

## Examining the effect of predictability on code-switching: A production experiment

Yanting Li (yantil5@uci.edu), Gregory Scontras and Richard Futrell  
Department of Language Science, University of California, Irvine

**Introduction** Bilinguals sometimes code-switch, that is, switch from one language to another during communication [1-3]. Corpus studies show that less predictable words tend to be switched more [4-6], but behavioral evidence remains limited [7], partially due to the difficulty of eliciting voluntary code-switches with controlled linguistic context. This study presents a production experiment designed to elicit Mandarin-English code-switches and test the role of predictability, manipulated through classifiers: to quantify nouns in Mandarin, speakers can either use the generic classifier *gè* that is compatible with most nouns, or specific classifiers that are restricted to narrower ranges of nouns [8]. **We hypothesize** that nouns following the generic *gè* (and thus are less predictable) are more likely to be code-switched to English compared to those following a specific classifier.

**Method** We designed a picture choosing game in which participants believed they were collaborating with another bilingual though no partner was actually involved (see Fig 1-Left for procedure). On each trial, participants read a preamble aloud and completed it by naming the circled target picture in either language, so the “partner” can choose it (Fig 1-Right). The key manipulation was the classifier preceding the target noun: either the generic *gè* (neutral condition) or a specific classifier (predictable condition)<sup>1</sup>. The task included 30 target items. To encourage natural code-switching, we gave no instructions about language choice, but all task instructions contained switches.

100 Mandarin-English bilinguals were recruited through Prolific. Participants were excluded if their L1 was not Mandarin (N = 20), if they produced more than five incompatible classifier-noun pairs (N = 2), or if they named all pictures in only one language, suggesting that our manipulation did not work for them (N = 16). We excluded trials if the participant failed to name the target in the English picture naming task later, or if the classifier-noun pair produced was incompatible, resulting in 1,636 valid trials. Our primary analysis is a mixed-effects logistic regression model predicting the language of the noun (English vs. Mandarin) using the type of classifier produced (generic vs. specific), with random intercepts and slopes for classifier type by speaker and noun.

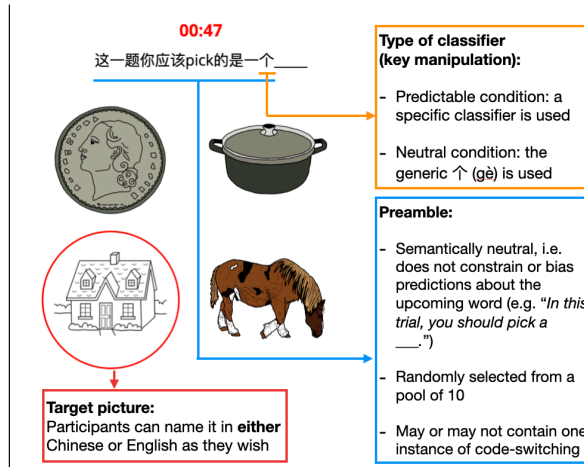
**Results** Overall, participants produced more nouns in Mandarin than in English (Fig. 2-Left), but a substantial 40.77% were produced in English, indicating that the task successfully elicited code-switching. Classifier type did not significantly affect language choice ( $p = 0.318$ ). However, less frequent nouns in Mandarin were more likely to be produced in English ( $p < 0.05$ , Fig 2-Bottom). To obtain a more fine-grained, continuous measure of predictability, we calculated the entropy over a list of 71,586 nouns [10] following each classifier using an LLM [9]: higher entropy means lower noun predictability. Replacing the categorical classifier type variable with entropy in the regression model revealed an effect in the expected direction ( $\beta = -1.206, z = -2.104, p < 0.05$ ).

**Conclusion** In this study, we behaviorally tested whether less predictable words are more likely to be code-switched during speech production. Although the manipulation classifier type did not appear to influence code-switching of the following nouns, we observed a **clear effect of word frequency**. Despite previous work showing that a frequency effect often disappears once in-context predictability is accounted for, frequency and predictability are tightly correlated, so the frequency effect is consistent with the broader predictability-based account [4-7]. Complementing this, our entropy-based measure of predictability showed the same pattern: nouns following classifiers with **higher post-classifier entropy** (and therefore less predictable) were more likely to be produced in English. Importantly, we successfully established an experimental paradigm that **elicits ecologically valid code-switches in communicative contexts with controlled meanings**, which can be used for future studies of code-switching production.

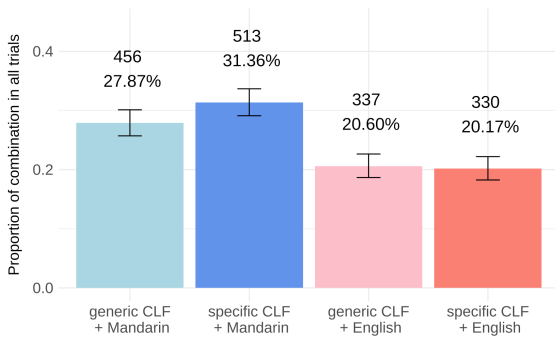
<sup>1</sup>Target nouns were confirmed to be less predictable in the neutral than in the predictable condition using an LLM [9].

**Experiment procedure:**

1. Language background questionnaire
2. Task instruction
3. Practice trials
4. Test trials (30 in total)
5. English picture naming
6. Feedback



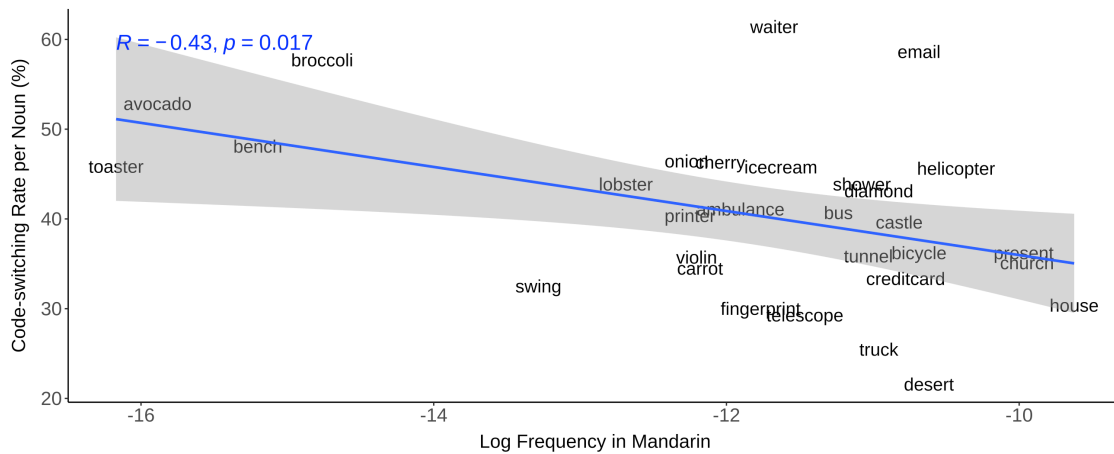
**Figure 1:** Left: Experiment procedure. The English picture naming task checks whether the participants can (and therefore have the choice to) name the target pictures in English; Mandarin is not tested as they are L1 speakers. Right: Sample test trial (box contents are explanatory and not shown to participants).



**Figure 2:** Overview of experiment results.

Left: Distribution of classifier–noun pairs with binomial confidence intervals.

Bottom: Code-switching rate per noun by log frequency<sup>[11]</sup> of nouns in Mandarin for all 30 items.



**References.** [1] Gumperz (1977) *RELC journal*; [2] Poplack (1980) *Linguistics*; [3] Myers-Scotton (1997) *uelling languages: Grammatical structure in codeswitching*; [4] Myslin & Levy (2015) *Language*; [5] Calvillo et al. (2020) *EMNLP*; [6] Bhattacharya & van Schijndel (2023) *HSP*; [7] de Bruin & Shiron (2024) *J. Exp. Psychol.*; [8] Ma (2015) *The Emergence and Development of Chinese Individual Classifiers*. [9] Zhang et al. (2021) *ScienceDirect*. [10] Qi et al. (2019) *arXiv preprint*. [11] Speer (2022) *rspeer/wordfreq: v3.0 (v3.0.2)*