A Network Science Approach to Bilingual Code-Switching

Qihui Xu[1] Magdalena Markowska[2] Martin Chodorow[1,3] Ping Li[4]

1. Graduate Center, CUNY
2. Stony Brook University
3. Hunter College, CUNY
4. The Hong Kong Polytechnic University

Qihui Xu

Magdalena Markowska

Martin Chodorow

Ping Li

qxu@gradcenter.cuny.edu

“刚才 我 不是 跟 你 讲 我 apply 那个 job”

Haven’t I told you that I applied for that job?

(Lee et al., 2017)

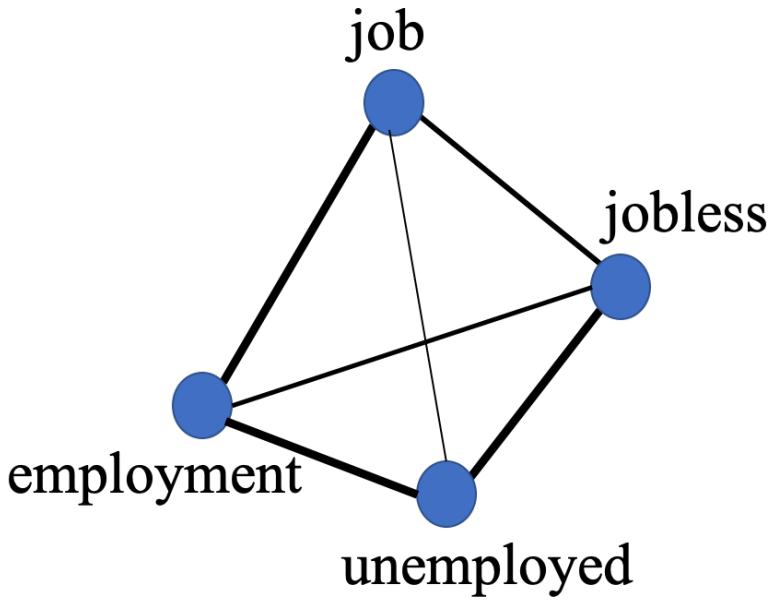
Why do bilinguals code-switch?

Lexical Accessibility

- Word frequency (Gollan & Ferreira, 2009; Gross & Kaushanskaya, 2015)
- Familiarity (Gollan et al., 2014)
- RT (de Bruin et al., 2018)

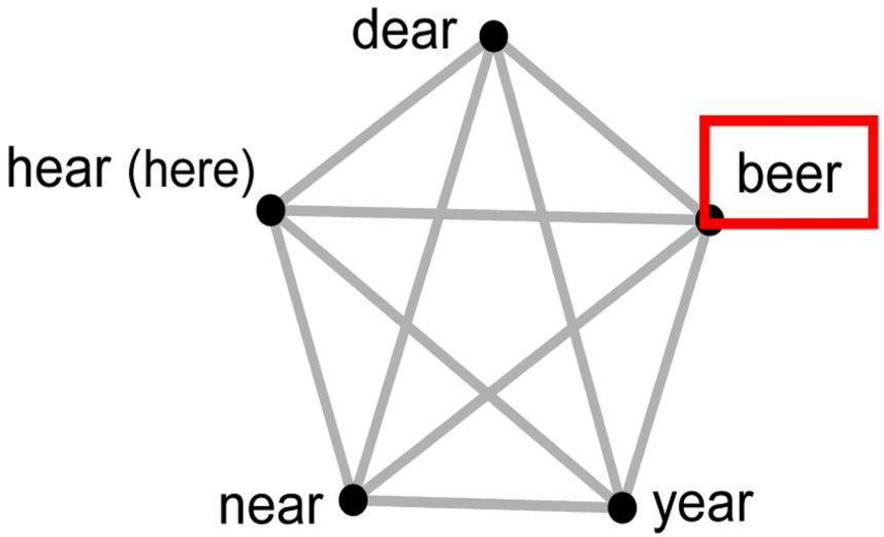
From local to global

Job



Words in network science

Clustering coefficient

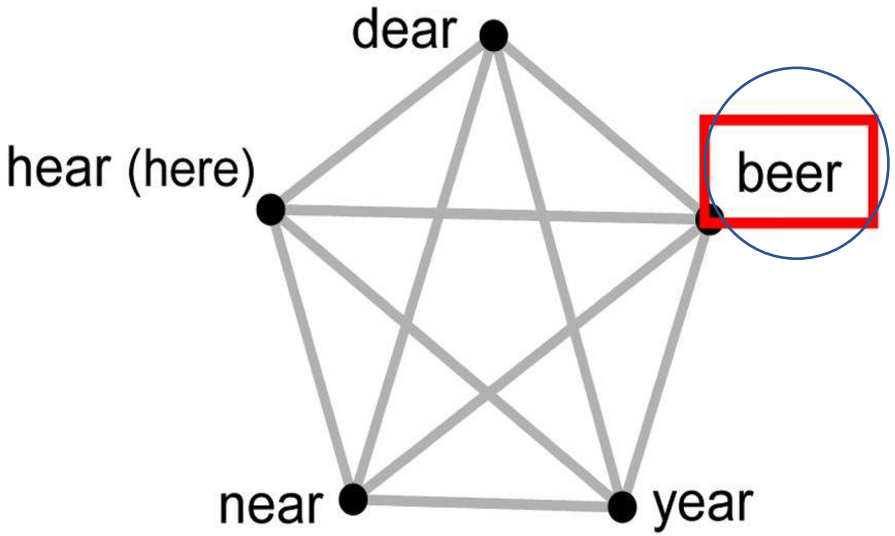


Chan & ViteVitch, 2009, 2010;

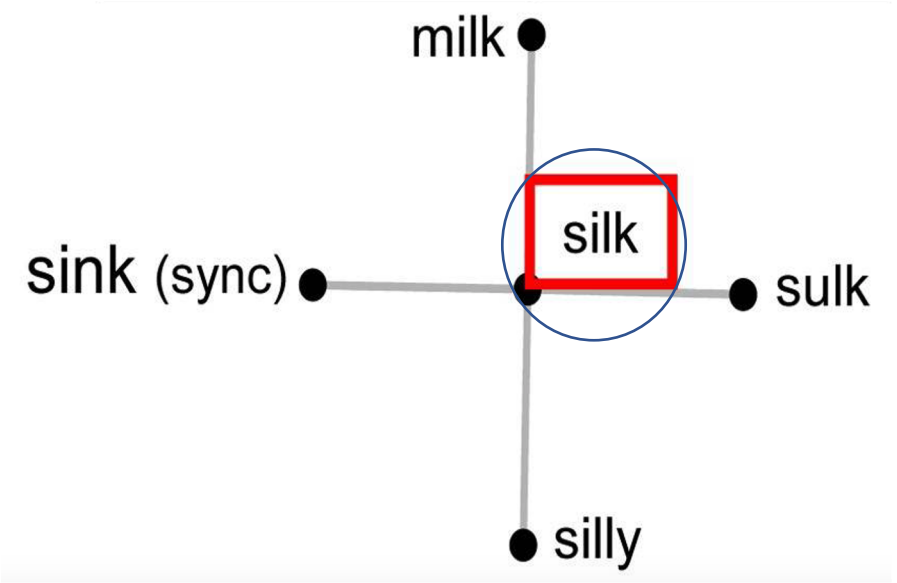
Sample image from Karuza et al., 2016

Words in network science

High clustering coefficient



Low clustering coefficient

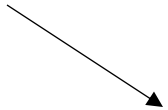


Chan & ViteVitch, 2009, 2010;

Sample image from Karuza et al., 2016

Present study

Clustering coefficient

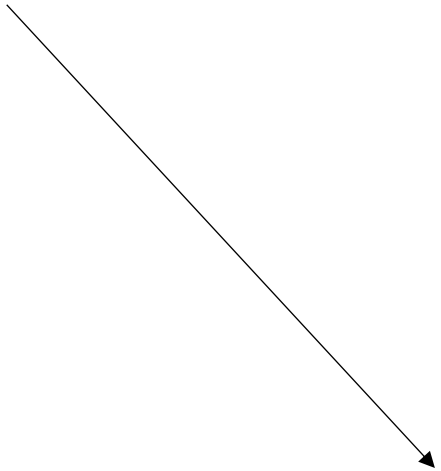


Lexical Accessibility

CS speech

Present study

Clustering coefficient



CS speech

Present study

Clustering coefficient (CS word) < Clustering coefficient (translation equivalent)

surprised

惊讶

我 很 surprised

Data

Mandarin-English Code-Switching in South-East Asia

156 balanced bilinguals

155,979 sentences

Data

刚才 我 不是 跟 你 讲 我 apply 那个 job

Haven’t I told you that I applied for that job?

Data

刚才 我 不是 跟 你 讲 我 apply 那个 job

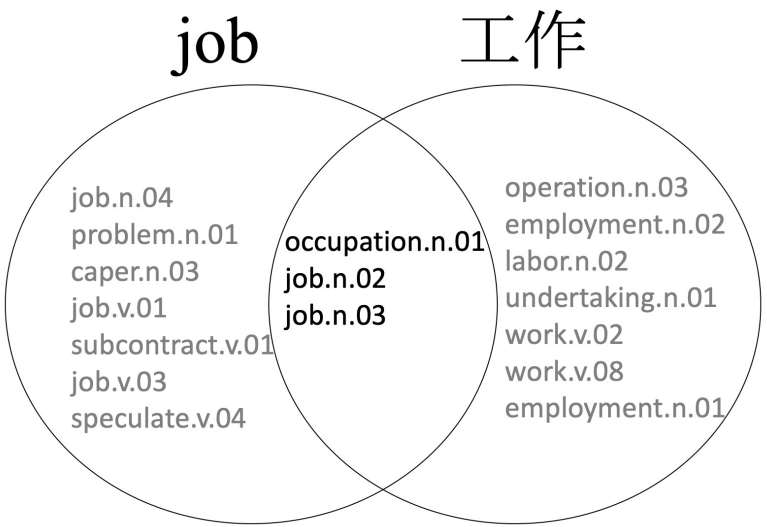
Haven’t I told you that I applied for that job?

CS words
apply
那个
job

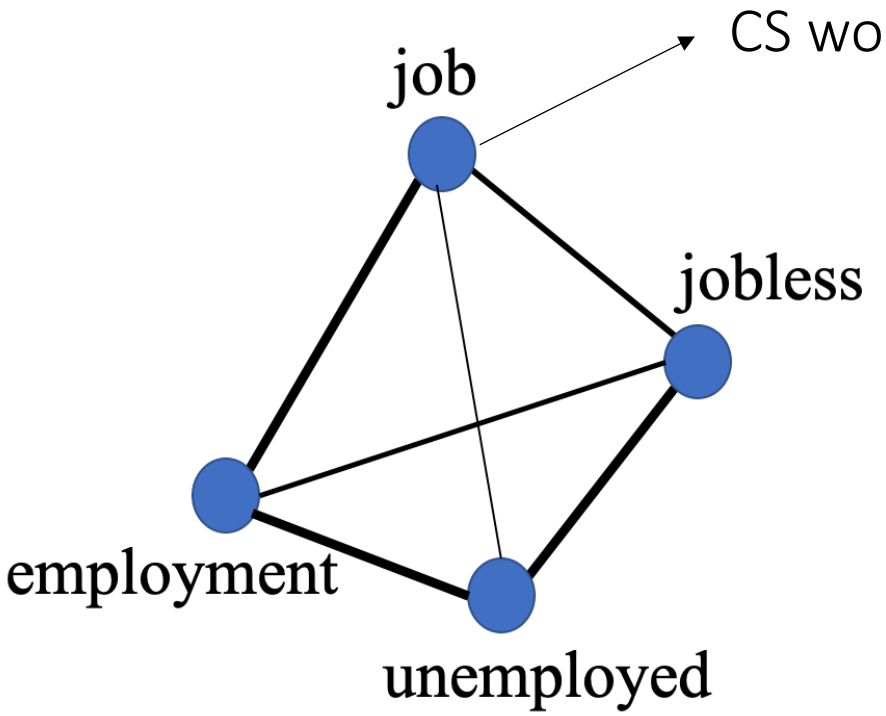
Data

刚才 我 不是 跟 你 讲 我 apply 那个 job

CS words	Translation equivalents
apply	申请
那个	that
job	工作

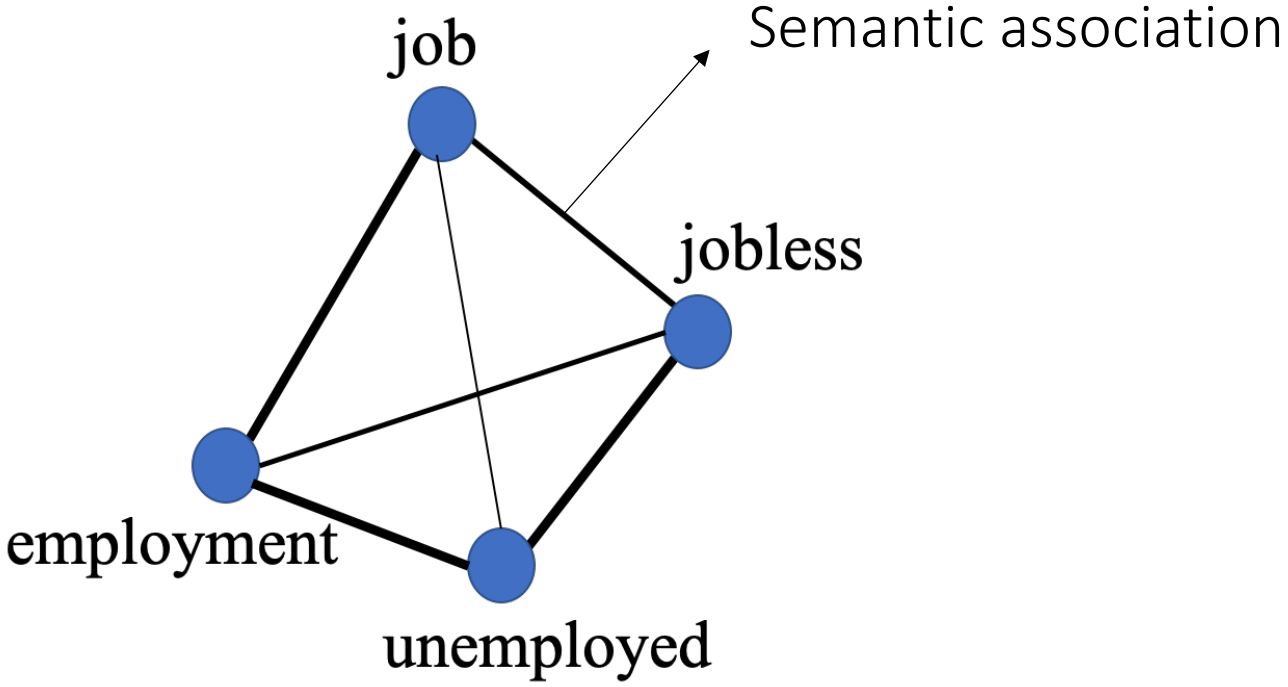


Semantic networks

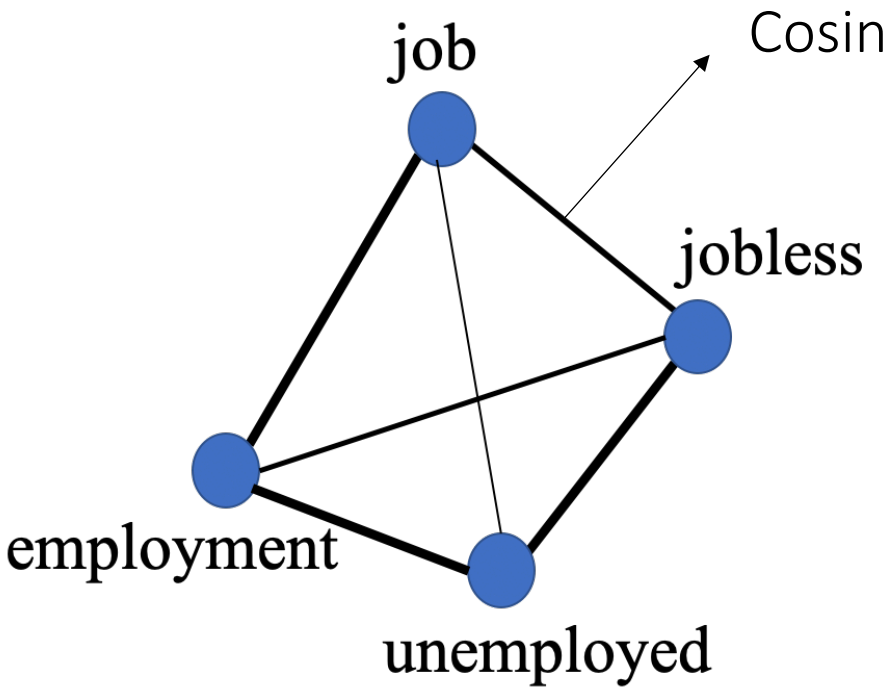


CS words & Translation equivalents

Semantic networks



Semantic networks

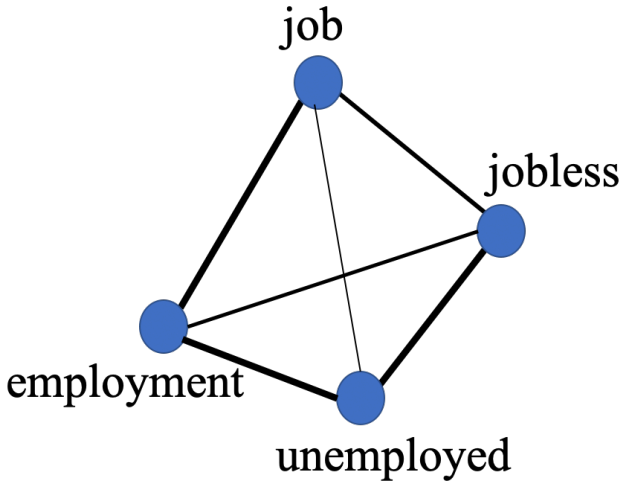


Pretrained fastText embeddings

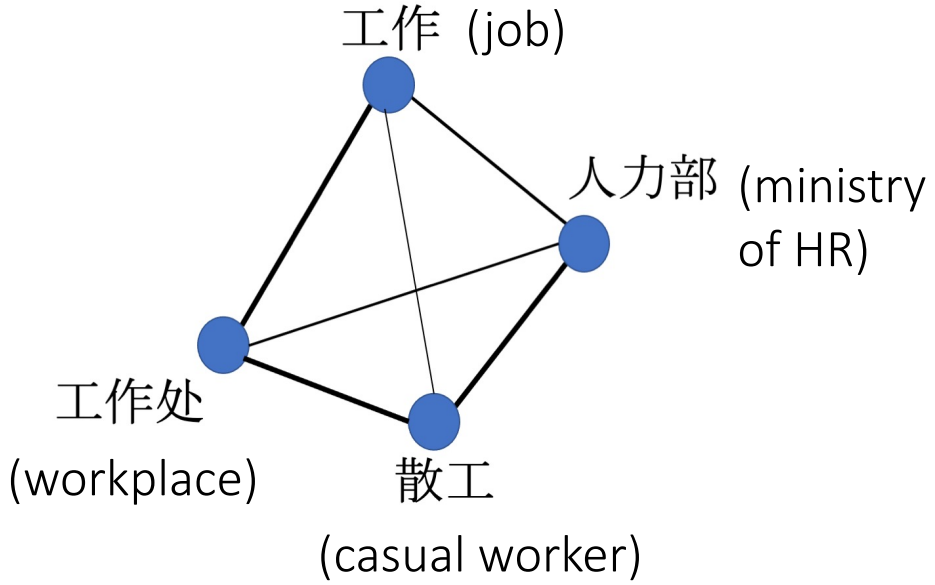
(Bojanowski et al., 2017; Grave et al., 2018)

Semantic networks

English

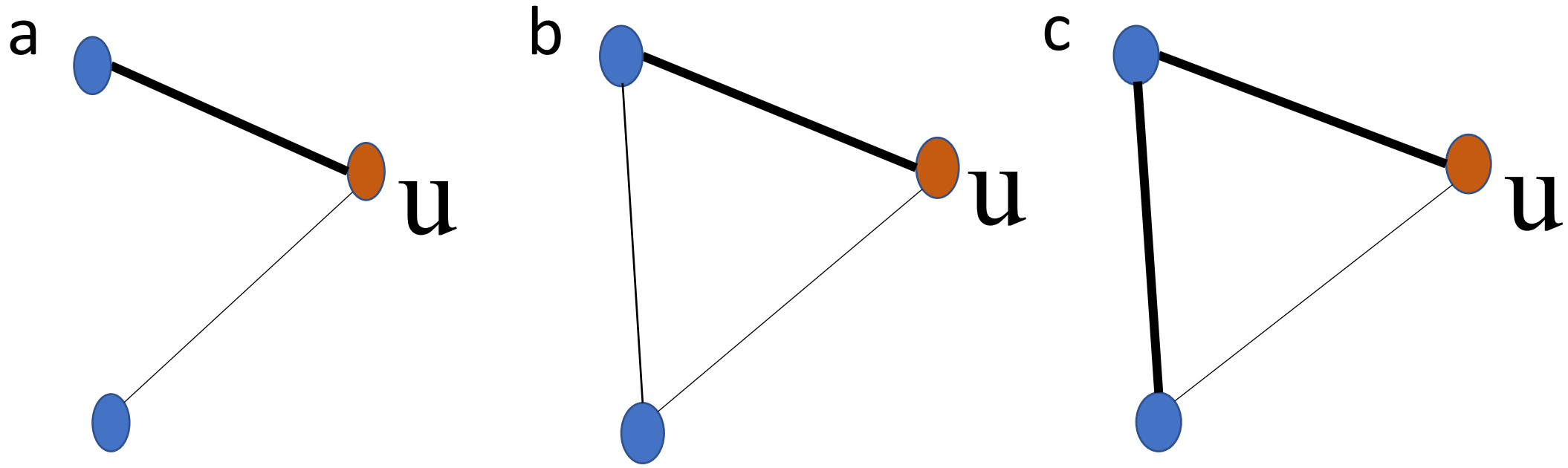


Chinese



Analysis

Clustering coefficient



Analysis

Word frequency

SUBTLEX databases of English and Chinese

(Brysbaert and New, 2009; Cai and Brysbaert, 2010)

Analysis

Data rescaling

Clustering coefficient vs. Word frequency

$r = .41$

Analysis

Clustering coefficient

Word frequency

CS word vs. Translation equivalent

English - Chinese

Chinese - English

Analysis

Clustering coefficient

Word frequency

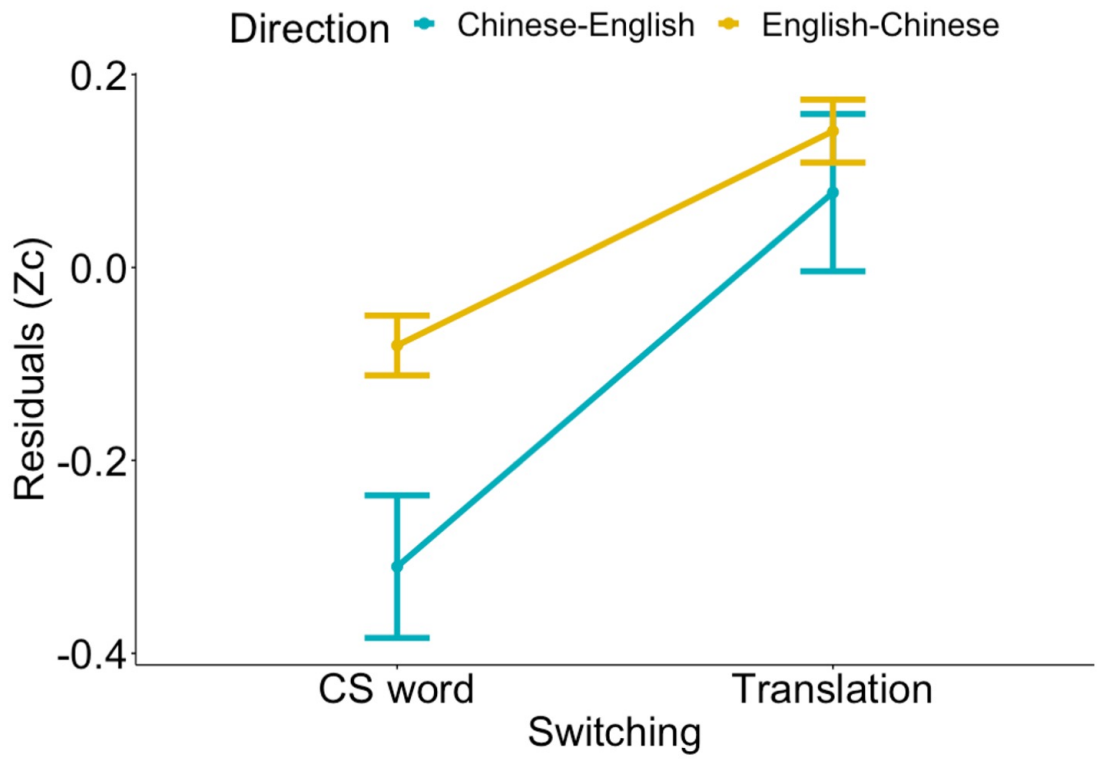
CS word vs. Translation equivalent

English - Chinese

Chinese - English

Results

Clustering coefficient

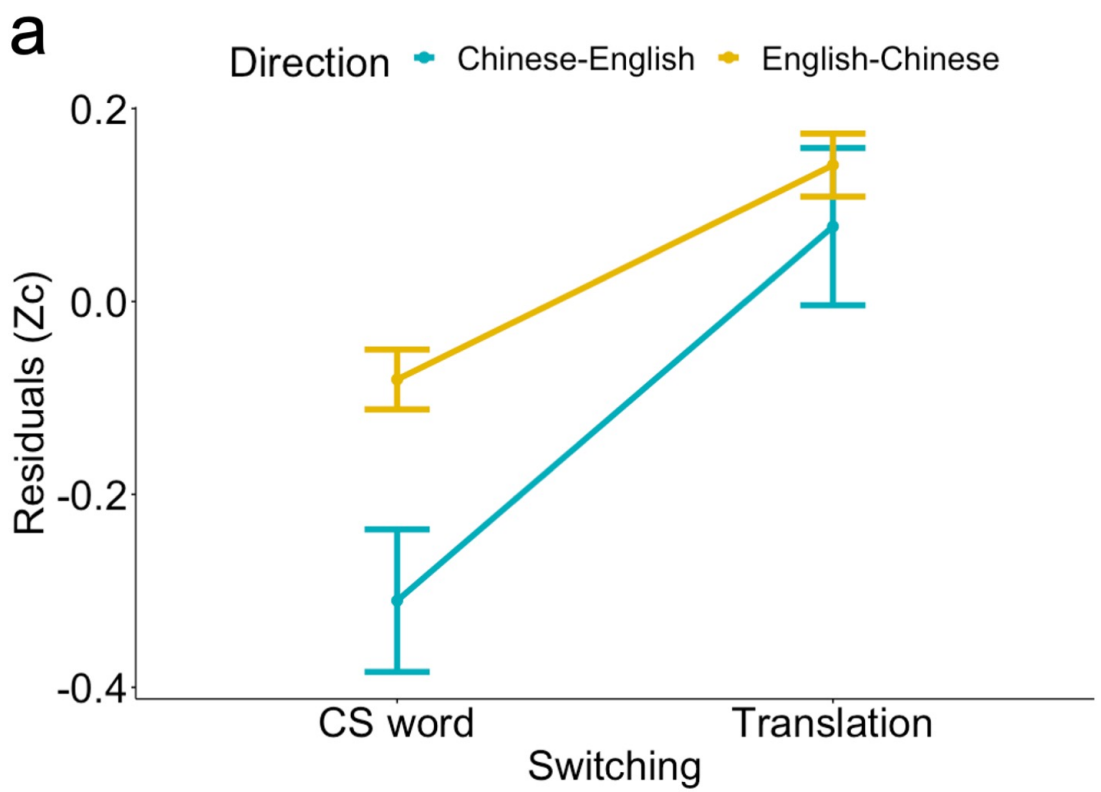


CS words have significantly **lower clustering coefficients** than their translation equivalents

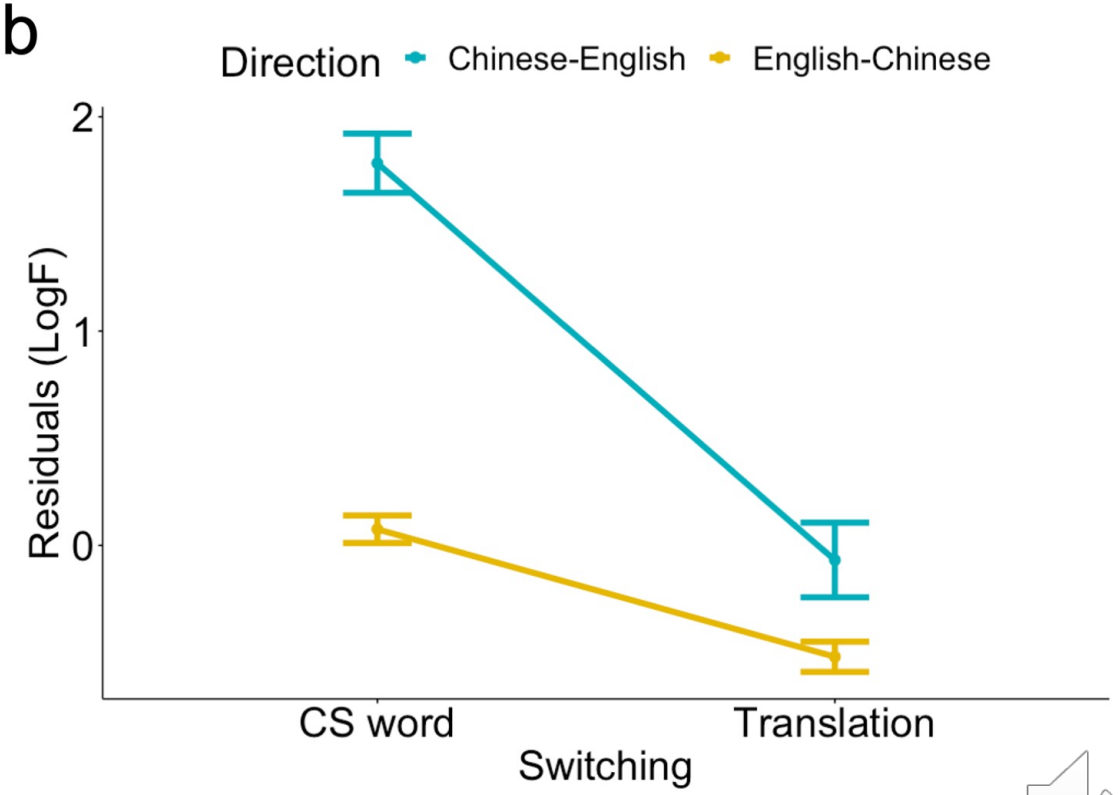
Xu, 2021

Results

Clustering coefficient



Word frequency



Xu, 2021

Conclusion

- **CS words** tend to have **lower clustering coefficients** than their translated equivalents.
- Clustering coefficient vs. Word frequency