

Computational considerations of comparisons and similes

Vlad Niculae
vlad@vene.ro

Victoria Yaneva
V.Yaneva@wlv.ac.uk

University of
Wolverhampton

I. Metaphor, Simile, Comparison

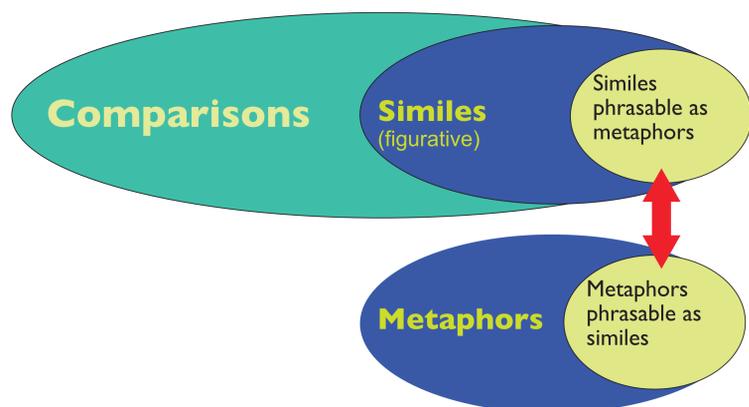
Classical view (Aristotle): "the simile also is a metaphor (...) the difference is but slight" (Rhetoric III, 4)

~ But... (Israel, 2004) ~

Metaphor: "The house had great bones." (Washington Post, 9/29/02)

Simile: "Two bears sat facing each other like two matrons having tea." (In "Wise Blood" by Flannery O'Connor)

Comparison: "He was like his father, except he had a crooked nose and his ears were a little lopsided." (In "Black cat" by Alex Krill)



II. Autism Spectrum Disorders and figurative language comprehension

Some people with ASD perceive only the literal meaning of phrases:

A request to "stick your coat down over there" is met by a serious request for glue (...) Tell him that his sister is "crying her eyes out" and he will look anxiously on the floor for her eyeballs... (Happé, 1995)

Literal comparisons can be understood by most people with ASD. Highly abstract similes or ambiguous use of "like" evoke confusion.

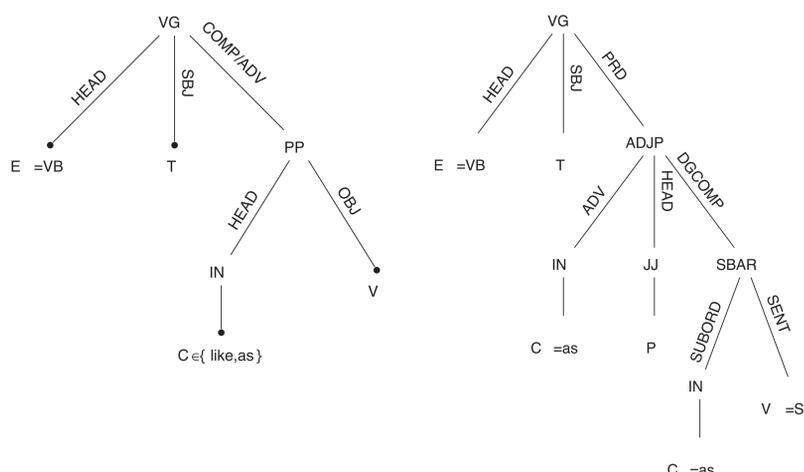
Understanding and making comparisons requires shifting of perception, which has been shown difficult to people with ASD.

III. Anatomy of a comparison

T: topic
E: eventuality
P: predicator
C: comparator
V: vehicle

[He **T**] [looked **E**] [like **C**] [a broiled frog **V**], [hunched **P**] over his desk, grinning and satisfied.

[His mouth **T**] [tasted **E**] [like **C**] [the bottom of a parrot's cage **V**]



IV. Recognizing comparison patterns

Pattern matching on GLARF enriched BLLIP parse trees

- 40 examples of similes from (Hanks, 2005)

	full	part	none	Partial match
comparison	24	5	4	P = 94%, R = 88%
non-comparison	1	1	5	

- Occurrences of *like* in BNC and VUAMC

	full	part	none	Partial match
comparison	0.17	0.07	0.33	P = 70.5%, R = 41.7%
non-comparison	0.05	0.05	0.33	

- Occurrences of *as* in BNC and VUAMC

	full	part	none	Partial match
comparison	0.11	0.05	0.09	P = 29.6%, R = 64.8%
non-comparison	0.26	0.11	0.39	

Causes of error:

- slightly different patterns: he didn't look much like a doctor
- phrases such as: hold your hands like this
- ambiguity of patterns: I feel like an ice cream

V. Future work

Comparison patterns

Shallow patterns: reduce parser errors, apply across languages (e.g. dependency edges, BIO tags)

Comparison pattern discovery: focus on other cues apart from "like" and "as"

[The retirement of Yves Saint Laurent **T**] [is **E**] [the fashion equivalent **C**] of [the breakup of the Beatles **V**]. (heard on the National Public Radio)

Snowball approach:
A ~ B frequent -> A ~ B frequent

Semi-automatic pattern annotation

Semantics

Classification tasks:
- comparison vs ambiguous patterns (e.g. ad hoc sets)
- simile vs comparison
- abstract vs concrete
- conventional vs creative

Simplification tasks:

- filling in missing **P**
- finding a better **E**
- representing meaning

Downstream applications

Textual entailment as extrinsic evaluation

Localization in machine translation: use appropriate stereotypes

Flexible IR queries: comparison normalization

Stylistic fingerprint: **V**'s used more frequently inside a community

Stereotype learning (Veale et al., 2008, 2011, 2012) using more patterns

Similes in other languages

The simile trope works very similarly in English, French, Bulgarian, Romanian at least.

Pattern discovery with minimal supervision should be applicable across languages.

What annotation and resources are necessary? What can be transferred?