

Background

What is the theoretical status of the relationship between meaning and morphosyntactic category? There is a great degree of cross-linguistic consistency in the relationship between the meaning of lexical items and their syntactic behaviour: the vast majority of languages clearly handle object words differently from action words. Yet exceptions exist both within classes and across languages.

How can a theory explain both these strong universal tendencies and well-established deviations from them? We focus on (visual) groundedness. Groundedness formalizes the notion of how much information a word conveys about an utterance's "meaning" in context: how meaningful vs. grammatical a word is, providing a continuous analogue of the lexical–functional distinction.

Hypothesis: Variation in word class organization will be associated with variation in groundedness.

Japanese adjectives

Unusually, Japanese adjectives form two distinct morpho-syntactic categories, rather than a single unified class. The split between the classes (called *i*-adjectives and *na*-adjectives) is not clearly phonological or semantic. Morpho-syntactically, *i*-adjectives are more verb-like and *na*-adjectives are more noun-like:

VERB-LIKE: INFLECTS

- (1) *yama-ga takai / takakatta.*
mountain-nom high / high.past
'The mountain is/was tall.' (*i*-adjective)

NOUN-LIKE: NEEDS COPULA

- (2) *Taroo-ga shizuka da / shizuka datta*
Taro-nom quiet cop / quiet cop.past
'Taro is/was quiet.' (*na*-adjective)

Read
more:



Visual groundedness as an organizing principle for word class: Evidence from Japanese

Coleman Haley

coleman.haley@ed.ac.uk

Sharon Goldwater

sgwater@ed.ac.uk

Edoardo Ponti

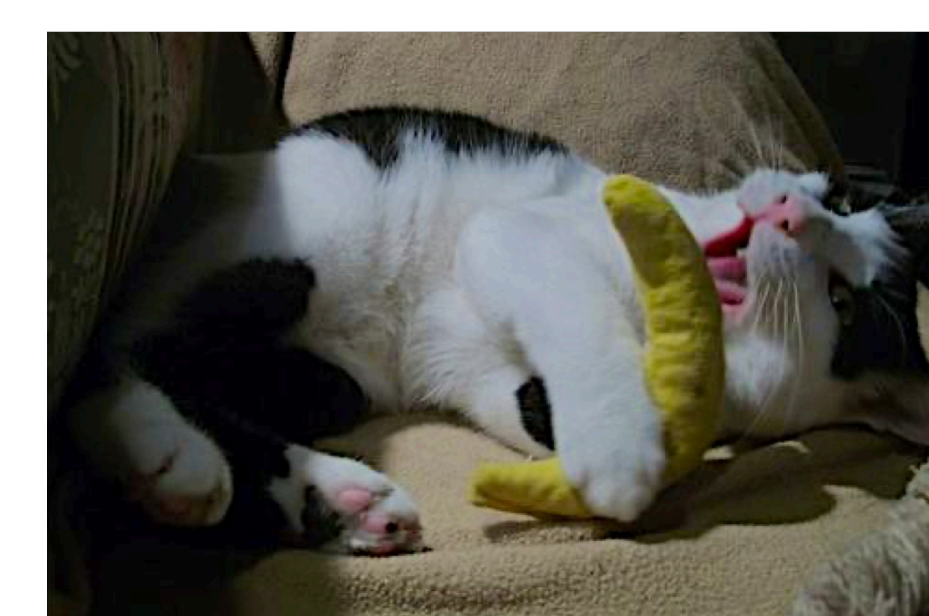
eponti@ed.ac.uk

Noun-like adjectives
more grounded than verb-like!

Quantifying groundedness

We define groundedness as the pointwise mutual information between a word *in context* and the meaning of an utterance. We represent meaning with an image, which allows us to estimate groundedness as a *surprisal difference* between a language model and an image captioning model:

$$\text{Groundedness}(w_t) = \log p(w_t | I, \mathbf{w}_{<t}) - \log p(w_t | \mathbf{w}_{<t})$$



$p(\text{cat} | A, \text{Image})$

Image captioning model

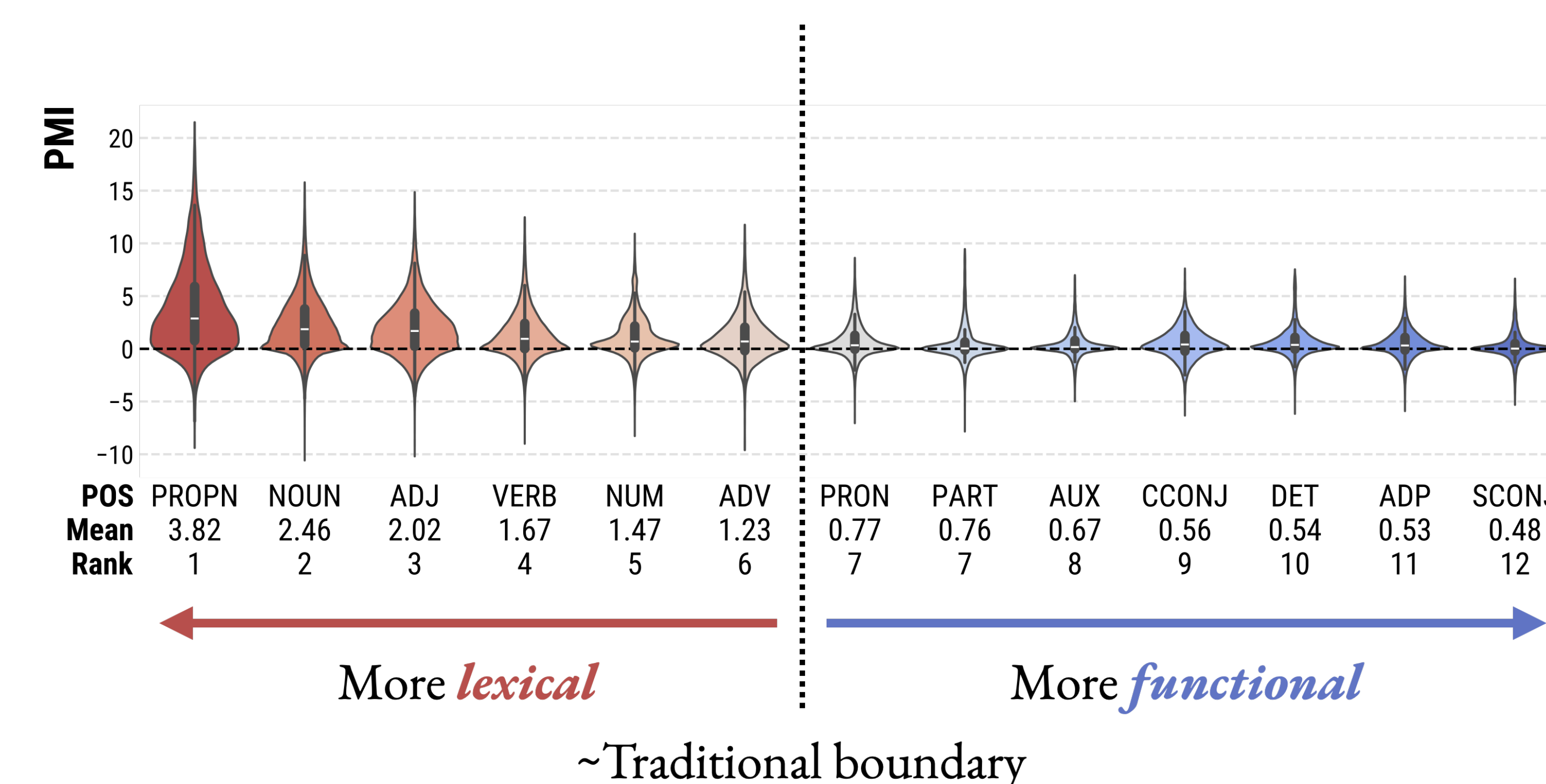
$p(\text{cat} | A)$

Language model

"A cat plays with a toy banana."

Groundedness cross-linguistically

In our prior work, we found groundedness captures a gradient lexical–functional distinction across 30 languages (incl. Japanese), showing a cline within traditional lexical classes (noun > adjective > verb).



Results in Japanese

Because Nouns > Verbs in terms of groundedness, we hypothesize that *na*-adjectives (noun-like) will have higher groundedness than *i*-adjectives (verb-like). We use *sudachipy* for tagging adjectives, filtering for adjective types which occur at least 5 times. We find *na*-adjectives have significantly higher groundedness in our manually-captioned corpora, and a null result in the machine-translated corpus.

To compute significance, we use a linear mixed-effects model with fixed effects of position and word class and a random effect for word type.

bits	Effect(<i>na</i>)	p-value	Dataset	MT?
0.16	0.68		COCO-35L	Y
0.90	0.029		XM3600	N
0.94	0.015		STAIR	N

Ruling out alternatives

Our groundedness scores are composed of two terms: LM surprisal and captioning surprisal. Could one of these explain the effect? (e.g. *na*-adjectives are less frequent). On our XM3600 corpus, we failed to find a significant effect of either LM surprisal ($p=0.133$; $\beta=1.17\pm0.77$) or captioning surprisal ($p=0.591$; $\beta=0.38\pm0.61$) alone. It is only through interaction between these factors that an association with word class emerges.