Similarity as the basis for meaning extensions

Annika Tjuka

Max Planck Institute for the Science of Human History

April 28, 2020

Starting Point

"Since metaphor is based on the perception of **similarities**, [...] when an analogy is **obvious**, it should give rise to the same metaphor in various languages; hence the **wide currency** of expressions like the 'foot of a hill' or the 'leg of a table'."

(Ullmann 1963)

What is “obvious similarity”?

Dimensions of Similarity: SHAPE

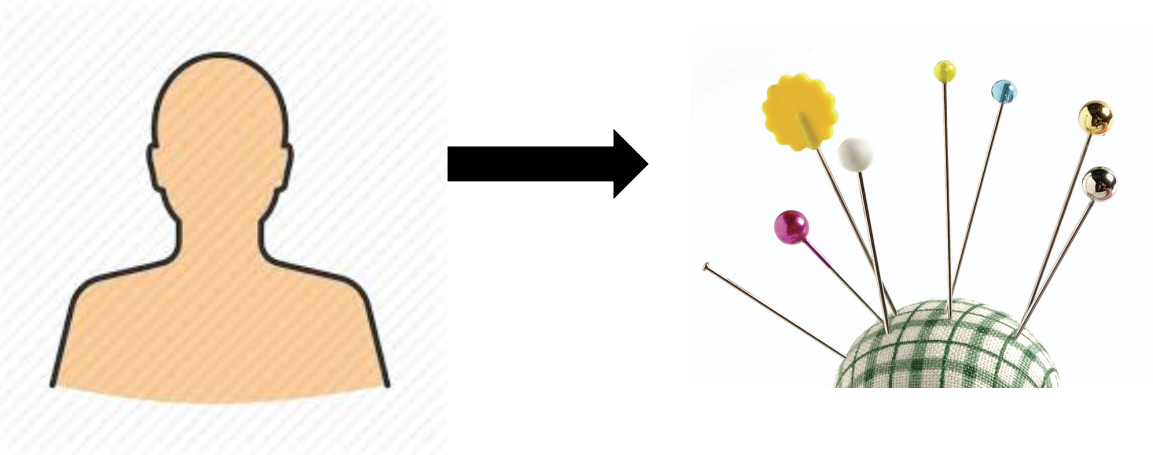

Figure 1: Example of the shape dimension: ‘pinhead’

Dimensions of Similarity: SPACE

Figure 2: Example of the space dimension: 'foot of the mountain'

Dimensions of Similarity: FUNCTION

Figure 3: Example of the function dimension: 'mouth of the river'

How can we quantify the notion of similarity?

Measures of Similarity

Data on psycholinguistic attributes, for example,

- concreteness
- imageability
- sensory modalities (i.e., hearing, taste, touch, smell, and vision)

Databases

- Concepticon: a reference catalog that links concept lists across different language varieties (List et al. 2019)
- CLICS: a database with cross-linguistic colexification patterns (Rzymski et al. 2020)

Databases

- Concepticon: a reference catalog that links concept lists across different language varieties (List et al. 2019)
- CLICS: a database with cross-linguistic colexification patterns (Rzymski et al. 2020)
- **A database with norms, ratings, and relations (NoRaRe) of words and concepts from psycholinguistic studies**

The NoRaRe Database

- wealth of data on properties such as frequency norms, concreteness ratings, sensory modality ratings, semantic field categorizations, etc.
- link the datasets to Concepticon
- hand-curated and automatic mapping workflow
- test-driven data curation

Outlook

Possible Applications for the NoRaRe Database

- Cognition: What are the factors that influence meaning extensions?

Possible Applications for the NoRaRe Database

- Cognition: What are the factors that influence meaning extensions?
- Culture: Do we see areal differences in the ratings of subjective word properties (e.g., emotional values)?

Possible Applications for the NoRaRe Database

- Cognition: What are the factors that influence meaning extensions?
- Culture: Do we see areal differences in the ratings of subjective word properties (e.g., emotional values)?
- Evolution: Is there a link between psycholinguistic criteria and language change?

Possible Applications for the NoRaRe Database

- Cognition: What are the factors that influence meaning extensions?
- Culture: Do we see areal differences in the ratings of subjective word properties (e.g., emotional values)?
- Evolution: Is there a link between psycholinguistic criteria and language change?

Thank you for listening!

References

Forkel, R., Greenhill, S. J., & List, J. M. (2017). Cross-Linguistic Data Formats (CLDF). Max Planck Institute for the Science of Human History: Jena.

List, J.-M., Rzymski, C., Greenhill, S., Schweikhard, N., Pianykh, K., Tjuka, A., Wu, M.-S. & Forkel, R (2020). Concepticon 2.3.0. Jena: Max Planck Institute for the Science of Human History. (Available online at http://concepticon.clld.org, Accessed on 2020-04-26.)

Rzymski, C., Tresoldi, T., Greenhill, S. J., Wu, M. S., Schweikhard, N. E., Koptjevskaja-Tamm, M., ... & Chang, S. (2020). The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. Scientific Data, 7(1), 1–12.

Ullmann, Stephen (1963). Semantic universals. In Joseph Greenberg (ed.), Universals of language 2, 373–396.

Method

Figure 4: Elicitation material.

Result

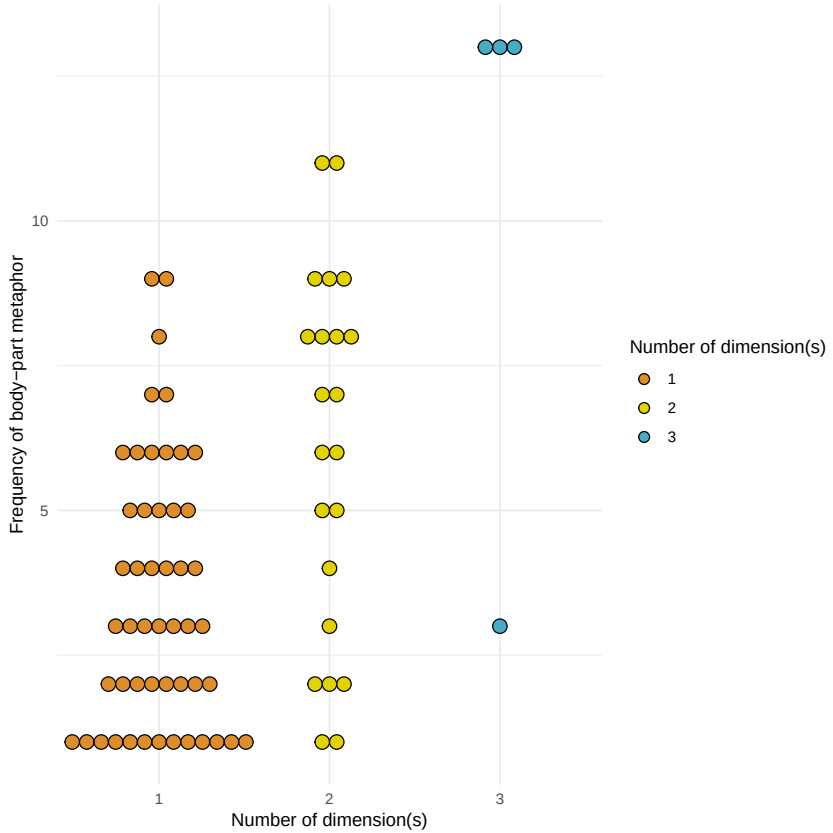

Figure 5: Frequency of body part extensions in relation to their classification into the three dimensions.

Research Question and Result

- Does the frequency of body part extensions across languages depend on the number of similarity dimensions?

Research Question and Result

- Does the frequency of body part extensions across languages depend on the number of similarity dimensions?
  - Yes, a body part extension like ‘leg of the table’ which is categorized in all three dimensions occurs throughout the entire language sample. In comparison, ‘foot of the mountain’ (space and function dimension) occurs in 7 out of 13 languages and ‘head of the house’, as in *roof*, (space dimension) is used in one of the sample languages.