Do early Spanish-English bilinguals use more low-frequency words and cognates than late bilinguals?

Melissa Castillo[1,2], Kelly Medina[1,2], Daniela Castillo [1,2], Eve Higby[3]

[1]CUNY Graduate Center, [2]Queens College, [3]University of California, Riverside

Motivation

- Researchers often assess language production abilities using verbal fluency tasks in which participants produce as many exemplars as they can to a semantic or phonemic category within 60 seconds.
- Bilinguals' performance on these tasks have been compared to monolinguals', but it is unclear whether early and late bilinguals differ on these tasks.
- Age of second-language (L2) acquisition affects vocabulary size and lexical organization, with early bilinguals showing a more integrated semantic network (Hernandez, Li, & MacWhinney, 2005).
- Early bilinguals have been shown to produce just as many responses in their dominant language as monolinguals (Blumenfeld et al., 2016) or fewer responses (Sandoval et al., 2010). Among late bilinguals, those with lower L2 proficiency produced as many responses in the dominant language as monolinguals, while highly proficient bilinguals produced more responses than monolinguals for both semantic and phonemic categories (Mathison, under review). In the less-dominant language, late bilinguals produce fewer responses (Van Assche et al., 2013).
- Early bilinguals produced more low-frequency words in the semantic category than monolinguals but they did not differ for the phonemic category (Sandoval et. al., 2010). Both groups produced lower-frequency words on the phonemic task than the semantic task.
- Both bilinguals and monolinguals tend to produce more cognates in the phonemic than semantic task. However, some have reported higher production of cognates in English than Spanish (Blumenfeld et al., 2016) while others found the same rate of cognates in both languages (Sandoval et al., 2010).
- Bilinguals living in the U.S. typically produce more correct responses in English than in Spanish (Beatty-Martinez et al., under review; Blumenfeld et al., 2016; Sandoval et al., 2010).
- In sum, comparisons across studies paint a somewhat mixed picture of bilingual word production on verbal fluency tasks and potential age of acquisition effects. Differences in the categories chosen (Tombaugh et al., 1999) and the environment that the bilinguals are in (Beatty-Martinez et al., under review) may help explain these differences.

Purpose

The purpose of this study was to determine whether the age of second-language acquisition influences the type and number of words retrieved during semantic and phonemic fluency tasks in the two languages of highly proficient bilinguals.

Predictions

- Early Spanish-English bilinguals would produce more responses in English than late bilinguals, reflecting greater English dominance and a larger English vocabulary.
- We expected lower-frequency responses on the phonemic task than the semantic task, but no difference between early and late bilinguals.
- Early bilinguals would produce a similar number of cognates in both languages. Late bilinguals would produce more cognates in English, their L2, compared to their Spanish.

Methods

Participants:

|	Early Bilinguals	Late Bilinguals
N	23	19
Sex (Female; Male)	13; 10	9; 10
Age (mean, SD)	25.5 years (6.6)	28.6 years (5.4)
Age of English Acquisition (mean, SD)	4.9 years (2.2)	13.9 years (3.1)

Task:
- Semantic categories
  - English: Animals and Clothing
  - Spanish: Fruits and Vegetables, and Furniture
- Phonemic categories
  - English: Letters D and N
  - Spanish: Letters M and F

Analysis:
- Responses were transcribed from recordings. Word frequency per million words (log-transformed) were obtained from the SUBTLEX database in English (Brysbaert & New, 2009) and Spanish (Cuetos et al., 2011).
- Cognates were considered those with a significant degree of overlap between the response and its closest English translation (e.g., *bicycle and bicicleta*).
- Separate 2x2x2 ANOVAs (Group: Early and Late Bilinguals, Task: Semantic and Phonemic, Language: English and Spanish) were conducted for number of correct responses, log-transformed word frequency, and percent cognates.
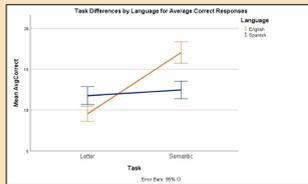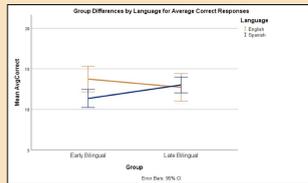
Results

Average Correct
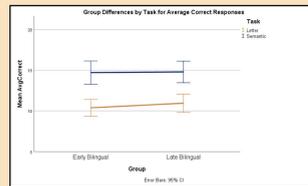
Group Difference by Task
Both groups produced more correct responses in the semantic task (m = 14.74) than they did in the phonemic task (m = 10.64) (p <.001).
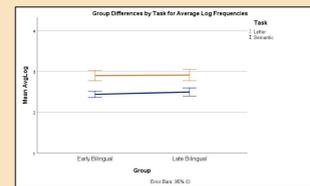

Task Differences by Lang.
Both groups produced more correct responses in Spanish (m = 11.75) than in English (m = 9.52) on the phonemic task, but more correct responses in English (m = 17.04) than in Spanish (m = 12.45) on the semantic task (p < .001).
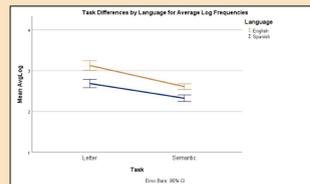

Group Differences by Lang.
Late bilinguals produced almost the same number of correct responses in English and Spanish, while early bilinguals produced more correct responses in English (m = 13.74) than they did in Spanish (m = 11.35) (p = .02).

Average Log Frequency

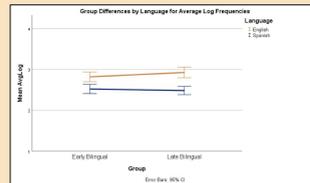**Group Differences by Task**
Both groups produced more high-frequency words in the phonemic task (m = 2.91) than in the semantic task (m = 2.47) (p <.001).
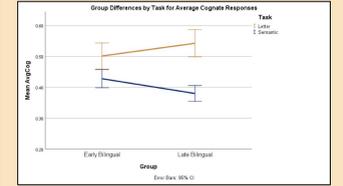

Task Differences by Lang.
More high-frequency words were produced in English than in Spanish for both the phonemic (m = 3.13) and semantic tasks (m = 2.61) (p <.001).
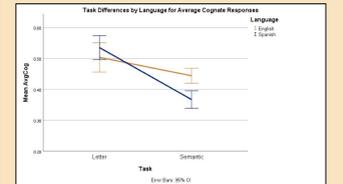

Group Differences by Lang.
Both early and late bilinguals produced more high-frequency words in English (m = 2.87) than they did in Spanish (m = 2.51) (p < .001).

Average Percent Cognate

Group Differences by Task
Both groups produced more cognates in the phonemic task (m = .52) than they did in the semantic task (m = .41), but the task difference was greater for late bilinguals (p= . 01).

Task Differences by Lang.
Both groups produced more cognates in Spanish (m = .54) than they did in English (m = .50) on the phonemic task but more cognates in English (m = .44) than in Spanish (m = .37) on the semantic task (p = .003).
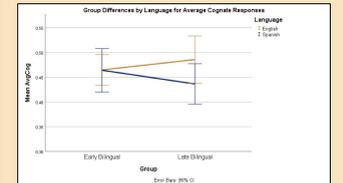
Group Differences by Lang.
Both groups produced a similar number of cognates in English (p = .17).

Findings Summary
- The semantic task elicited more responses, more high-frequency words, and more cognates than the phonemic task.
- More high-frequency words were generated in English than in Spanish and on the phonemic than the semantic task.
- Early bilinguals produced more responses in English than Spanish while late bilinguals produced a similar number of responses in both languages.
- Late bilinguals produced more cognates on the phonemic task than on the semantic task; early bilinguals produced a similar number of cognates on both tasks.

Discussion

- Despite having a later age of English acquisition, late bilinguals performed as well as early bilinguals in their L2 (English), but early bilinguals were not as proficient in Spanish as late bilinguals. Early bilinguals growing up in the U.S. are typically English dominant, but late bilinguals are usually assumed to be less proficient in their L2. We intentionally recruited highly proficient bilinguals, so this shows that late L2 learners can become balanced in both languages.
- The phonemic task may have been more difficult than the semantic task as seen in the higher production of cognates and high-frequency words. In the process of lexical retrieval, semantic associations between words are activated. Participants can rely on that semantic activation to optimize their performance on the semantic task, but the semantically related words may need to be inhibited during the phonemic task (Shao et al., 2014).
- In contrast to Sandoval et al. (2010), bilinguals produced more high-frequency words in the phonemic than in the semantic category. This discrepancy could be due to task differences as they used double-letter categories (e.g., ab-), which are more restrictive than single letters (e.g., f-).
- The higher frequency of responses and more cognates produced on the phonemic task versus the semantic task may reflect strategies employed when the task is more difficult. Bilinguals may be able to turn to their shared lexical representations (i.e., cognates) when the task is more challenging.

Conclusion
- The results reveal a complicated set of interactions between age of acquisition and task demands.
- The way the lexicon is organized for bilinguals has implications for the accessibility of certain words (such as cognates), and lexical organization is likely to differ for early and late bilinguals.