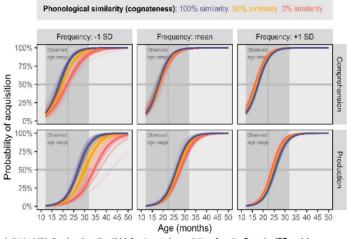
## Cognateness, frequency, and vocabulary size: an interactive account of bilingual lexical acquisition

Gonzalo Garcia-Castro<sup>1</sup>, Daniela S. Avila-Varela<sup>1,2</sup>, Ignacio Castillejo<sup>3</sup>, Núria Sebastian-Galles<sup>1</sup>

<sup>1</sup>Universitat Pompeu Fabra (UPF); <sup>2</sup>Universidad de Lisboa (UL); <sup>3</sup>Universidad Autónoma de Madrid (UAM) gonzalo.garciadecastro@upf.edu; daniela.avila@upf.edu; ignacio.castillejo@uam.es, nuria.sebastian@upf.edu

A prominent feature of the bilingual mental lexicon is that the activation of representations spreads in a language non-selective fashion. An instance of such non-selectivity is embodied by cognateness (i.e., form-similarity between translation equivalents, such as flower-flor in English and Spanish), which impacts lexical processing in bilinguals. Recent studies have suggested that cognateness facilitates vocabulary acquisition in bilingual toddlers—who show larger vocabulary sizes when their languages share many cognates—and an earlier age-of-acquisition for cognates than for non-cognates. The specific mechanisms underpinning such facilitation are unclear. We present an account of bilingual early lexical acquisition in which cognateness interacts with lexical frequency and language exposure to facilitate word acquisition. We evaluated this model against vocabulary data from 436 Catalan-Spanish bilinguals aged 12 to 34 months. We used a Bayesian Exploratory Item Response Models to estimate participants' probability of acquisition of 604 words, conditional to the cognate status and lexical frequency of the word-form, and the age and degree of exposure to each language of the toddler. We found converging evidence for an earlier age-of-acquisition for cognate words, and for such effect being mediated by lexical frequency and language exposure. Low-frequency words, and words from the language of least exposure were more strongly benefitted by their cognate status than high-frequency words. Our findings support an interactive account of bilingual vocabulary acquisition in which the lexical representations in one language interact with the acquisition of words in the other language.



Individual (thin lines) and median (thick lines) posterior predictions from the Bayesian IRT model.