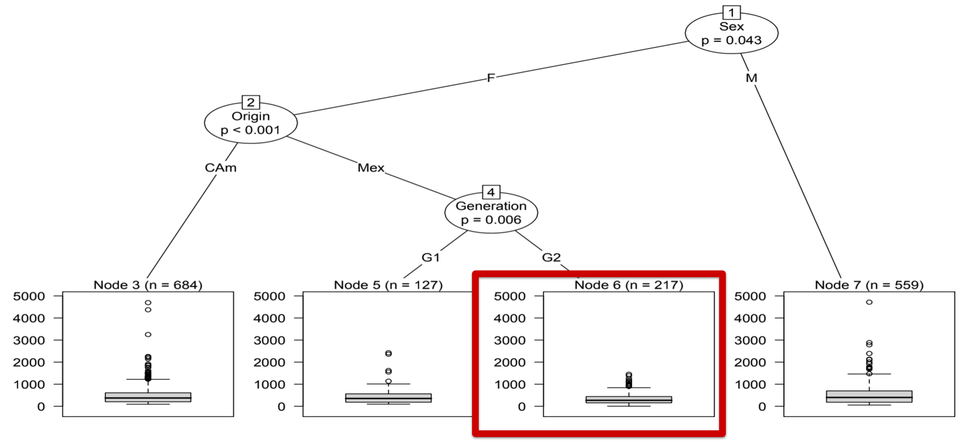

Figure 2. Conditional Inference Tree of **SP duration**

Silent Pauses (59% of data)

- Men produce longer SPs than women.
- **G2 Mexican Female** speakers produced the shortest SP tokens of all speaker groups (node 6).

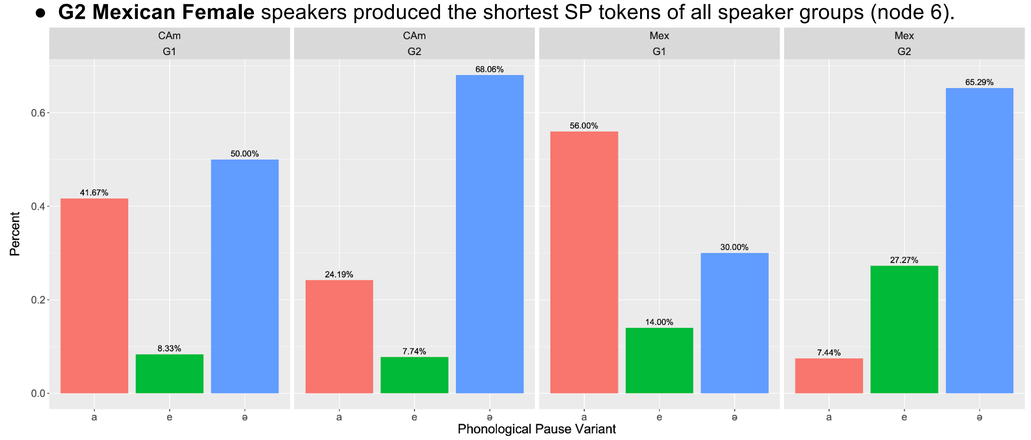

Figure 3. Boxplot of **PP type** according to *generation* and *origin*.

Phonological Filled Pauses (21% of data)

- Significant main effects were found for *generation*, *origin*, and *sex* in a mixed effects model of [e] vs central vowels, [a] and [ə].
- Nonnative [ə] present in all speaker groups but increases among G2 speakers.
- [e] decreases for G2 CAm speakers but increases among G2 Mexican speakers.

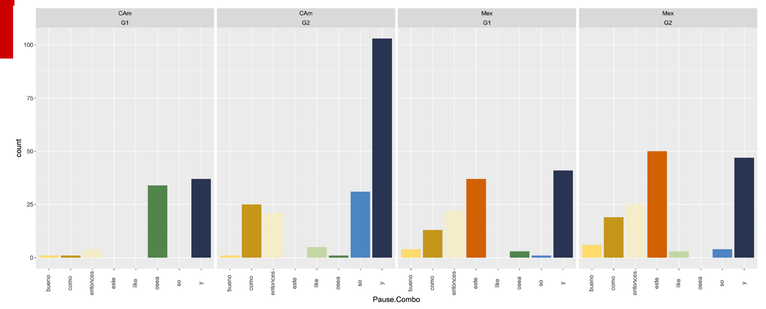

Figure 4. Barplots of **LPs** according to generation and origin

Lexical Filled Pauses (20% of data)

- LP choice among CAm speakers increase in variety among G2, especially *so* and *como*.
- *Este* tokens only occur among Mex speakers and increase in popularity among G2.
- *Como*, already common filler for G1 Mex speakers, increases for second generation speakers.
- There were no Mex *so* tokens and only one among CAm speakers.

Discussion

- **Shorter SP** duration among **G2 speakers** does not align with previous findings (Cuartero et al., 2023; García-Amaya, 2022).
  - Mexican, especially G2 speakers, may be assimilating English SP norms
    - Standard English SP is **493 ms** compared to Standard Spanish SP of **587 ms** (Campione & Véronis, 2002).
    - Average Female Mexican SP among my participants is 393 ms.
- The most **popular PPs** were **[a]** and **[ə]** among all demographics.
  - G1 speakers from both origins use English **[ə]** filler.
  - Still, **G2 use higher rates of [ə]** overall.
  - Fronted [e] increases for Mexican speakers in G2.
- **LPs** vary due to both **origin** and **generation.**
  - G1 CAm have little variety in LP distribution, but this increases for G2.
    - G2 increases use of *so*, an English LP, and *como* which is a calque of English LP *like*.
  - G1 & G2 Mex speakers demonstrate LP choice variety.
    - *como* LP established in G1 and increased in G2
  - There is an absence of *so* tokens among both generations but an abundance of *este*, especially G2.
    - *Este* is a dominant LP in Mexico City Spanish but is not salient in other varieties such as the Medellín variety (Graham, 2018)

Conclusions

NC Spanish is a complex case for Spanish linguistic variation study.

- The emerging norms related to pause behavior in this region align with previous studies of FPs because we find that PPs do become more **centralized** among speakers with more English contact.

Still, this region offers new insight about the complexities of contact variation, which is **not a homogenous process**.

- ***So*** is a dominant LP among speakers in other established areas contact regions in the US – New York, Boston, or Chicago.
- ***Este*** reduced in popularity in other regions with higher contact of English but may be an emerging marker of Mexican pride in NC.