In search of meaning in the (bilingual) lexicon

Adel Chaouch-Orozco
Research Assistant Professor at the Hong Kong Polytechnic University

LT Research Forum – June 12, 2024

Introduction

- Our languages are activated simultaneously (e.g., Kroll et al., 2006; Thierry & Wu, 2007; van Assche et al., 2012).

Introduction

- Our languages are activated simultaneously (e.g., Kroll et al., 2006; Thierry & Wu, 2007; van Assche et al., 2012).

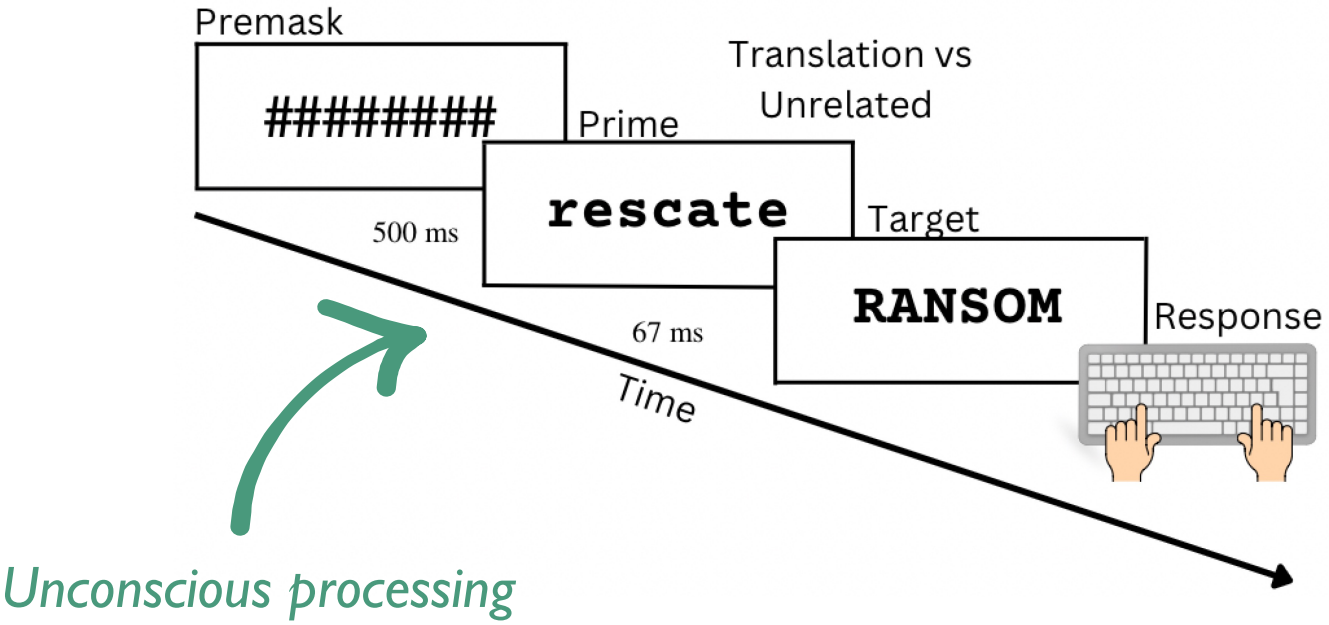
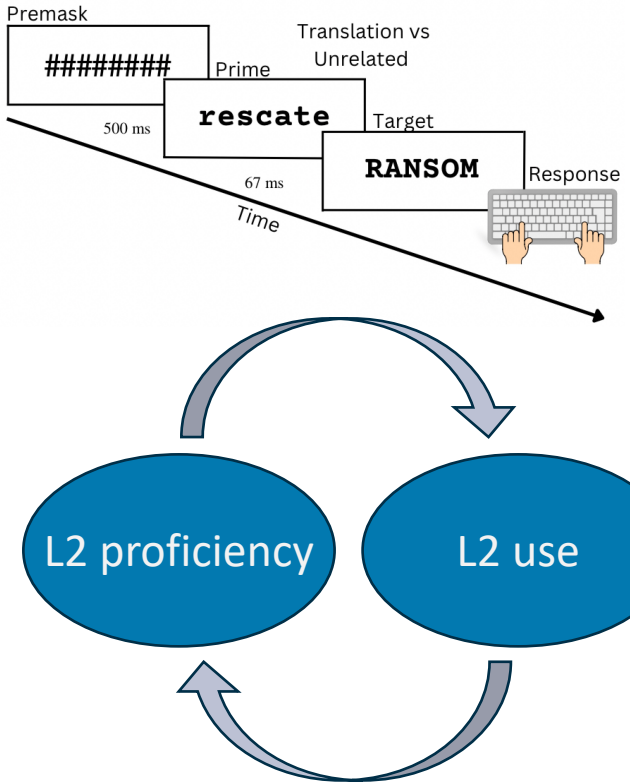


Image from Scrimshire et al. (2023)

Introduction

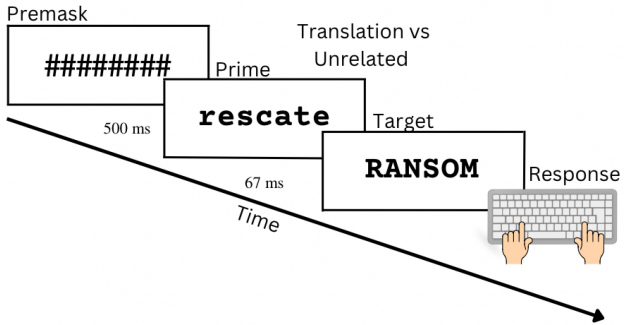
- **Chaouch-Orozco, González Alonso & Rothman (2021):** Translation priming asymmetries are explained by L2 use and not L2 proficiency. But…

- **Chaouch-Orozco, González Alonso, Duñabeitia & Rothman (2022):** No effects of L2 use when L2 proficiency is almost native-like.

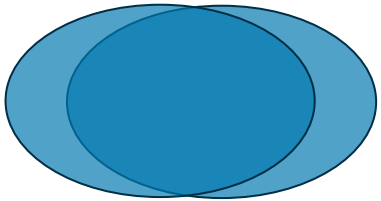


Introduction

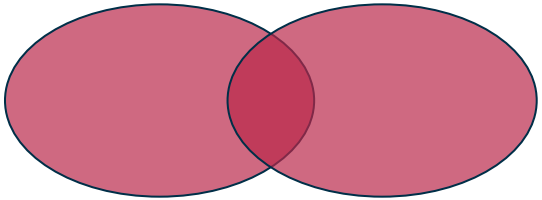
- **Chaouch-Orozco, González Alonso, Duñabeitia & Rothman (2023)**



Semantic overlap between **concrete** translation pairs

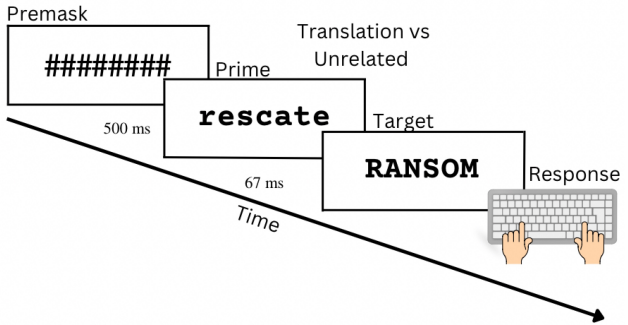


Semantic overlap between **abstract** translation pairs

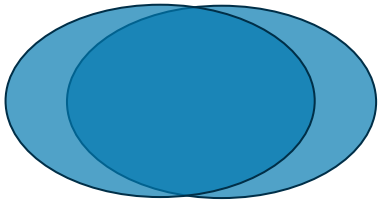


Introduction

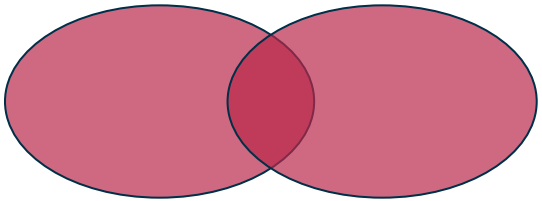
- **Chaouch-Orozco, González Alonso, Duñabeitia & Rothman (2023):** Translations are not equivalent: Semantic overlap between translations differ for concrete and abstract pairs.



Semantic overlap between **concrete** translation pairs



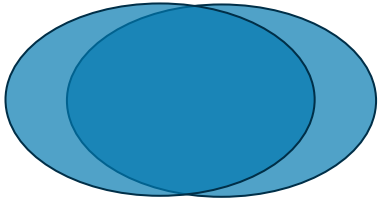
Semantic overlap between **abstract** translation pairs



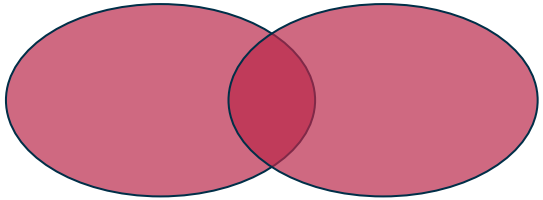
Introduction

1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?

Semantic overlap between **concrete** translation pairs



Semantic overlap between **abstract** translation pairs

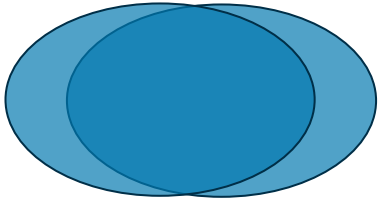


Introduction

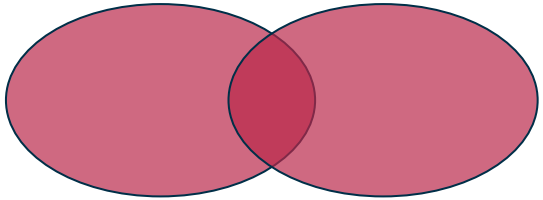
1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?

2. How does **culture** determine meaning?

Semantic overlap between **concrete** translation pairs



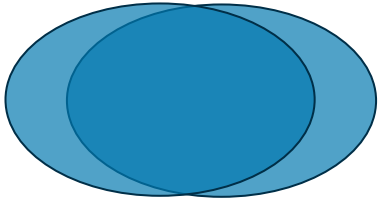
Semantic overlap between **abstract** translation pairs



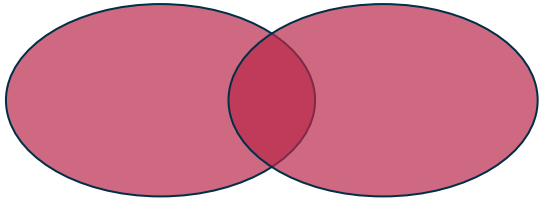
Introduction

1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?
2. How does **culture** determine meaning?
3. How does **(bilingual) semantic representation** impact processing?

Semantic overlap between **concrete** translation pairs



Semantic overlap between **abstract** translation pairs



1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?
2. How does **culture** determine meaning?
3. How does **(bilingual) semantic representation** impact processing?

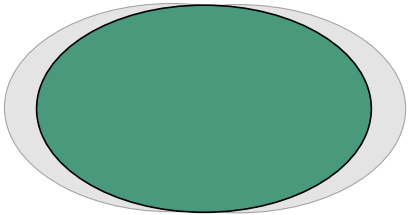
The bilingual lexicon

1. How do bilinguals represent **competing semantic representations** from their two languages? And how can we *precisely measure* them?
    - Bilingual semantic representation and processing is a *neglected field* of research (e.g., Šipka, 2015; Thompson et al., 2020).

The bilingual lexicon

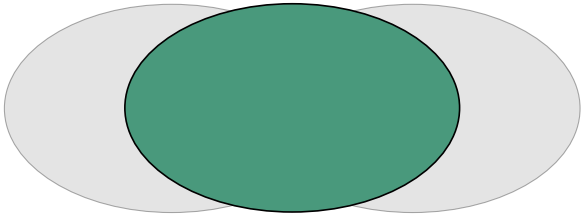
1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* them?
    - The traditional view assumes *one-to-one semantic mappings* (e.g., De Deyne et al., 2021; Dijkstra et al., 2019).

Semantic overlap between **concrete** translation pairs



dog -狗

Semantic overlap between **abstract** translation pairs

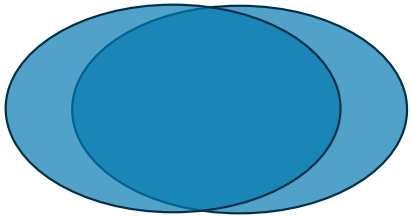


humility - 谦逊

The bilingual lexicon

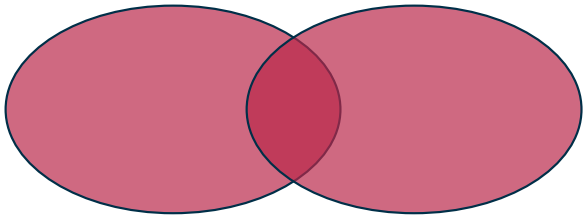
1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* them?
    - The traditional view assumes *one-to-one semantic mappings* (e.g., De Deyne et al., 2021; Dijkstra et al., 2019).

Semantic overlap between **concrete** translation pairs



dog -狗

Semantic overlap between **abstract** translation pairs



humility - 谦逊

The bilingual lexicon

1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* them?
    - The traditional view assumes *one-to-one semantic mappings* (e.g., De Deyne et al., 2021; Dijkstra et al., 2019).

Semantic overlap between **concrete** translation pairs

dog -狗

Semantic overlap between **abstract** translation pairs

humility - 谦逊

The bilingual lexicon

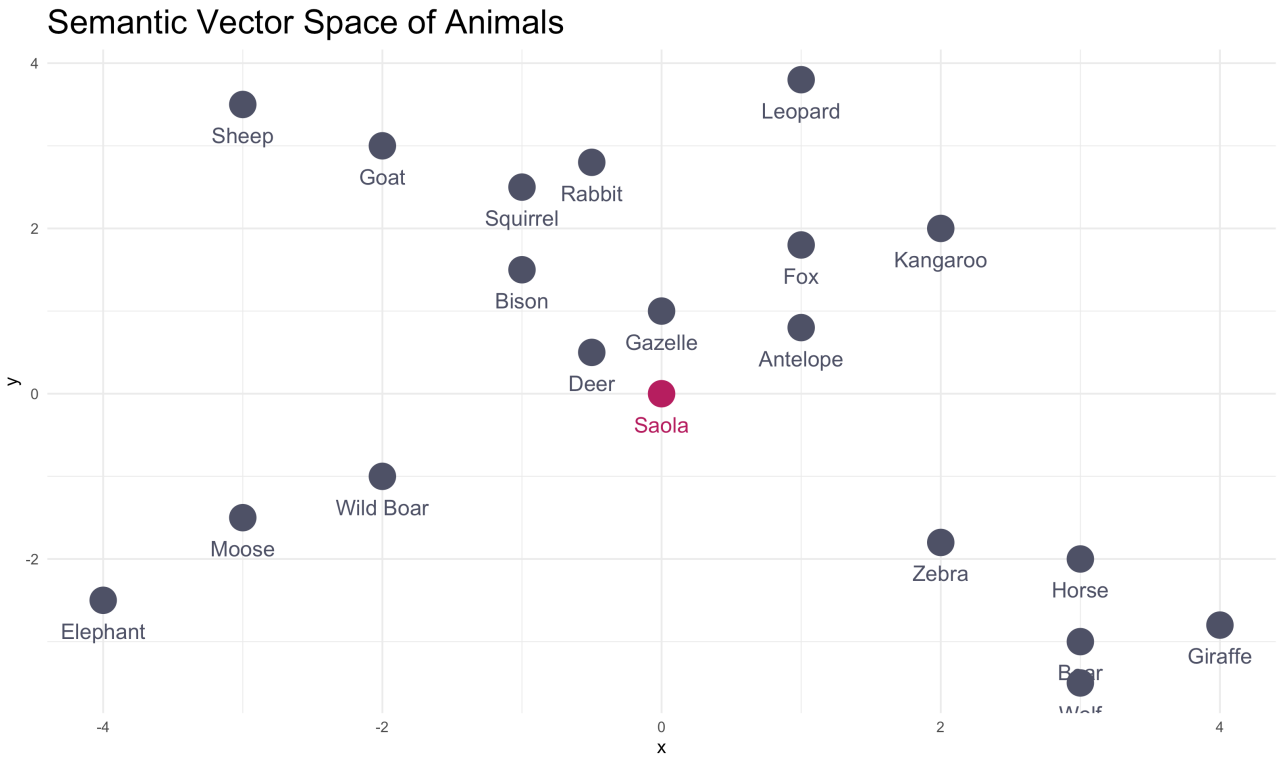
1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* them?
    - Measuring meaning is challenging:
        - Semantic feature norms (e.g., Lynott et al., 2022).
        - Neuroimaging techniques (e.g., Huth et al., 2016).
        - Distributional semantic models (e.g., Günther et al., 2019).
            - Based on the distributional hypothesis (Firth, 1957; Harris, 1954)

"You shall know a word by the company it keeps" (Firth, 1957)

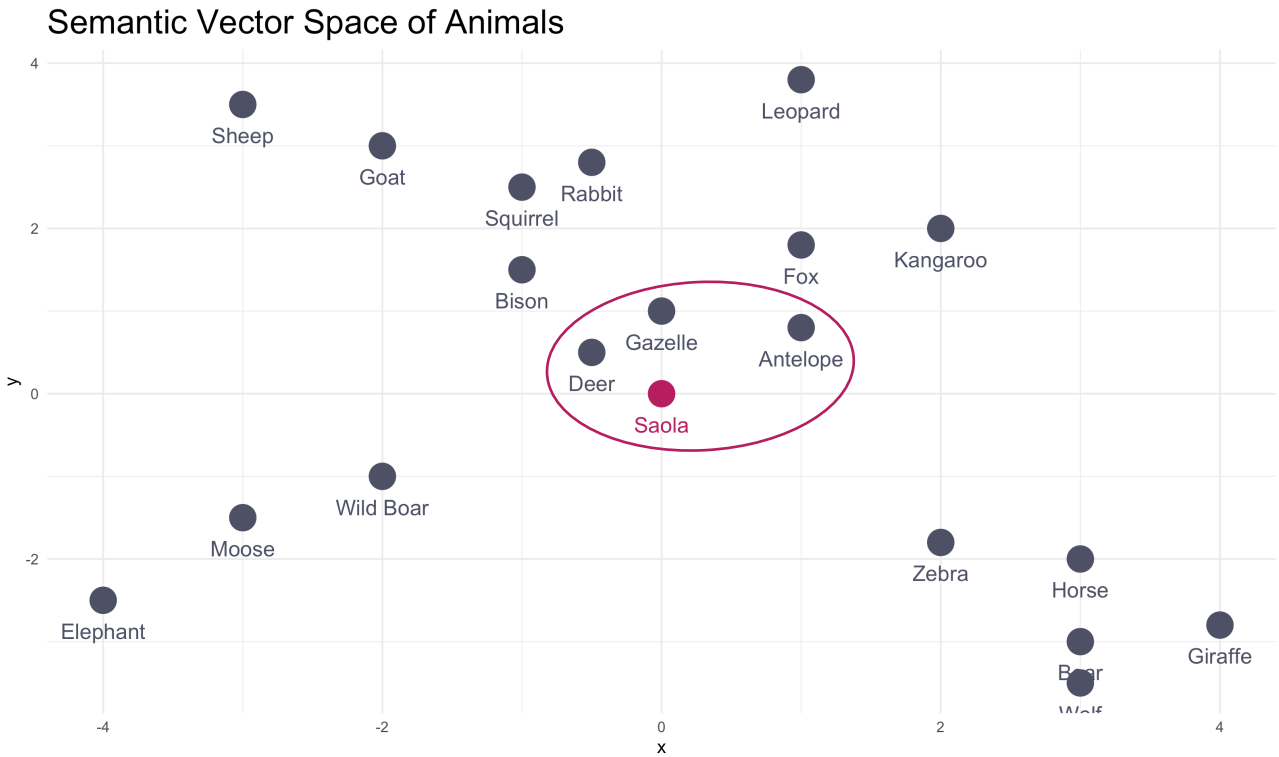
The bilingual lexicon

- What is a *saola*?

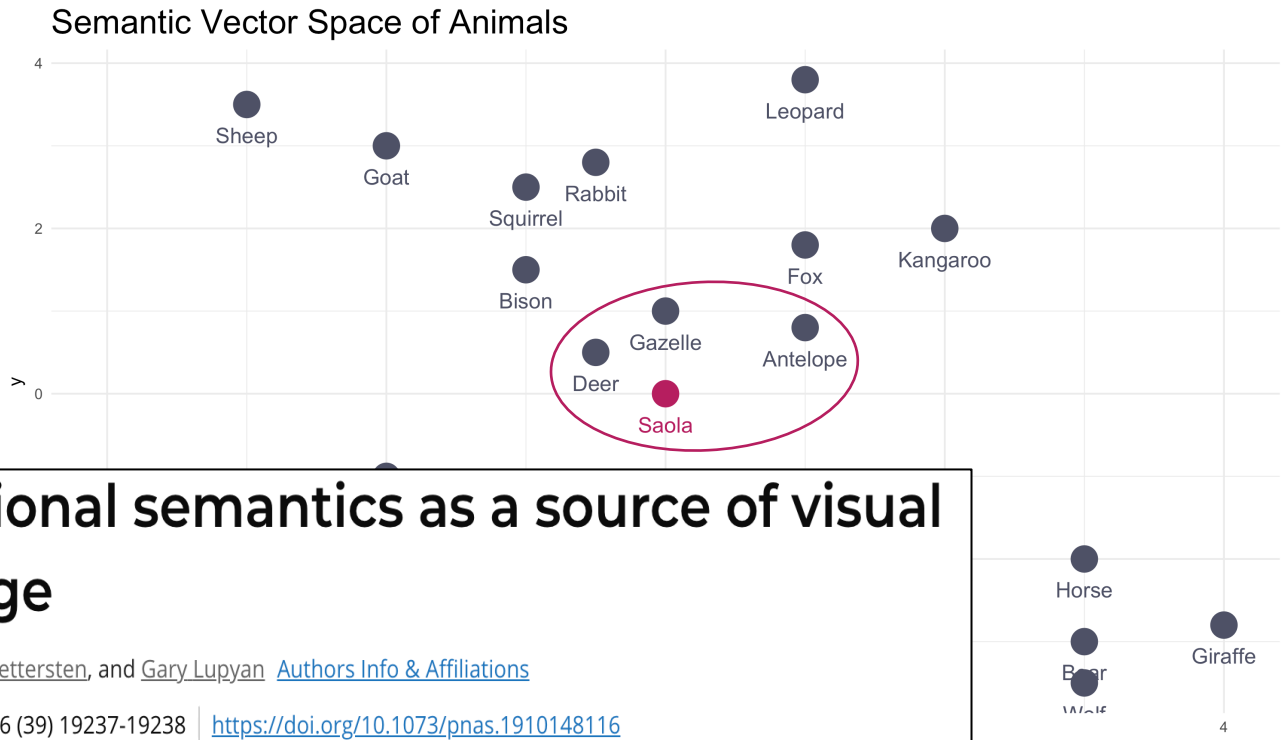
The bilingual lexicon



The bilingual lexicon



The bilingual lexicon

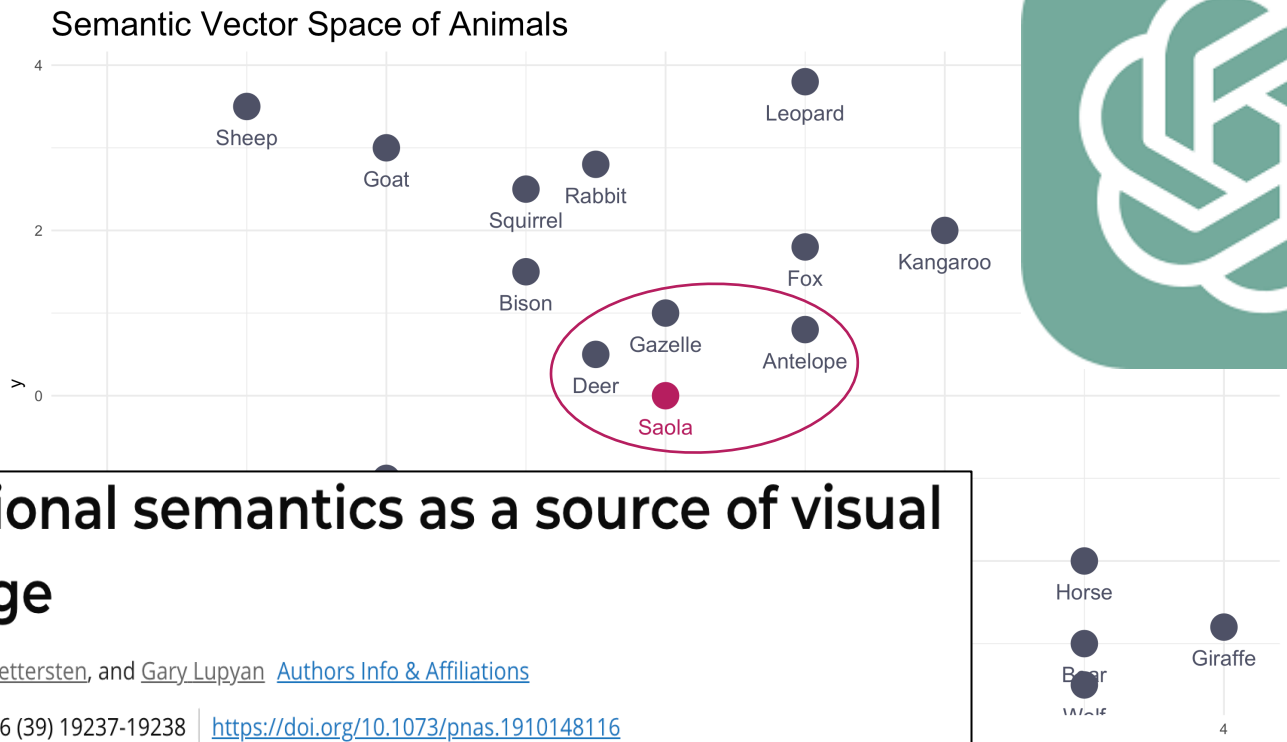


Distributional semantics as a source of visual knowledge

Molly Lewis, Martin Zettersten, and Gary Lupyan Authors Info & Affiliations

September 5, 2019 | 116 (39) 19237-19238 | https://doi.org/10.1073/pnas.1910148116

The bilingual lexicon

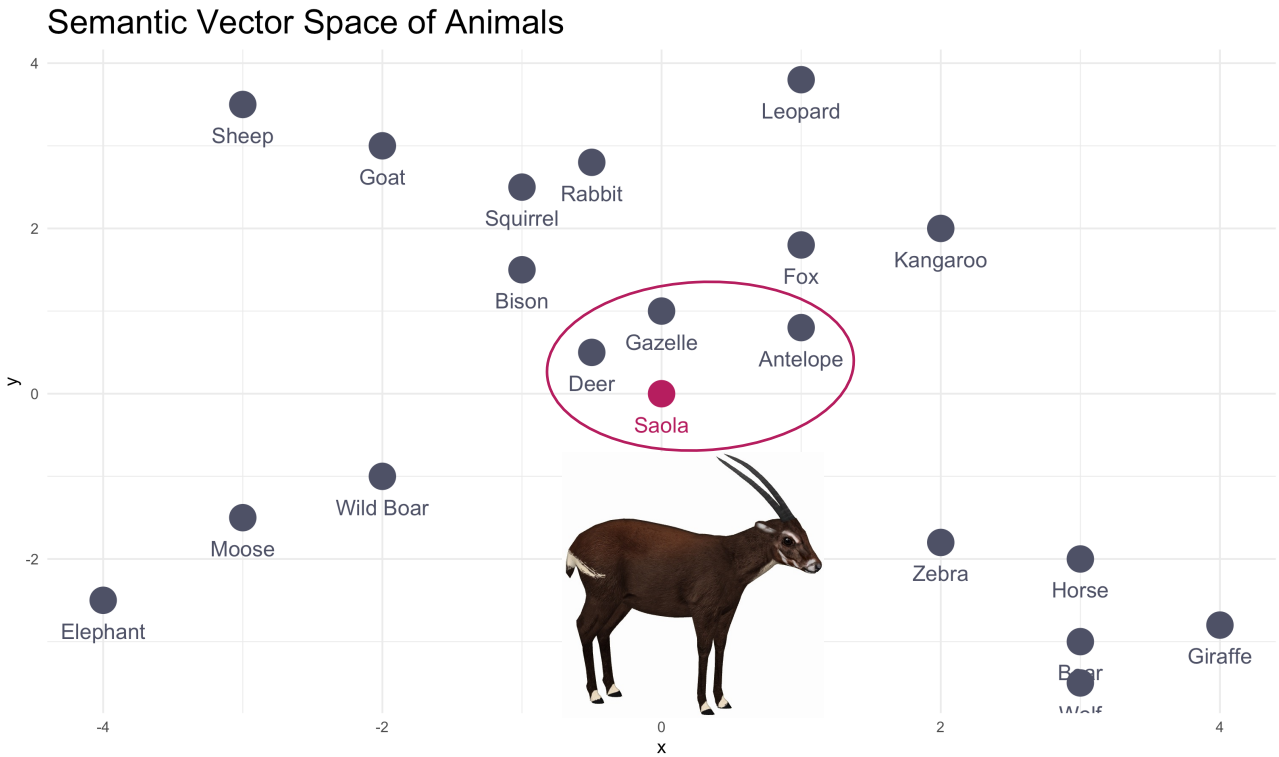


Distributional semantics as a source of visual knowledge

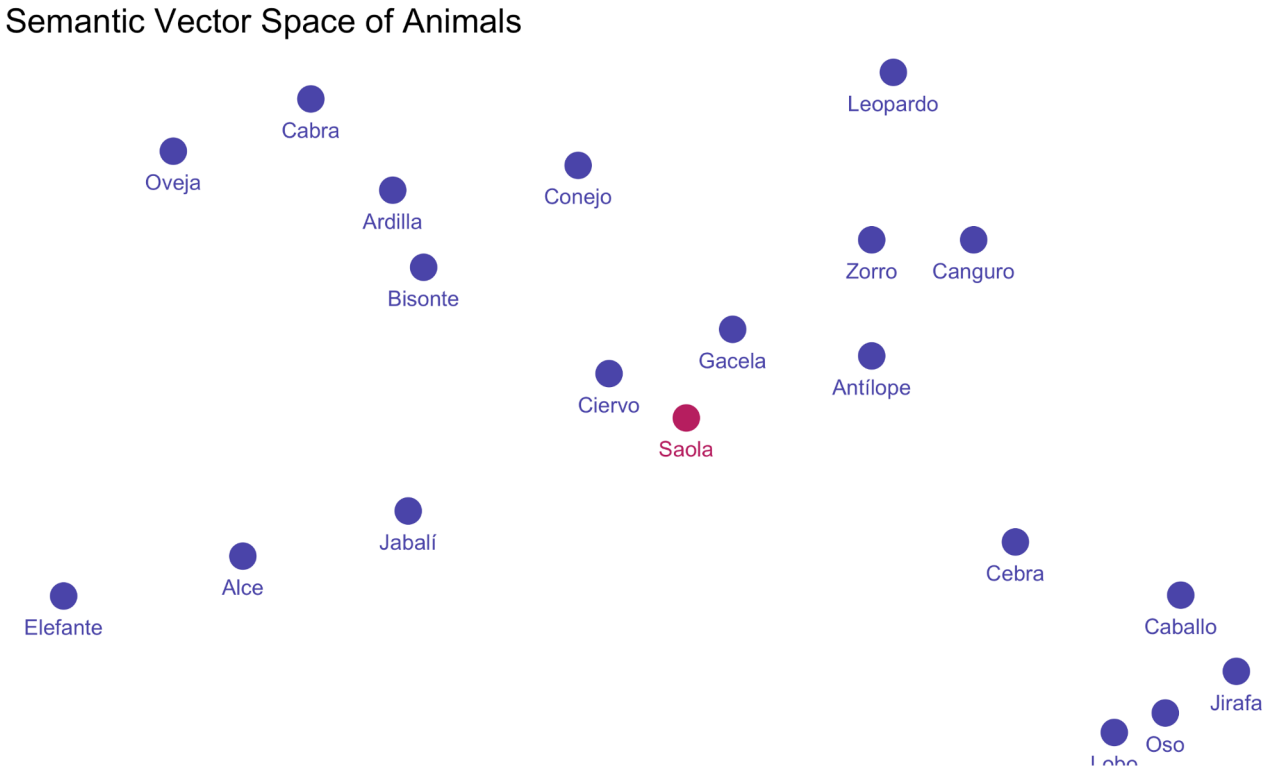
Molly Lewis, Martin Zettersten, and Gary Lupyan Authors Info & Affiliations

September 5, 2019 | 116 (39) 19237-19238 | https://doi.org/10.1073/pnas.1910148116

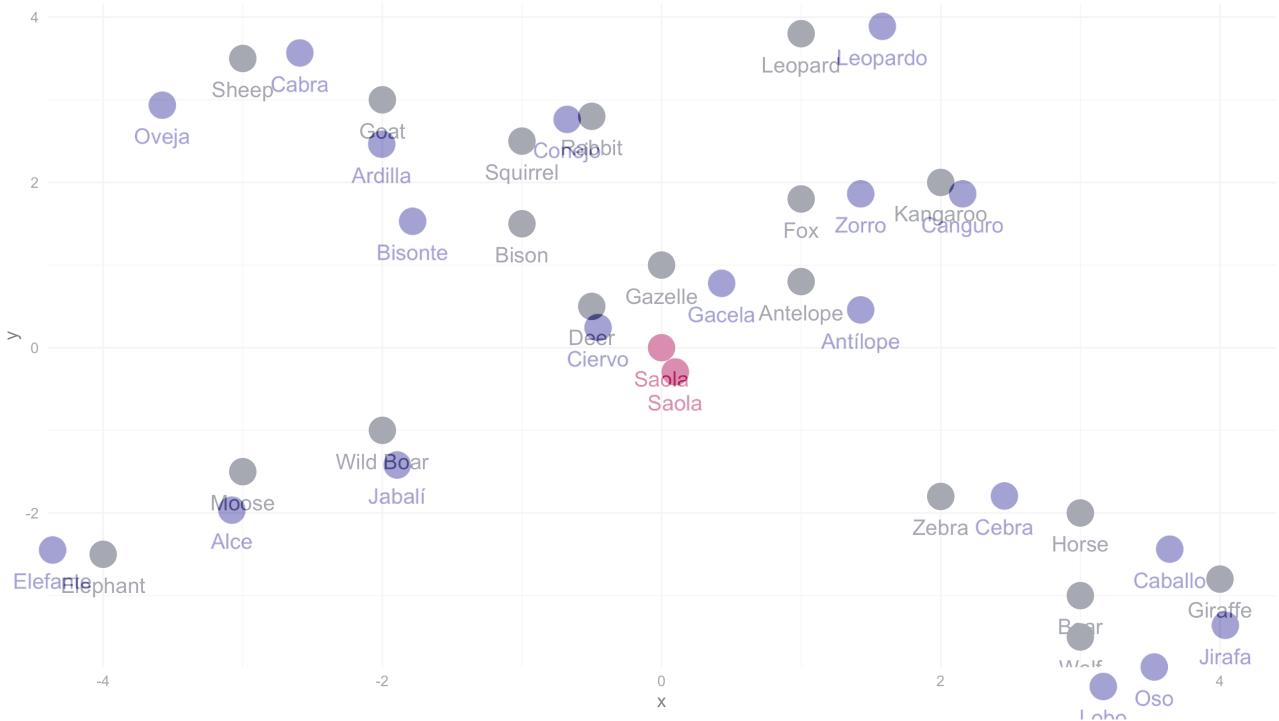
The bilingual lexicon



The bilingual lexicon



The bilingual lexicon



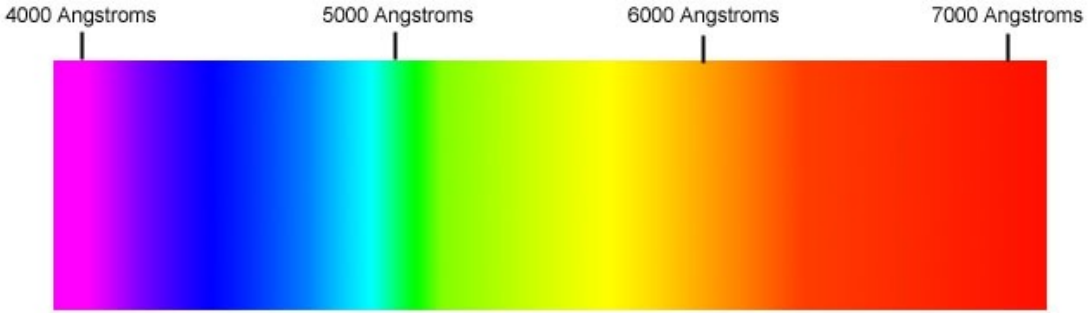
1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?

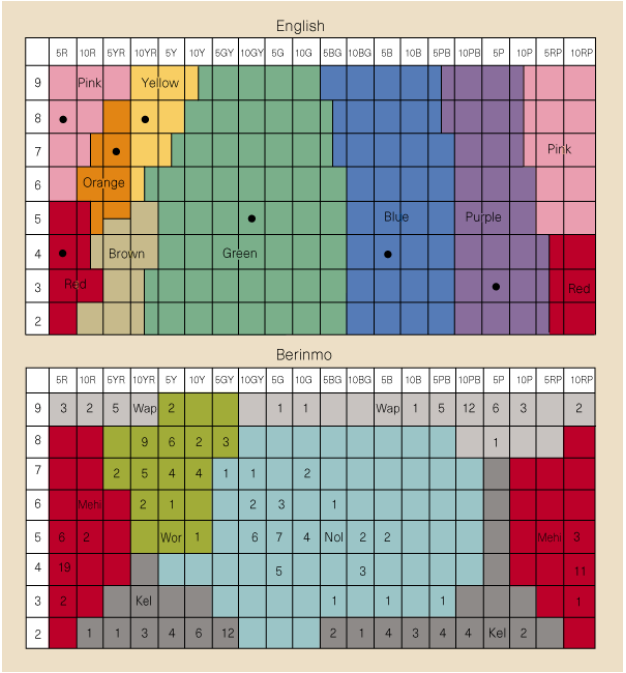
2. How does **culture** determine meaning?

3. How does **(bilingual) semantic representation** impact processing?

Cultural lexical semantics

2. How does **culture** determine meaning?





Davidoff et al. (1999)

Cultural lexical semantics

2. How does **culture** determine meaning?

ELSEVIER

Cognition

Volume 108, Issue 3, September 2008, Pages 819-824

COGNITION

Brief article

Number as a cognitive technology: Evidence from Pirahã language and cognition ☆

Michael C. Frank [a], Daniel L. Everett [b], Evelina Fedorenko [a], Edward Gibson [a]

Cultural lexical semantics

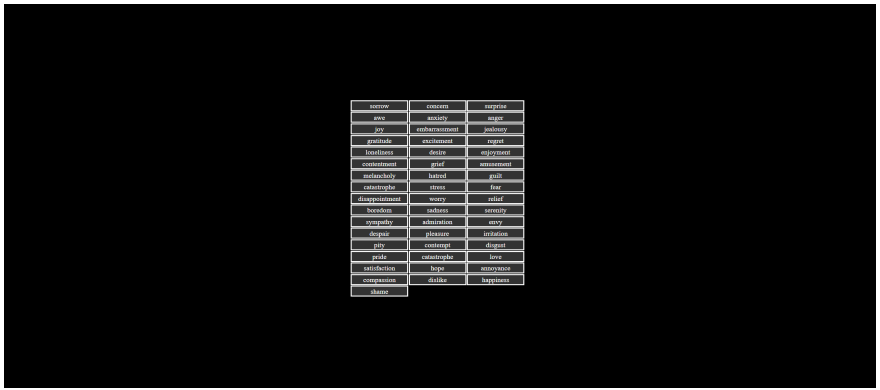
2. How does **culture** determine meaning?

Cultural lexical semantics

- **Chaouch-Orozco, Li & Li (in preparation):**
  - Research question: Which specific cultural dimensions are associated with how emotion semantic spaces vary across languages?

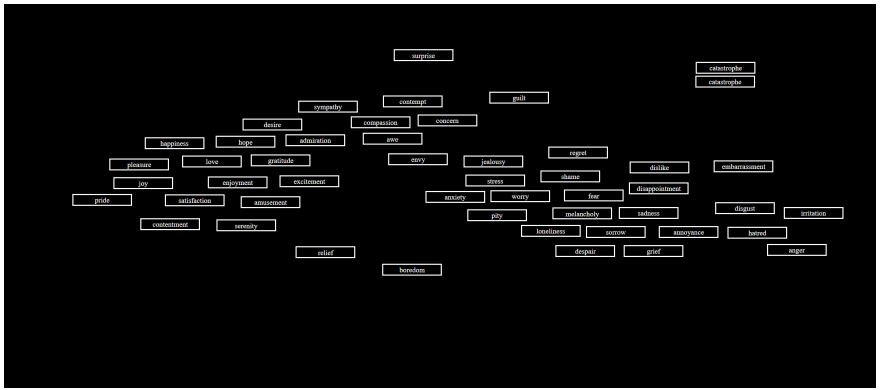
Cultural lexical semantics

- **Chaouch-Orozco, Li & Li (in preparation):**
  - Method:
    - Hofstede's cultural dimensions (Hofstede, 2001): power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence.
    - 50 native speakers of 15 languages.
    - Q-SpAM (Koch et al., 2021) with 47 emotion words.



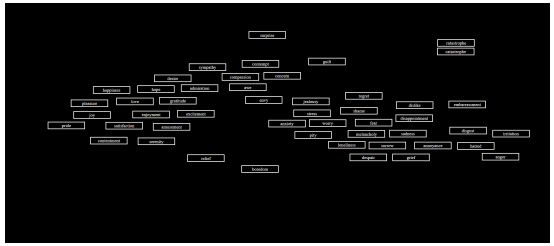
Cultural lexical semantics

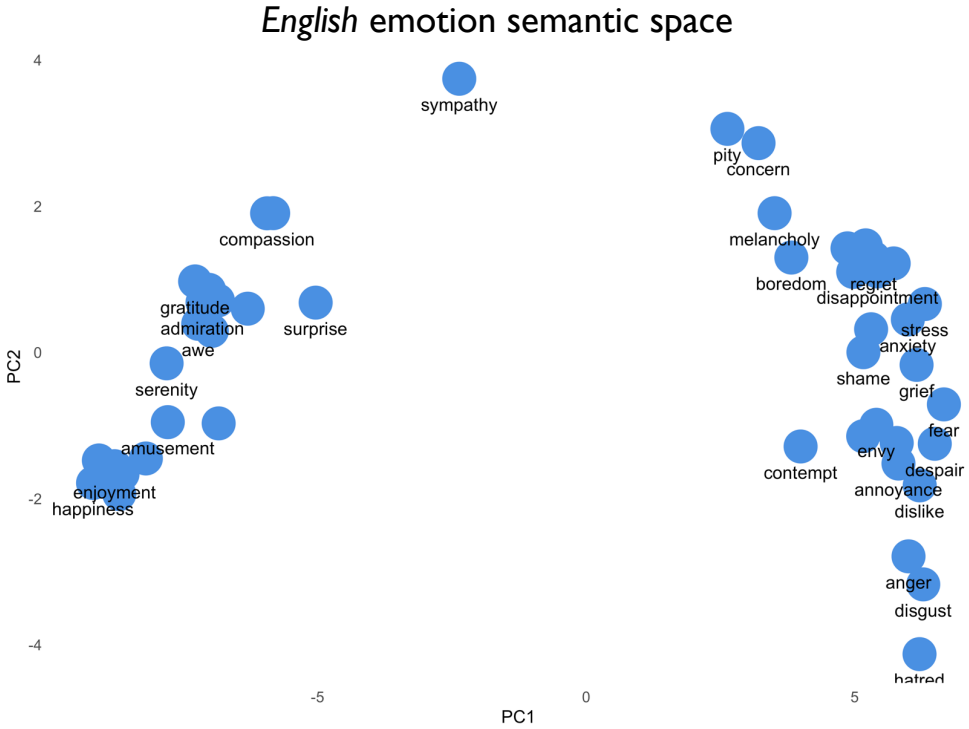
- **Chaouch-Orozco, Li & Li (in preparation):**
  - Method:
    - Hofstede's cultural dimensions (Hofstede, 2001): power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence.
    - 50 native speakers of 15 languages.
    - Q-SpAM (Koch et al., 2021) with 47 emotion words.



Cultural lexical semantics

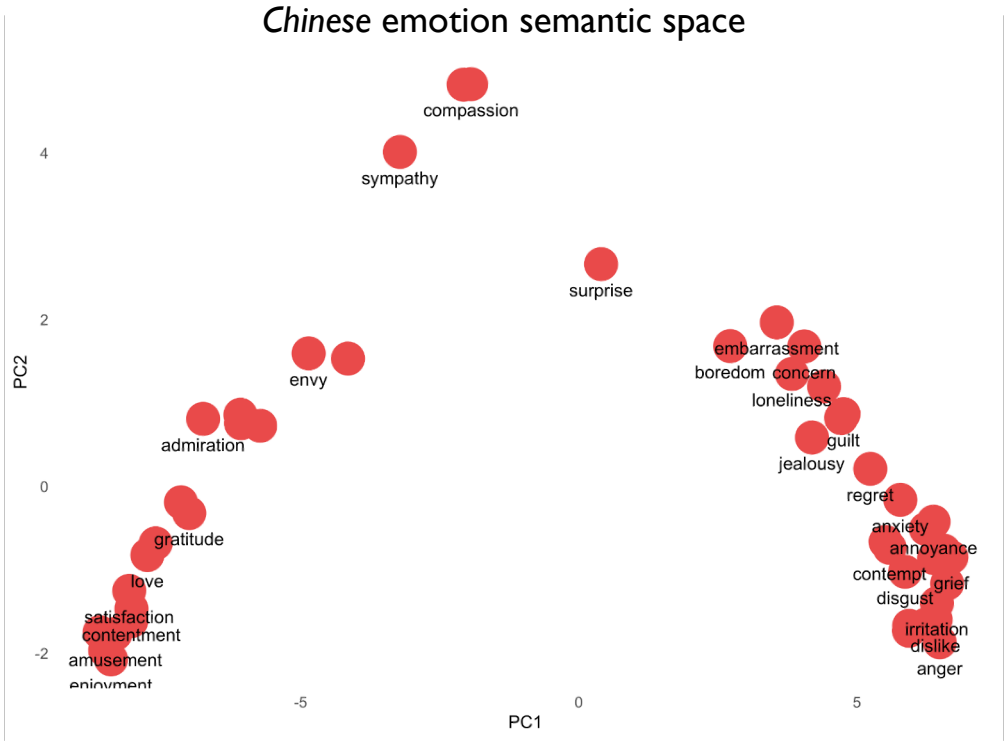
- **Chaouch-Orozco, Li & Li (in preparation):**





Cultural lexical semantics

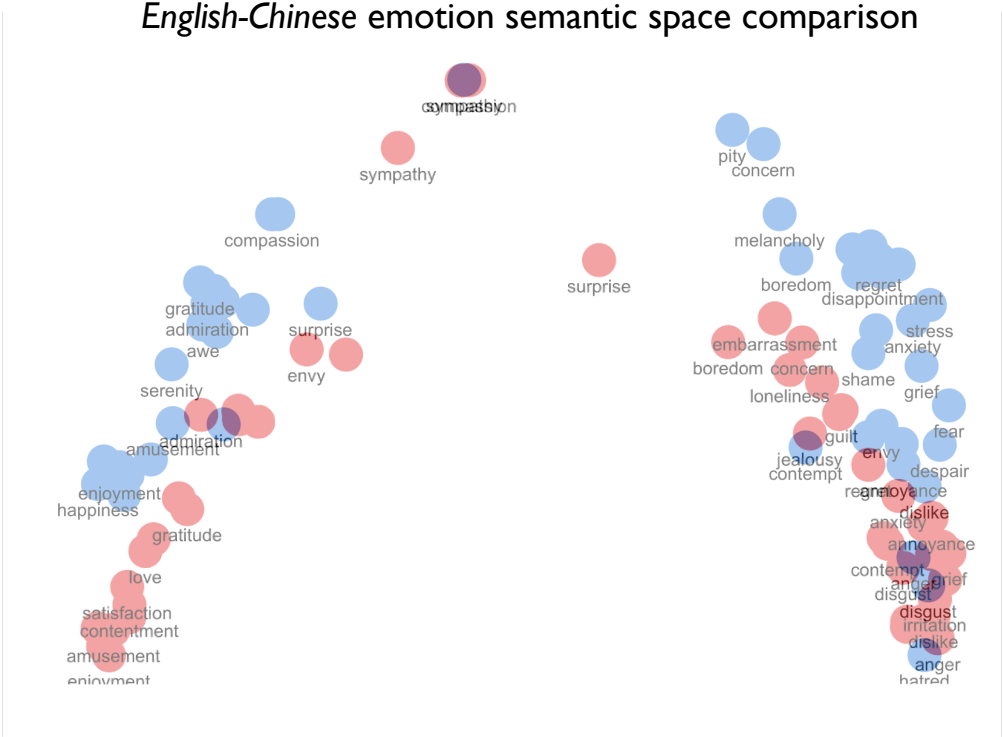
- **Chaouch-Orozco, Li & Li (in preparation):**



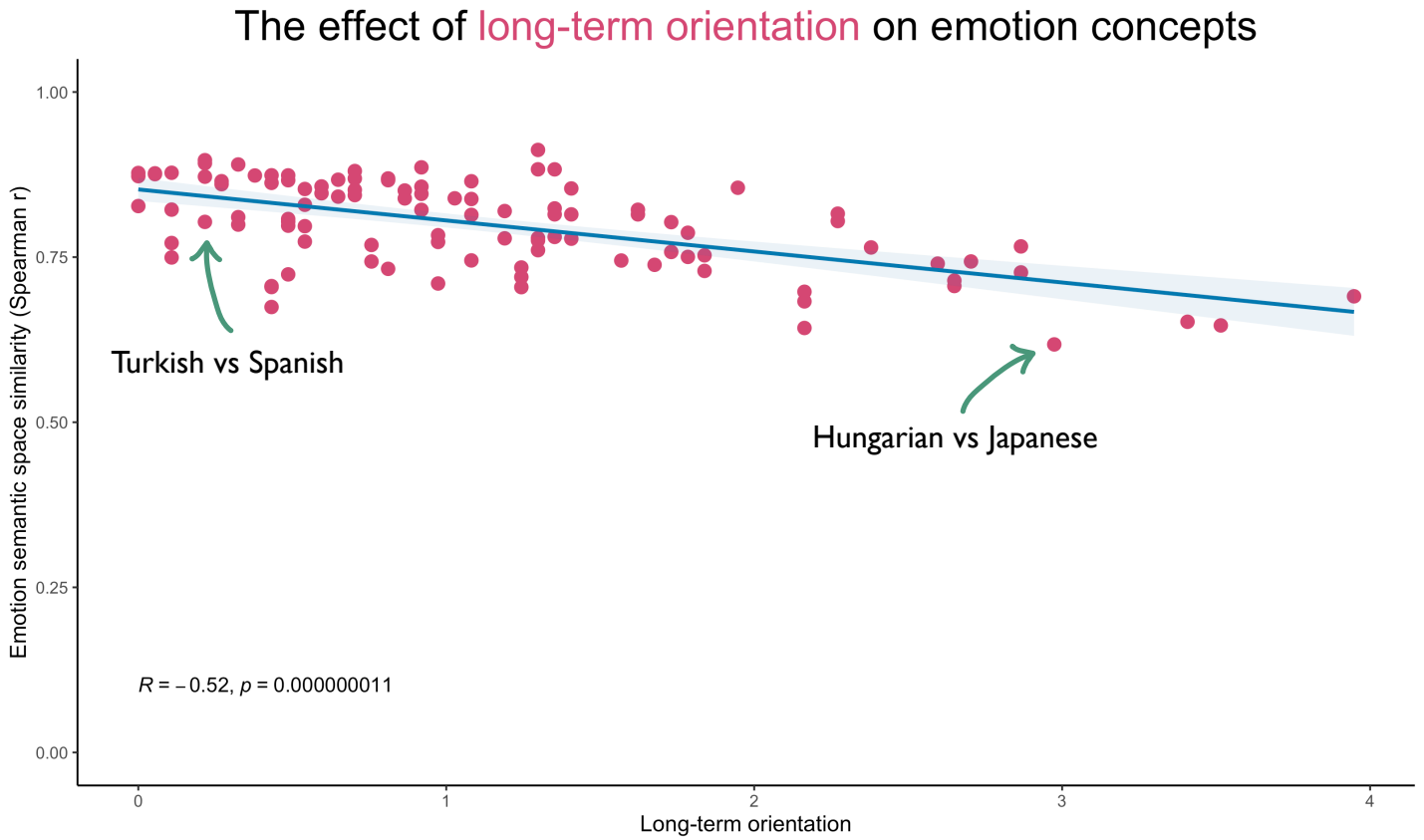
Cultural lexical semantics

- **Chaouch-Orozco, Li & Li (in preparation):**

English-Chinese emotion semantic space comparison

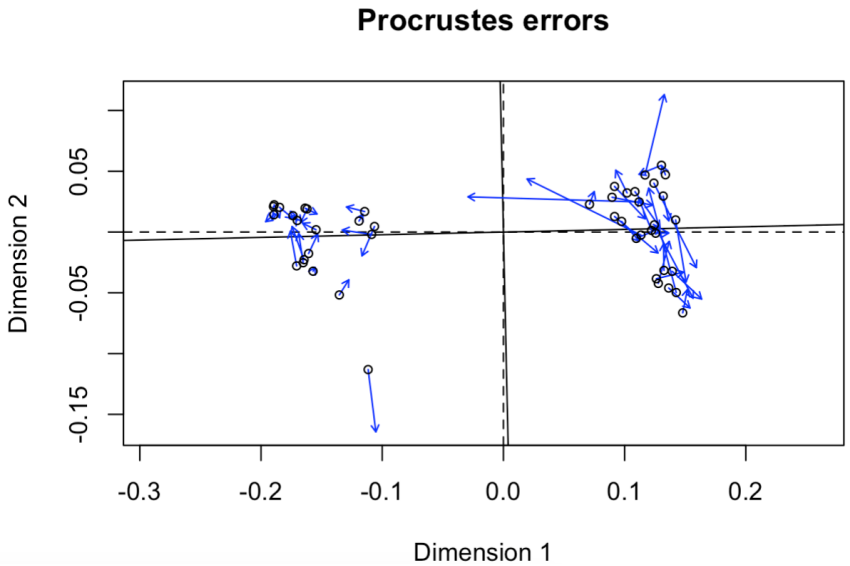


Cultural lexical semantics

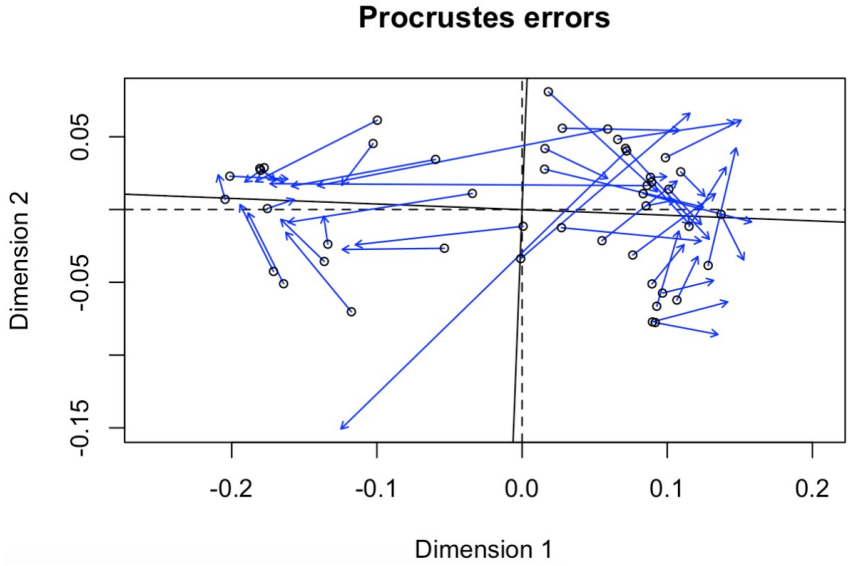


Cultural lexical semantics

Two short-term oriented cultures
(Turkish vs Spanish)

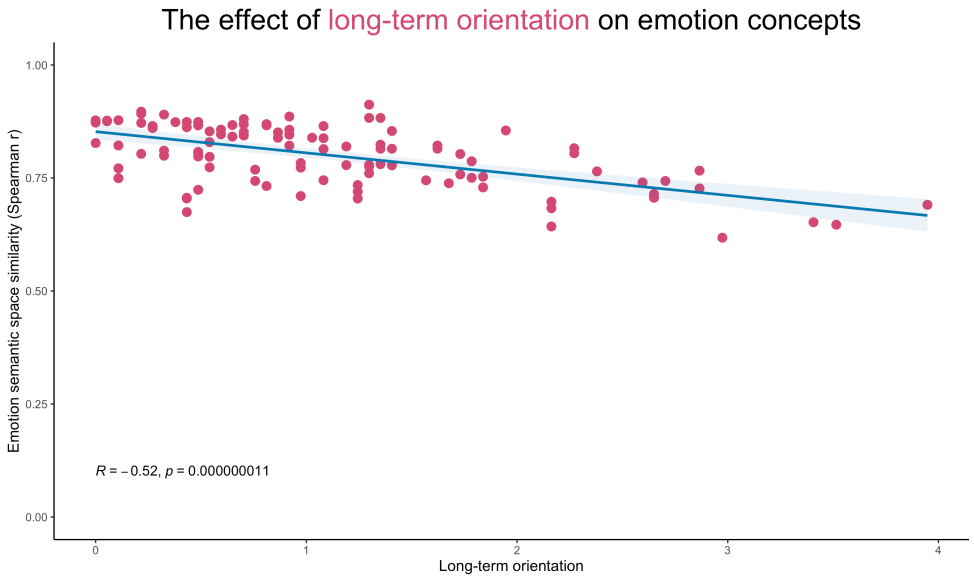


Long- vs short-term oriented cultures
(Hungarian vs Japanese)



Cultural lexical semantics

- **Chaouch-Orozco, Li & Li (in preparation):**



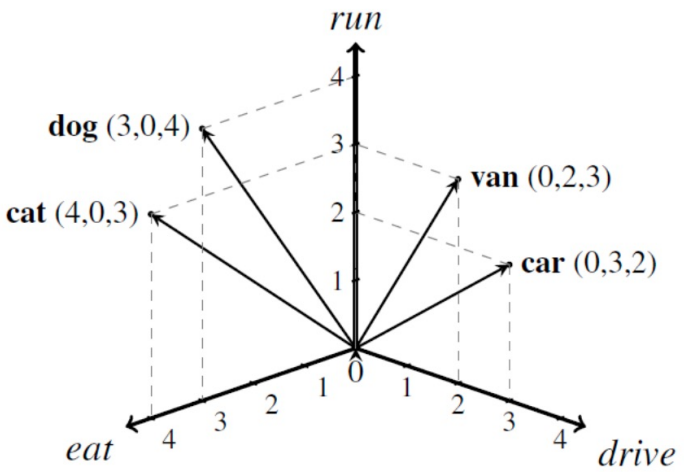
- Significant effect of culture (long-term orientation).
- Larger effect for *negative emotion words*.
- The effect remains when controlling for language family, script, and religion.

Cultural lexical semantics

- **Follow-up:** How do Q-SpAM-based semantic spaces correlate with those obtained from word embeddings?

Cultural lexical semantics

- **Follow-up:** How do Q-SpAM-based semantic spaces correlate with those obtained from word embeddings?
    - Word embeddings build vector representations from *text corpora*.



Cultural lexical semantics

- **Follow-up:** How does culture influence emotion semantic representation in **bilinguals**?
  - **Tsoi & Chaouch-Orozco (in preparation):**
    - Emotion categorization (Q-SpAM).
    - Native Cantonese speakers (Hong Kong), native Japanese speakers (Japan), Cantonese-Japanese late sequential bilinguals (immersed in Japan for at least two years).
  - **Chaouch-Orozco & Chattopadhyay (in preparation):**
    - Emotion categorization (Q-SpAM).
    - Native Nepalese speakers (Nepal), native Cantonese speakers (Hong Kong), Nepalese-Cantonese heritage bilinguals.

1. How do bilinguals represent **competing semantic information** from their two languages? And how can we *precisely measure* that?

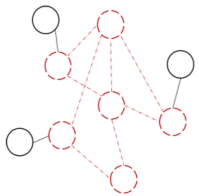
2. How does **culture** determine meaning?

3. How does **(bilingual) semantic representation** impact processing?

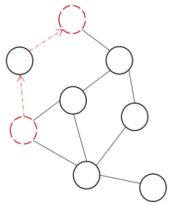
Language processing and cognition

1. Chaouch-Orozco & Martín-Villena (2024):

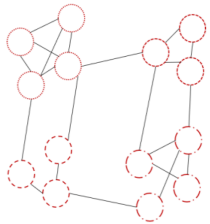
- *Research question:* Does L2 immersion erode the L1 semantic network's organization?
- *Method:*
  - 94 immersed and 80 non-immersed Spanish-English late sequential bilinguals.
  - Two semantic fluency tasks: fruits and vegetables (L1), animals (L2) → Correlation networks (Kenett et al., 2013).



Cluster coefficient (CC)



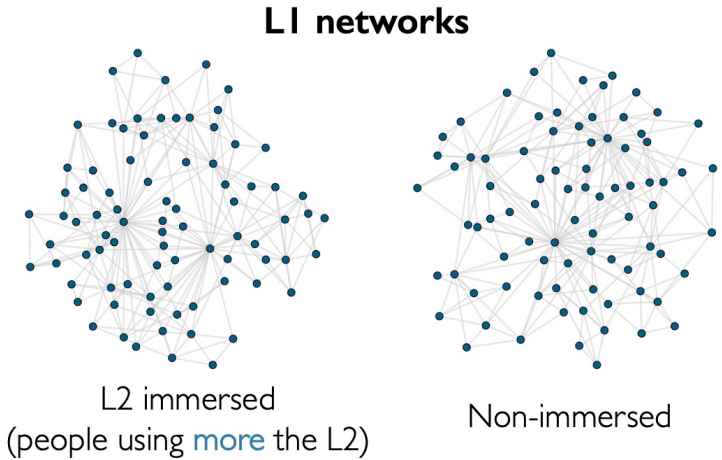
Shortest Path Length (SPL)



Modularity (Q)

Language processing and cognition

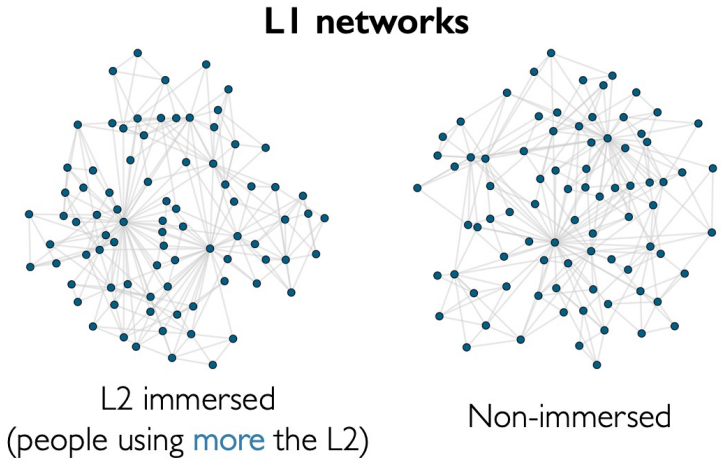
1. **Chaouch-Orozco & Martín-Villena (2024):**
   - L2 immersion *impacts* the structural organization of the L1 semantic network.



Language processing and cognition

1. Chaouch-Orozco & Martín-Villena (2024):

- L2 immersion *impacts* the structural organization of the L1 semantic network.



The LeAF framework

Hierarchy of changes	Processing effects
Clustering coefficient (reduced connectedness)	Hesitations, pauses, circumlocutions
Modularity (reduced integration)	Lexical substitutions, semantic inaccuracy, L2 borrowings
Average shortest-path length (reduced navigability)	Retrieval slowing down, repetitions, language switches

Language processing and cognition

2. Liu & Chaouch-Orozco (2023)
3. Chaouch-Orozco & Liu (in preparation)

Language processing and cognition

- Our languages are activated simultaneously (e.g., Kroll et al., 2006; Thierry & Wu, 2007; van Assche et al., 2012).

Language processing and cognition

- Our languages are activated simultaneously (e.g., Kroll et al., 2006; Thierry & Wu, 2007; van Assche et al., 2012).
- How do we produce the target language so efficiently?

Language processing and cognition

- How do we produce the target language so efficiently?

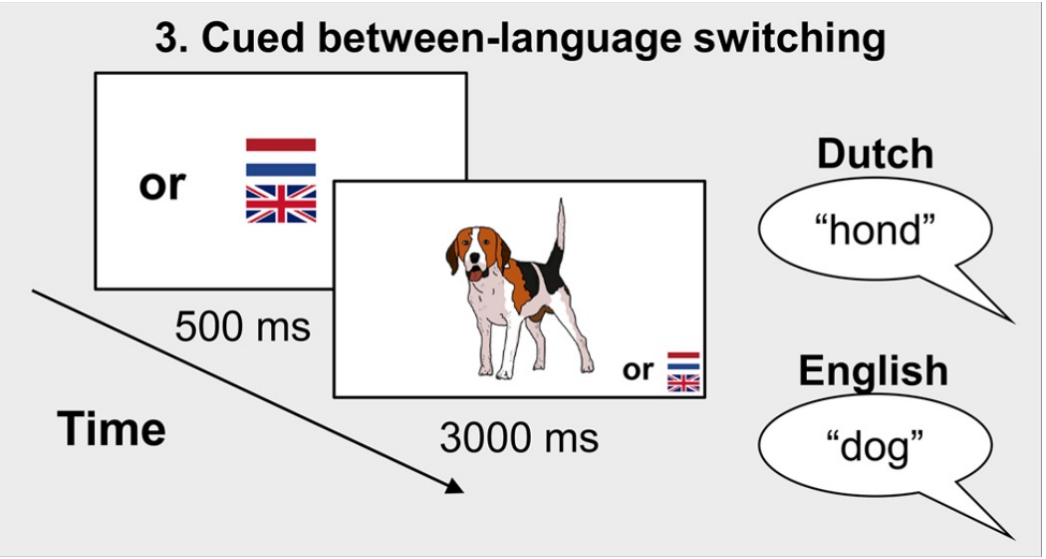


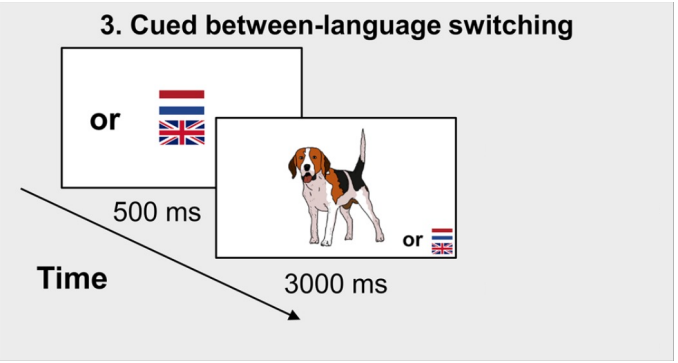
Image from Moojiman et al. (2023)

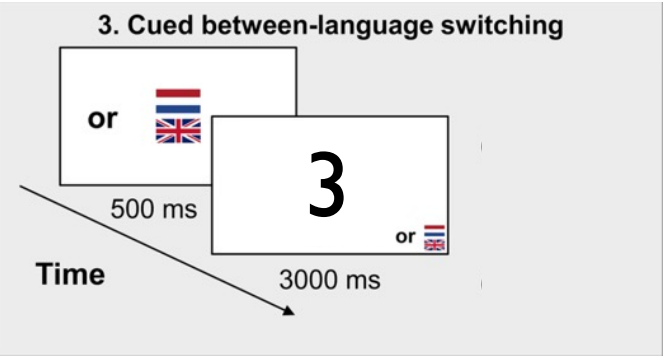
Repeat trials
vs
Switch trials

Switching costs
(faster RT in repeat trials)

Language processing and cognition

- *Pictures* and *digits* are used indistinctively, but are they comparable?
  - *Declerck et al. (2012):*
    - Digit effect (i.e., larger switching costs with pictures than digits).
    - Caused by phonological overlap.





Language processing and cognition

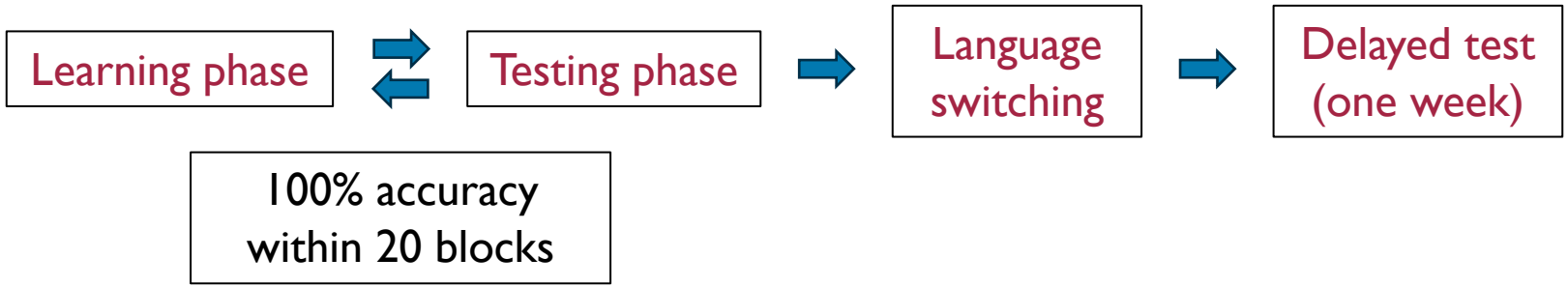
2. **Liu & Chaouch-Orozco (2023):**
    - Chinese-English-French trilinguals.
    - We observed an *inverse digit effect:* larger switching costs for digits.
        - The *inverse* digit effect is not explained by *phonological overlap* nor by *semantics* (similar magnitudes; i.e., numerical distance effect).
        - But maybe by **associative relationships** (Macizo and Alvarez, 2018)?

Language processing and cognition

3. **Chaouch-Orozco & Liu (in preparation):**
    - Research question: How do *semantic* and *associative relationships* influence language control when naming pictures and digits?
    - Hypotheses:
        - **Associative connections** result in *larger switching costs*.
            - Increased activation of within-language associates makes switching more effortful, particularly for digits (Macizo & Alvarez, 2018; Liu & Chaouch-Orozco, 2023).
        - **Semantic connections** result in *smaller switching costs*.
            - Cross-language semantic priming activates related concepts across languages, facilitating switching (Shen & Chen, 2023).

Language processing and cognition

3. **Chaouch-Orozco & Liu (in preparation):**
    - Method:
        - 240 Chinese native speakers are taught novel L2 Turkish words labelling objects and numbers.



Language processing and cognition

3. Chaouch-Orozco & Liu (in preparation):

Unrelated objects

Digits – sequence (+ *association*; - *semantics*)

1, 2, 3, 4, 5, 6, 7, 8, 9

Digits – random order (- *association*; - *semantic*)

3, 9, 2, 5, 6, 8, 4, 1, 7

Magnitudes – sequence (+ *association*; + *semantics*)

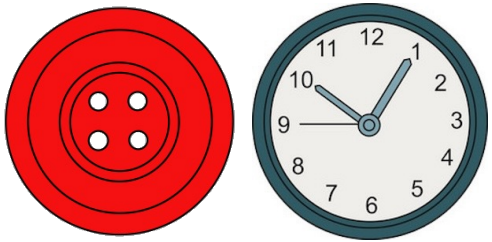
Magnitudes – random order (- *association*; + *semantics*)

Language processing and cognition

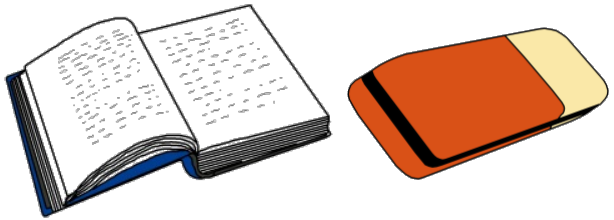
3. **Chaouch-Orozco & Liu (in preparation):**

Unrelated objects

Semantically similar objects



Associated objects



Language processing and cognition

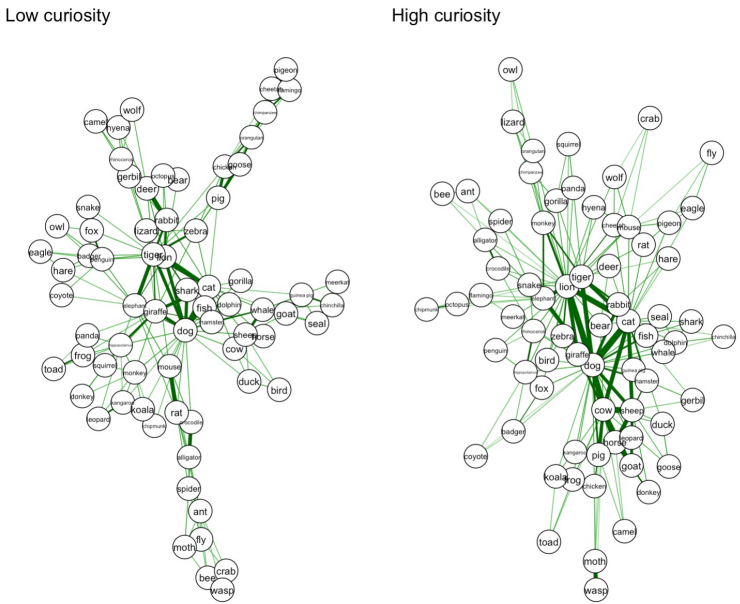
3. **Chaouch-Orozco & Liu (in preparation):**
    - *Preliminary findings:*
        - **Associative connections** result in *larger switching costs*.
        - **Semantic connections** result in *smaller switching costs*.
    - We have to wait for the rest of the picture data!

Language processing and cognition

4. **Chaouch-Orozco & Li (in preparation):**
    - Research question: Do *more curious people* exhibit more efficiently organized semantic networks and better verbal analogical reasoning skills?

Language processing and cognition

4. Chaouch-Orozco & Li (in preparation):



- *Recombinant knowledge search* (Schilling & Green, 2011).
- **More efficient organization:** Faster access to semantic information, and a more integrated network.

Language processing and cognition

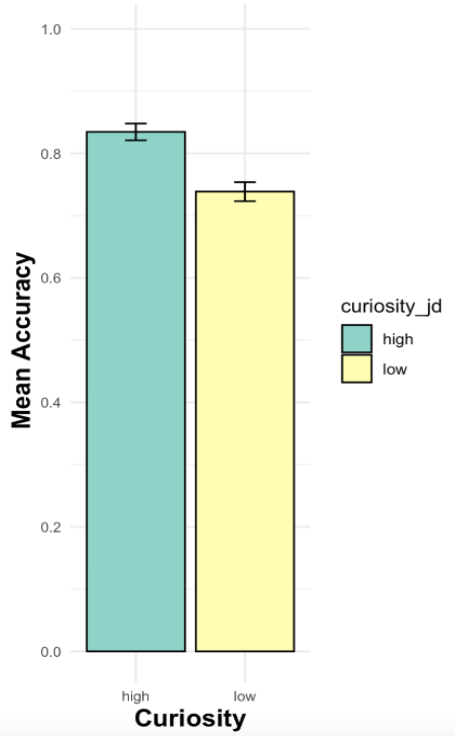
4. **Chaouch-Orozco & Li (in preparation):**
    - *Analogical reasoning task*:
        - wire : copper :: knife : **steel** (*valid* analogy)

Language processing and cognition

4. **Chaouch-Orozco & Li (in preparation):**
    - *Analogical reasoning task*:
        - wire : copper :: knife : **steel** (*valid* analogy)
        - wire : copper :: knife : **water** (*incorrect* analogy)

Language processing and cognition

4. **Chaouch-Orozco & Li (in preparation):**
   - *Analogical reasoning task*:
     - wire : copper :: knife : **steel** (*valid* analogy)
     - wire : copper :: knife : **water** (*incorrect* analogy)
   - Results:
     - *More curious people* exhibit better organized semantic networks and enhanced verbal analogical reasoning skills.
     - But no differences in *vocabulary size* and *attention*.



Language processing and cognition

4. **Chaouch-Orozco & Li (in preparation):**
    - Next steps:
        - Establishing causal relationships (priming exploration/curiosity).
        - Examining individual networks.
        - Incorporating ERPs.
        - **What about verbal analogical reasoning in *bilinguals' L1*?**

Conclusions

Conclusions

- *Translations are not equivalent.* Computational models of the multilingual lexicon should incorporate distributed semantic representations (Chaouch-Orozco et al., 2023).
- *Culture* determines how we categorize reality *in very specific ways* (Chaouch-Orozco et al., in preparation).
- Semantic relationships in the (bilingual) lexicon have effects on *language processing:*
  - Lexical attrition (Chaouch-Orozco et al., 2024).
  - Language switching and control (Chaouch-Orozco et al., in preparation).
  - Curiosity → Semantic networks → Verbal analogical reasoning.

Conclusions

- *Translations are not equivalent*. Computational models of the multilingual lexicon should incorporate distributed semantic representations (Chaouch-Orozco et al., 2023).
- *Culture* determines how we categorize reality *in very specific ways* (Chaouch-Orozco et al., in preparation).
- Semantic relationships in the (bilingual) lexicon have effects on *language processing:*
  - Lexical attrition (Chaouch-Orozco et al., 2024).
  - Language switching and control (Chaouch-Orozco et al., in preparation).
  - Curiosity → Semantic networks → Verbal analogical reasoning.

Conclusions

- *Translations are not equivalent.* Computational models of the multilingual lexicon should incorporate distributed semantic representations (Chaouch-Orozco et al., 2023).

- *Culture* determines how we categorize reality *in very specific ways* (Chaouch-Orozco et al., in preparation).

- Semantic relationships in the (bilingual) lexicon have effects on *language processing:*

  - Lexical attrition (Chaouch-Orozco et al., 2024).

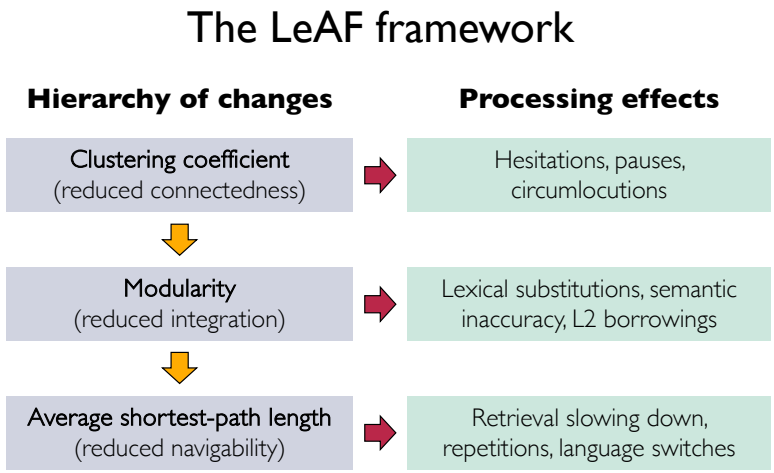
  - Language switching and control (Chaouch-Orozco et al., in preparation).

  - Curiosity → Semantic networks → Verbal analogical reasoning.

Future directions

Future directions

- Improving the **LeAF framework.**
  - Representation and *processing*.
  - From lexical attrition to *dynamics within the bilingual lexicon:* all sort of bilingual populations in Hong Kong.



Future directions

- Exploring the "Whorfian turn": Does emotion semantic representation impact emotional processing?
    - Some evidence (Gendron et al., 2012, 2013; Lindquist et al., 2006).
    - Autistic children (**Zhang & Chaouch-Orozco, in preparation**).
    - Mood disorders (**The Hong Kong Emotion Map**).
        - Semantic and associative relationships.
        - Emotional granularity.
        - Thought processes and rumination.

Future directions

- Investigating negative emotion words' **semantic evolution in the lab.**
  - *Potential factors:*
    - Cultural dimensions (long-term orientation).
    - Allostatic dysregulation (response to stress).
    - The "range effect" (Alves et al., 2017).
  - *Method:* Serial reproduction task.

Thank you!
Questions?

Special thanks to everybody involved in these studies:

Jason Rothman

Ping Li

Jorge González Alonso

Jon Andoni Duñabeitia

Hong Li

Eloi Puig-Mayenco

Fernando Martín-Villena

Xiyuan Li

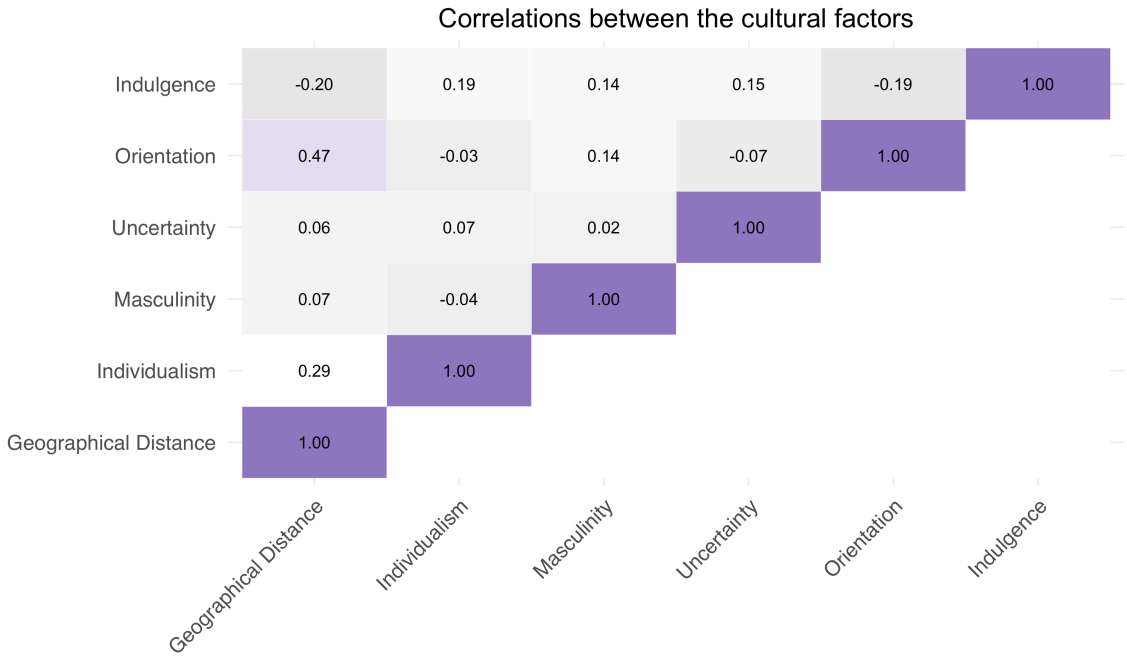
Yixin Zhang

Ryan Tsoi

Pallabi Chattopadhyay

Future directions

- Why *long-term orientation?*



Future directions

- Why *long-term orientation?*
  - Originally labelled as *Confucian Work Dynamism* (Chinese Culture Connection, 1987).
  - Related to **ethical values** (Nevins et al., 2007).
    - Future planning
    - Tradition
    - Service to others
    - Status and shame

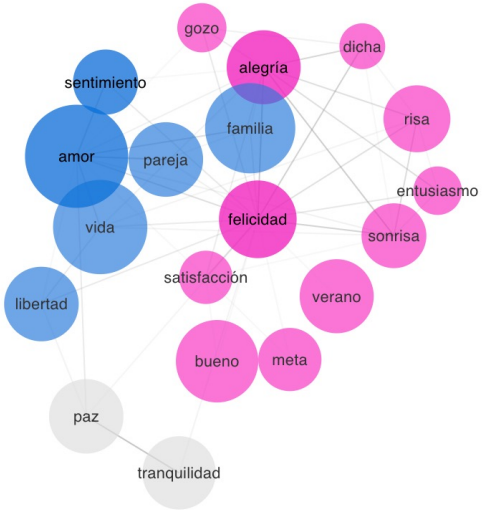
Future directions

- Why *long-term orientation?*
  - Originally labelled as *Confucian Work Dynamism* (Chinese Culture Connection, 1987).
  - Related to **ethical values** (Nevins et al., 2007).
    - Future planning → anxiety, hope, fear, worry
    - Tradition → boredom
    - Service to others → sorrow, compassion, sympathy
    - Status and shame → shame, guilt, embarassment

Emotion networks

- **Follow-up:** How do Q-SpAM-based semantic spaces correlate with those obtained from word embeddings and word association models?
  - Word associations *(Small World of Words;* De Deyne et al., 2019).

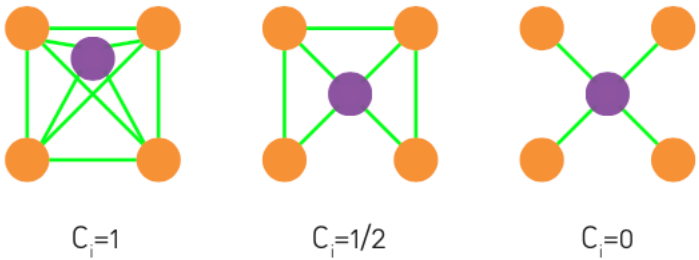
Association network for "happiness"



Association network for "felicidad"

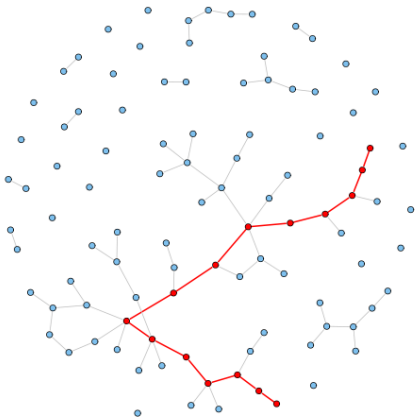
Future directions

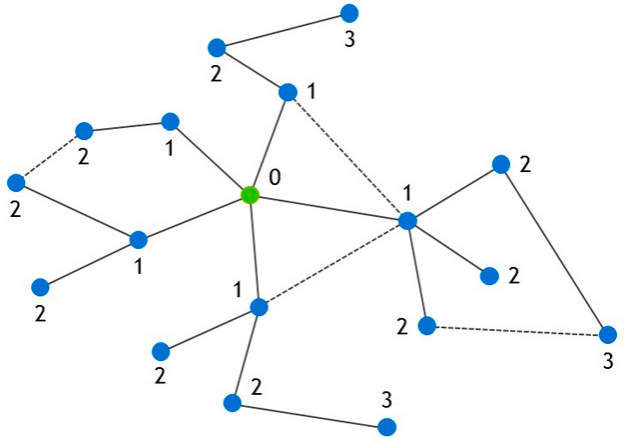
- Clustering-coefficient (CC): The degree to which nodes tend to group together.



Future directions

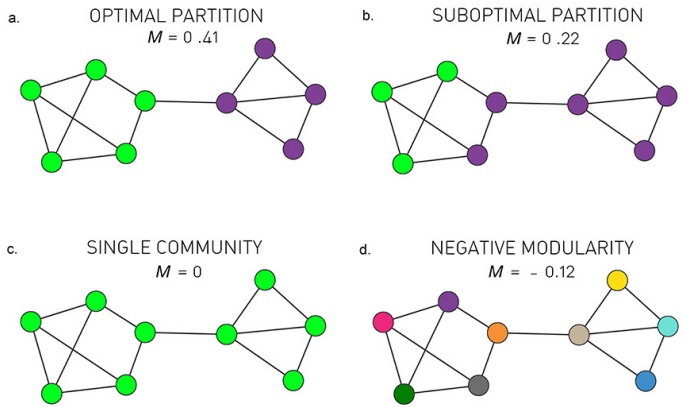
- Average shortest-path length (ASPL): The average distance between each pair of nodes.





Future directions

- Modularity (Q): The degree to which the network comprises distinct communities.



Future directions

Three critical indices of structural organization

- **High clustering coefficient (CC)** → Better semantic organization in monolinguals (Christensen et al., 2018; Cosgrove et al., 2021), and in the L2 of bilinguals (Feng & Liu, 2023).
- **Low average shortest-path length (ASPL)** → Faster navigability within the lexicon (Siew et al., 2019; Siew & Guru, 2023).
- **Optimal modularity (Q)** → Increased knowledge (Siew & Guru, 2023) and verbal creativity (Kenett et al., 2014).