Perceptual salience of creak and duration as prosodic boundary cues in english and spanish

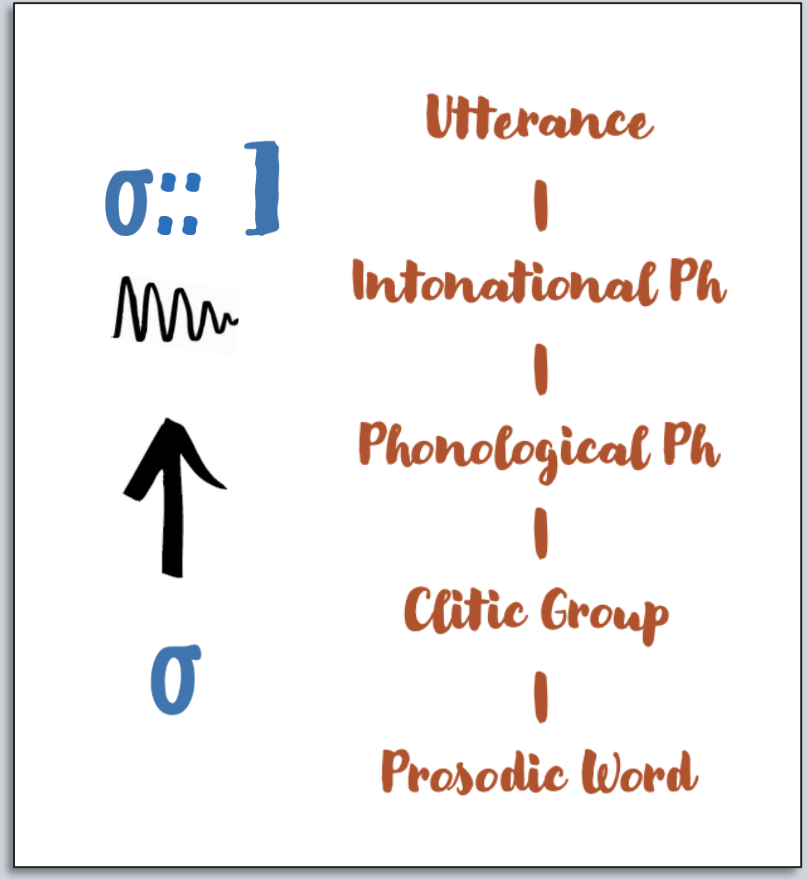
ELIZABETH WOOD, MEGAN CROWHURST AND MACKENZIE WALTERS
The University of Texas at Austin

96th Annual Meeting of the Linguistic Society of America
Washington, D.C.
January 8, 2022

Background

- *Phrase-final lengthening:* Syllables in speech are lengthened at ends of prosodic constituents in English,[1] Spanish,[2] and possibly all languages in some form.[3]

- *Phrasal creak:* Ends of prosodic constituents in speech can also be marked by laryngealization (e.g., creaky voice).[4] PC is most prevalent in higher-order prosodic constituents (especially Intonational Phrases and Utterances).[4]



Goals

- Crowhurst (2018) reports that English speakers can use both phrase final lengthening and phrasal creak cues to locate medial Phonological Phrase (PhP) boundaries,[5] but the perceptual salience of these cues for Spanish speakers is poorly documented.

- We set out to…

  a) Replicate Crowhurst (2018) for English

  b) Investigate Spanish speakers' use of final lengthening and creak cues to identify medial PhP boundaries using a similar study design, and

  c) Study how English/Spanish bilinguals might pattern compared to monolingual speakers of both languages.

Overview of the Experiment

- Stimuli were structurally ambiguous English and Spanish sentences in which an {X and Y} expression was followed by a syntactic complement. *Examples:*

  (1) They had {burgers and french fries} with ketchup.

  (2) Llevaban {faldas y blusas} de colores. *(Trans: They wore colored skirts and blouses)*

- (1) and (2) (etc.) can have a conjoined N structure, (3), where the syntactic complement has scope over both X and Y (the "together" reading); OR, they can have a conjoined NP structure, (4), where the complement applies only to Y (the X and Y are "separate" reading). A PhP boundary follows the X term in the "separate" but not in the "together" reading.

  (3) [N and N]$_N$ + PP — burgers and french fries]$_{PhP}$ with ketchup
  "Together" reading; both items have ketchup

  (4) [NP and [N PP]$_{NP}$]$_{NP}$ — burgers]$_{PhP}$ and french fries]$_{PhP}$ with ketchup
  "Separate" reading; only the french fries have ketchup

- ***Expectation:*** If participants use final lengthening and/or creak to locate medial PhP boundaries, then associating these cues with the X term should increase the odds of "separate" interpretations.

Participants

Groups

- 22 adult native American English speakers
- 25 adult native Spanish speakers, residents of Sonora, Mexico, who reported no L2 fluency
- 19 English/Spanish bilinguals (learned both languages before age 6 and use both in normal life now)

Recruitment

- Most participants were recruited through the researchers' networks and tested online on *gorilla.com*.
- 6 English monolinguals were recruited on *Prolific.com*.
- Paid for their participation

Stimuli

- Baseline versions of 3 English and 3 Spanish sentences were recorded by a female native English/ Spanish bilingual who kept the “together” reading in (3) in mind (no PhP boundary after X).
- f0 was resynthesized and normalized to remove any intonational cues
- Modal series: The X term was modally voiced. Duration was manipulated to produce a 5-step scale – baseline (as recorded) and 4 tokens in which X was lengthened by increments of ~ 25 ms
- Creaky series: identical except that the modal X term was replaced by a naturally creaky X term
- Full set: 10 tokens per sentence (5 with modal and 5 with creaky X); 3 sentences x 10 tokens = 30 utterances for each language.

Testing procedures

- Participants saw slides like **Fig. 1**. Images for the "together" and "separate" readings were paired with response keys on a standard keyboard.
- *Task:* 2-alternative forced choice; after hearing each token, participants were to press the key for the reading they preferred.
- Every sentence was presented in 5 stimulus blocks, each consisting of a slide and the 10 tokens for that sentence. Block and token order were randomized by the *gorilla* program.

Fig. 1. Example of slide shown to participants.

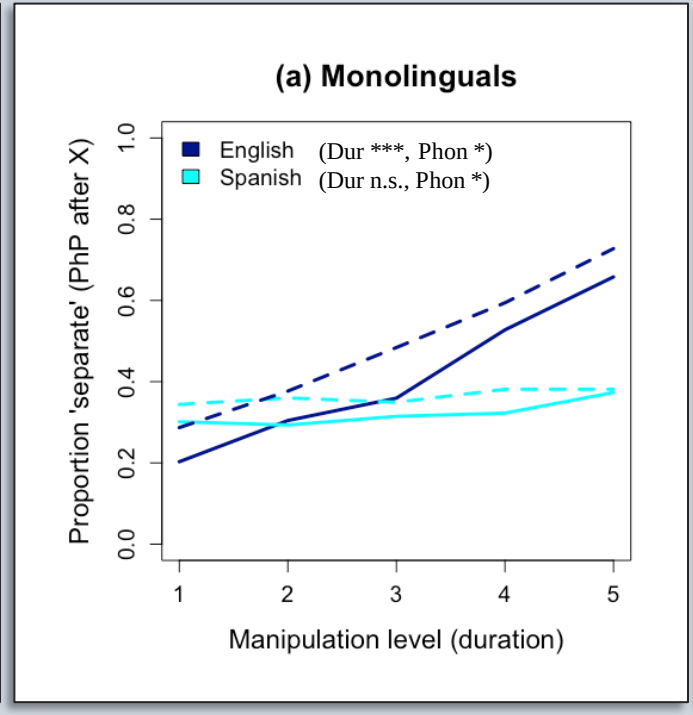
- Monolinguals heard the sentences for their language; bilinguals heard the full set, counterbalanced for which language came first.

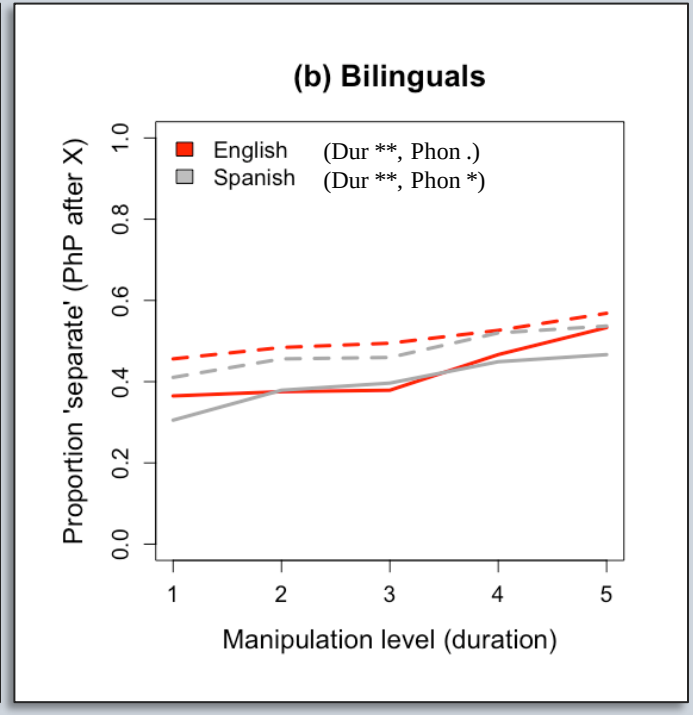
Statistical procedures

- Repeated measures; 150 observations per language per participant (3 sentences x 10 tokens x 5 repetitions)
- Linear mixed effects models (*glmer* in R), constructed separately for English and Spanish.
- DV measured proportion of "separate" decisions (implied the perception of a PhP boundary after X).
- *Variables:* Background (BG; mono- vs. bilingual), Duration (Dur; 5 levels), Phonation (Phon; modal vs. creaky); interaction BG*Dur. An optimizer, "bobyqa", was included.
- Random intercepts were included for Participant, Item (sentence); random slopes for Dur, Phon.

Results (fixed effects)

Fig 2. Proportion of 'separate' decisions as a function of varied Duration and Phonation. (Solid lines: modally voiced X; dashed lines: creaky X.)





Results (fixed effects)

- **Duration** (fixed effect). English: significant (β=.186, SE=.084, z=2.222, p=0.0263); Spanish: highly significant (β=.156, SE=.046, z=3.402, p< .0007).
- **Phonation** fixed effect. English: significant (β=.390, SE=.128, z 3.042, p= .0024); Spanish: highly significant (β=.336, SE=.101, z=3.328, p= .0009). (See legends in Fig. 2 for detail.)

Translating…

1. For all groups, the odds of a "separate" decision were lowest at baseline when term X was modally voiced. See Duration level 1 on solid lines in Figs. 2 and 3.
2. The fixed effects for Dur and Phon mean that overall, the odds of a "separate" decision were significantly higher **(a)** as X got longer (see differences along the x-axis in Figs. 2 and 3), and **(b)** when X was creaky (see the dashed lines in Fig. 2a and 2b).

Results: interaction

- **The interaction BG*Dur.** Significant for both languages:
  - English (β = 0.372, SE = 0.115, z = 3.248, p = 0.0012)
  - Spanish (β = -0.120, SE = 0.061, z = -1.959, p = 0.0501)
- The significant interaction means that mono- and bilinguals responded differently to varied duration.

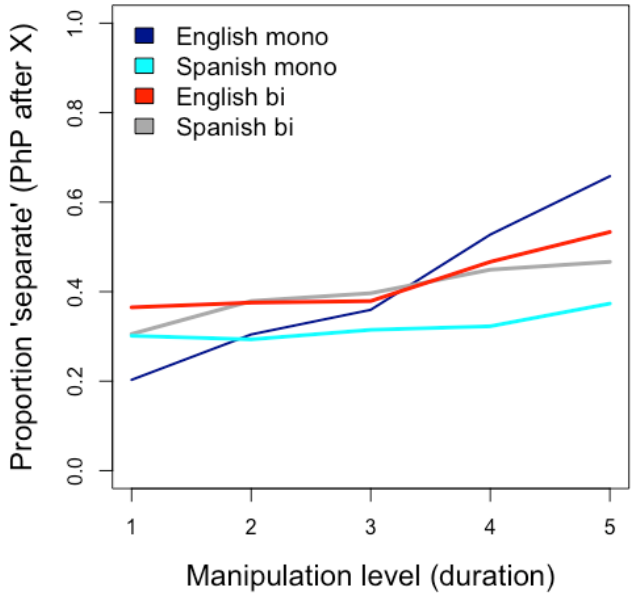


Fig. 3. ‘Separate’ decisions as a function of varied Duration (modally voiced)

Results: interaction

- Fig. 3 reveals that ***bilinguals responded differently from monolinguals in both their languages***:
- For Spanish, the proportion of "separate" decisions was higher overall than in the monolingual group (grey and cyan lines).
- For English, the positive trend was weaker than in the monolingual group (red and dark blue lines).

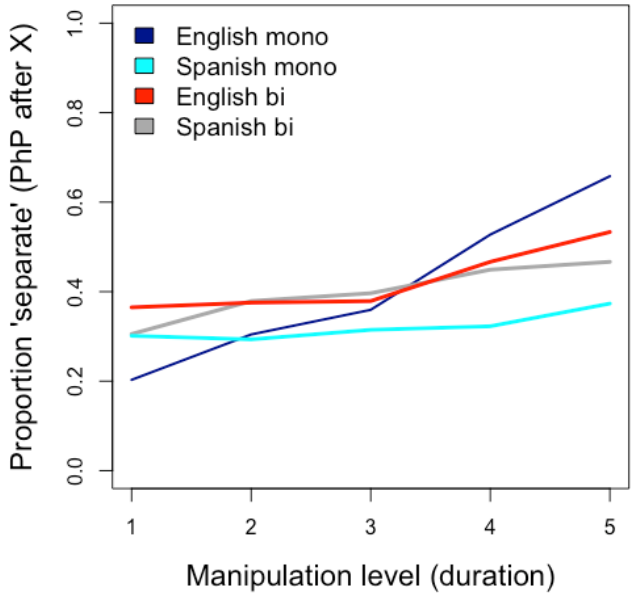


Fig. 3. 'Separate' decisions as a function of varied Duration (modally voiced)

Take-away points

- This study replicates findings for English reported in Crowhurst (2018) with a larger stimulus set.
- The study provides
  a) new information about the perceptual salience of final lengthening for Spanish speakers,
  b) the first evidence that creak is perceptually salient to Spanish speakers performing a linguistic task, and
  c) confirmation that bilinguals process language differently from monolinguals.

References

[1] Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *JASA* 91, 1707–1717.

[2] Ortega-Llebaria, M. & Prieto, P. 2007. Stress and focus in Spanish and Catalan: patterns of duration and vowel quality. In Prieto, P., Mascaró , J., & Solé, M.J. (Eds.), *Segmental and Prosodic Issues in Romance Phonology*. John Benjamins, NL, pp. 155–175.

[3] Fletcher, J. 2010. The prosody of speech: Timing and rhythm. In *The handbook of phonetic sciences*. Blackwell. pp. 521-602.

[4] Redi, L. & Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*, 407-429.

[5] Crowhurst, M. J. 2018. The joint influence of vowel duration and creak on the perception of internal phrase boundaries. *JASA*, 143(3), EL147-EL153.