Adjectives vs. Nouns:
Logical vs. non logical
Dimension integration operations

Galit W. Sassoon
ILLC, University van Amsterdam
Adjectives:

The challenge
Too many challenges...

Some basic issues pertaining to adjectives:
1. Dimensions and Multidimensionality
2. Difference from other categories, e.g., nouns
   Gradability, Scale structure, Measurement ...
3. **Nominalization**
4. Antonymy
5. **Context sensitivity** Vagueness, High flexibility
6. Modifier distribution,
7. ...
1. Multi-Dimensionality

• Being tall is merely a matter of height
• Being ‘healthy’ is a matter of degree in a variety of dimensions, such as blood-pressure, cholesterol and sugar.  
  Kamp 1975; Klein 1980
• One can be ‘healthy’ in some respects but not others.
• How to incorporate dimensions within formal semantics?
Formal semantics with dimensions

- In models $M = \langle D, [[ \cdot ]], f, c, \text{DIM} \rangle$, predicates $P$ are assigned:
  - denotations (entity sets) $[[P]]_M$
  - degree functions $f(P): D_x \rightarrow D_r$
  - a cutoff point $c(P) \in D_r$
  such that: $[[P]]_M = \{x \in D_x \mid f(P)(x) > c(P)\}$
  - a set of dimensions $\text{DIM}(P)$

- Dimensions are predicates whose interpretation restricts (and is restricted by) the interpretation of the predicate they are dimensions of.
  - $\text{DIM}(\text{tall}) = \{\text{tall}\}$;
  - $\text{DIM}(\text{healthy}) = \{\text{hwrt blood pressure, hwrt cholesterol, hwrt sugar, ...}\}$;
  - $\text{DIM}(\text{bird}) = \{\text{small, winged, feathered, flies, sings, nests, eats insects, ...}\}$.

How?
Nouns
Nominal concepts as similarity-based

- Extensive evidence that speakers associate nouns with rich clusters of dimensions, which create a typicality ordering and affect all cognitive tasks involving classification.

(most typical) ———— (most atypical)
Nouns
Nominal concepts as similarity-based

- Extensive evidence that speakers associate nouns with **rich clusters of dimensions**, which create a typicality ordering and affect all cognitive tasks involving classification.

- Psychologists associate nouns like ‘bird’ with graded structures – mappings of entities to degrees representing their **similarity to birds**, i.e. the extents to which their values on dimensions like ‘feathered’ and ‘small’ are close to the optimal values for birds.

- An entity is classified as a bird iff (roughly) its **average similarity on the dimensions** of bird, small size, flying, perching, etc. (or of some bird exemplar) **exceeds a threshold**.

(Murphy 2002; Hampton 1995; Lakoff 1987)
Nouns as similarity-based, support:

1- Relation between likelihood of categorization and similarity to the prototype.
2- Vagueness effects (borderline cases)
3- Online processing effects (typical < atypical).
4- No dimension is necessary / sufficient (Wittgenstein 1953)
5- Important dimensions (like "horse genotype") function as almost necessary + sufficient. They might be violated only if the values on the other dimensions may compensate (be sufficient to reach threshold; Hampton 1979).

... 

N- A finite memory representation for concepts (or predicate intensions):

Membership of infinitely many new instances can be determined on the basis of a finite set of known facts (dimensions and members).
Nouns as similarity-based categories:

-Thus, nouns are associated with multiple dimensions (characteristic features) which are incorporated through non logical operations such as averaging.
-at an early processing stage (implicit, procedural processing).
-In classification tasks, we do not consider for each dimension separately whether an entity has it or not, but rather, we sum* all the entity’s degrees and check whether the overall result is high enough.
Nouns as similarity-based categories:

- Thus, nouns are associated with multiple dimensions (characteristic features) which are incorporated through non logical operations such as averaging at an early processing stage (implicit, procedural processing).
- In classification tasks, we do not consider for each dimension separately whether an entity has it or not, but rather, we sum* all the entity’s degrees and check whether the overall result is high enough.

\[
\lambda x. \sum_{Q \in \text{DIM}(bird)} W(Q,bird) \times f(Q,x,bird) \geq \text{cutoff}(bird)
\]

The average on the set of bird dimensions of:

The relative importance \( W \) of each dimension \( Q \), times:

The similarity \( f \) between \( x \)'s value and the optimal value for birds in \( Q \).
No linguistic gradability

Degree morphology that cannot mark nominal gradability:

<table>
<thead>
<tr>
<th>English</th>
<th>Affixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tallest</td>
<td>#Birdest</td>
</tr>
<tr>
<td>The most beautiful</td>
<td>#The most bird</td>
</tr>
<tr>
<td>Too tall</td>
<td>#too bird</td>
</tr>
<tr>
<td>Tall enough</td>
<td>#bird enough</td>
</tr>
<tr>
<td>Very/perfectly/fairly tall</td>
<td>#very/....bird</td>
</tr>
<tr>
<td>(Twice) As tall as</td>
<td>#as bird as</td>
</tr>
<tr>
<td>Taller</td>
<td>#birder</td>
</tr>
<tr>
<td>More beautiful</td>
<td>#more bird</td>
</tr>
<tr>
<td>Units: 2 meters tall(er)</td>
<td>Only with nominalizations</td>
</tr>
</tbody>
</table>

Etcetra.
Kennedy 1997
no linguistic gradability, except...

-through (adjectival) mediation OR in between-noun comparisons:

• #Tweety is more a bird than Tan
• Tweety is a more typical bird than Tan
• Tweety is more (typical) of a bird than Tan
• Tweety is more a bird than a horse
• # Tweety is more tall than heavy

• Pretty much a chair

Cf. Marcin Morzycky’s work
2. Adjectives – Difference from Nouns

In psychology:

-Nouns (‘bird’) denote ‘object-categories’,
   Adjectives (‘red’) denote ‘properties’. But –

-What’s the difference?
   After all - ‘red’ refers to the set of red objects, and ‘bird’ to the property of being a bird.

This talk will try to say something about what underlies this intuitive ‘category’ vs. ‘property’ distinction.

But first more issues with adjectives.
3. Nominalization

This is not a talk about nominalization 😊
On this proposal, adjectival nominalizations

(\textit{height} for \textit{tall-}; \textit{health wrt bp} for \textit{healthy wrt bp})

are mass nouns such that

the adjective’s degree function maps entities to degrees representing quantities of the thing they denote.

Cf. Moltmann 2006
4. Antonymy:

- tall/short, - old/young, - safe/dangerous, - clean/dirty
- expensive/inexpensive/cheap/free
- identical/similar/different/dissimilar
- excellent/marvelous/fantastic/good/bad/terrible/…

• Why do adjectives come in antonym pairs?
• Or groups of near synonyms/antonyms?
• What underlies antonym polarity:
  - why do antonyms classify as positive and negative?
  - why the difference in cognitive salience? Pos > Neg

Lehrer 1985; Horn 1987; Paradis & Willner 2006; Giora 2006; Kamoen 2010, Tribushinina 2010
## Antonym morphology

<table>
<thead>
<tr>
<th>NON-essential</th>
<th>IRregular</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISapproved</td>
<td>DEvalued</td>
</tr>
<tr>
<td>Atypical</td>
<td>ABnormal</td>
</tr>
<tr>
<td>IMMoral</td>
<td>INappropriate</td>
</tr>
<tr>
<td>UNhealthy</td>
<td>UNDER-servant</td>
</tr>
<tr>
<td>PSEUDOintellectual</td>
<td>QUASI-judicial</td>
</tr>
<tr>
<td>SEMI-literate</td>
<td>SUB-human</td>
</tr>
<tr>
<td>COUNTER-attack</td>
<td>CONTRA-distinction</td>
</tr>
<tr>
<td>MALadministration</td>
<td>MISconstrued</td>
</tr>
<tr>
<td>ANTI-histamine</td>
<td>TasteLESS</td>
</tr>
</tbody>
</table>

Hamawand, 2009

- non-professional vs. unprofessional (not trained vs. under the exp. standard);
- non-essential vs. inessential (not necessary vs. not important);
- non-approval vs. disapproval (mere lack of approval vs. condemnation);
- atypical vs. untypical (only the latter can mean ‘bad’??);
- anti-revolution vs. counter-revolution
Positive and negative antonyms –

by assumption

Breakstone in progress & P.C.

Long - Short
Tall - Short
Deep - Shallow
Wide - Narrow
Old - Young

Warm - Cold
Heavy - Light
Expensive - Cheap
Rich - Poor
Late - Early
Fast - Slow

Good - Bad
Glad - Sad
Positive - Negative
Nervous - Relaxed

-still, the majority of intuitively negative antonyms in English are bare;
-In what sense are they negative? evidence?? Unclarity about the data
Negative antonymy

A generalization:

• The positive base (*tall, heavy*) denotes entities with **much** of a dimension (e.g., much height, much weight)

• The negative antonym (*short, light*) denotes entities **NOT having much** of it (having little height).
Negative antonymy

A generalization:

• The positive base (*tall, heavy*) denotes entities with *much* of a dimension (e.g., much height, much weight)
• The negative antonym (*short, light*) denotes entities *NOT* having *much* of it (having little height).

The syntactic negation theory of antonymy
(Rullmann 1995; Heim 2006 2008; Buring 2007):
- Instead of being specified in the lexicon, **antonyms are formed by a degree-negation operator, *little*,** hidden in the logical form of words such as *short*. 
Negative antonymy

BUT – in many adjectives the situation is reversed:

• The positive base (identical, similar, clean, healthy) denotes entities NOT having some/much of it (having little/no difference, dirt, sickness)

• The negative antonym (different, dissimilar, dirty, sick) denotes entities with some/much of a property (difference, dirt, disease).
  - objects count identical (or similar) in color iff no (or not much) color difference can be attested
  - entities count healthy if they do not possess much/any disease.
  - entities count clean if they do not possess much/any dirt, etc.
Negative antonymy

BUT – in many adjectives the situation is reversed:

• The positive base (*identical, similar, clean, healthy*) denotes entities *NOT* having some/much of it (having little/no difference, dirt, sickness)

• The negative antonym (*different, dissimilar, dirty, sick*) denotes *entities with* some/much of a property (difference, dirt, disease).
  - objects count *identical* (or *similar*) in color iff *no* (or *not much*) *color difference* can be attested
  - entities count *healthy* if they *do not possess much/ any* disease.
  - entities count *clean* if they *do not possess much/ any* dirt, etc.

-Kennedy 1997, 2001: dissociate linguistic and cultural negativity: *old* is linguistically positive, despite any cultural connotations;

-but- MANY dissociations between *negative cultural connotations* and *linguistic negativity* weaken the theory.
Negatives denote complement intervals

von Stechow 1984-Kennedy 2001

Positive adjectives map entities into bound intervals
Negative adjectives map entities into unbound intervals

But what about clean-dirty, safe-dangerous, healthy-sick,...?
### Linguistic negativity

<table>
<thead>
<tr>
<th>Measure-phrases</th>
<th>two meters tall</th>
<th>*short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio-phrases:</td>
<td>twice as tall as</td>
<td>&gt;</td>
</tr>
<tr>
<td>Nominalization</td>
<td>height</td>
<td>#‘shortness’</td>
</tr>
<tr>
<td>Un modification</td>
<td>unhappy</td>
<td>*unsad</td>
</tr>
</tbody>
</table>

*Many exceptions!!* 'Non-paradigmatic pairs'
Linguistic correlates of negative antonymy

<table>
<thead>
<tr>
<th>POS ☺ (☺)</th>
<th>NEG ☹</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Measure-phrases</td>
<td>two meters tall</td>
</tr>
<tr>
<td>Rare in most adjectives regardless of polarity</td>
<td></td>
</tr>
<tr>
<td>#two degrees warm</td>
<td>#two degrees cold</td>
</tr>
<tr>
<td>#80 kilos fat</td>
<td>#80 kilos skinny</td>
</tr>
<tr>
<td>twice as tall as &gt;</td>
<td>#twice as short as</td>
</tr>
<tr>
<td>~Balanced in most antonyms (literal reading- rare)</td>
<td></td>
</tr>
<tr>
<td>~twice as glad as ≅</td>
<td>~twice as sad as</td>
</tr>
<tr>
<td>~twice as similar as ≅</td>
<td>~twice as different as</td>
</tr>
<tr>
<td>• Ratio-phrases:</td>
<td></td>
</tr>
<tr>
<td>• Nominalization</td>
<td></td>
</tr>
<tr>
<td>height, width, depth, weight</td>
<td>#‘shortness’</td>
</tr>
<tr>
<td>similarity</td>
<td>≅</td>
</tr>
<tr>
<td>ease</td>
<td>≅</td>
</tr>
<tr>
<td>happiness</td>
<td>≅</td>
</tr>
<tr>
<td>• Un modification v.v.</td>
<td></td>
</tr>
<tr>
<td>unhappy</td>
<td>*unhappy</td>
</tr>
<tr>
<td>*utall, *unwide, *unold</td>
<td>*unsad</td>
</tr>
<tr>
<td>*unshort, *unnarrow…</td>
<td></td>
</tr>
</tbody>
</table>

van Cranenburgh, Sassoon & Fernandez 2010
5. Context sensitivity

• Vagueness  
  Kennedy 2007, van Rooij 2001  
  Sorites, gaps, borderlines, t.v. shifts

• High flexibility – ‘Secondary senses’:  
  - healthy (for my bank account);  
  - sick (abnormal);  
  - siiiick (= cool)…

• Dutch adjectives – difficulty with the incongruent condition (adj. >>. nouns)
An incongruent condition with adjectives?
Acceptability judgments, DUTCH
(with Tribushinina, Gulian and Timmer)

Our neighbors are

Nouns
2. Incongruent: #ideas.

Adjectives
4. Incongruent: #abstract.

1 = "totally unacceptable“, 7 = "totally acceptable“.

One way Anova with four correlated samples:

Congruent vs. Incongruent

<table>
<thead>
<tr>
<th>Nouns</th>
<th>Adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1 &gt;&gt; M2 P&lt;.01</td>
<td>M3 &gt;&gt; M4 P&lt;.01</td>
</tr>
</tbody>
</table>

Nouns vs. Adjectives

<table>
<thead>
<tr>
<th>Congruent</th>
<th>Incongruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1 &lt;&lt; M3 P&lt;.05</td>
<td>M2 &lt;&lt; M4 P&lt;.01</td>
</tr>
</tbody>
</table>

If adjectival predicates are in fact generally more acceptable than nominal ones, this supports an analysis whereby their interpretation is heavily dependent on the thing they are predicated of.
An incongruent condition with adjectives?

Acceptability judgments, DUTCH

"These bushes are ideas" vs. "These bushes are abstract".

‘unacceptable’

‘poetic’, ‘metaphoric’
### An incongruent condition with adjectives?
#### Acceptability judgments, DUTCH

**Instructions:**
A sentence is **acceptable** if it sounds normal, natural, you would not be surprised if you come across its meaning; a sentence is **unacceptable** if it sounds incorrect or strange.
1 = "totally unacceptable“, 7 = "totally acceptable”.

<table>
<thead>
<tr>
<th>English</th>
<th>Dutch</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>These events are tiny.</td>
<td>Deze gebeurtenissen zijn piepklein.</td>
<td>3.67</td>
</tr>
<tr>
<td>Drunk drivers are clear.</td>
<td>Dronken bestuurders zijn duidelijk.</td>
<td>3.00</td>
</tr>
<tr>
<td>His letters are greedy.</td>
<td>Zijn brieven zijn gierig.</td>
<td>2.91</td>
</tr>
<tr>
<td>All stories are unbreakable.</td>
<td>Alle sprookjes zijn onbreekbaar.</td>
<td>3.90</td>
</tr>
<tr>
<td>The kids are raw.</td>
<td>De kinderen zijn ruw.</td>
<td>4.36</td>
</tr>
<tr>
<td>The shrimp are relaxed.</td>
<td>De garnalen zijn relaxed.</td>
<td>3.70</td>
</tr>
<tr>
<td>Purple butterflies are logical.</td>
<td>Paarse vlinders zijn logisch.</td>
<td>3.33</td>
</tr>
<tr>
<td>Some liquids are messy.</td>
<td>Sommige vloeistoffen zijn rommelig.</td>
<td>3.50</td>
</tr>
<tr>
<td>These streets are strong.</td>
<td>Deze straten zijn sterk.</td>
<td>3.82</td>
</tr>
<tr>
<td>Her ideas are intact.</td>
<td>Haar ideeën zijn intact.</td>
<td>4.73</td>
</tr>
<tr>
<td>Our ceilings are harsh.</td>
<td>Onze plafonds zijn hardvochtig.</td>
<td>4.25</td>
</tr>
</tbody>
</table>
A proposal

Logical vs. non logical operations in the semantics of adjectives vs. nouns

-A trivial observation: If you are ‘healthy’, you cannot have any serious disease, whereas if you are ‘sick’, you must have some disease or other.
A proposal

Logical vs. non logical operations in the semantics of adjectives vs. nouns

-A trivial observation: If you are ‘healthy’, you cannot have any serious disease, whereas if you are ‘sick’, you must have some disease or other. Thus, ‘healthy’ means healthy in ALL dimensions that count (bp AND cholesterol AND sugar...), and ‘sick’ – sick in SOME dimension that counts (bp OR sugar OR...).
A proposal

Logical vs. non logical operations in the semantics of adjectives vs. nouns

-A trivial observation: If you are ‘healthy’, you cannot have any serious disease, whereas if you are ‘sick’, you must have some disease or other. Thus, ‘healthy’ means healthy in ALL dimensions that count (bp AND cholesterol AND sugar...), and ‘sick’ – sick in SOME dimension that counts (bp OR sugar OR...).

-Context restriction: We may consider one to be healthy despite, say, high blood-pressure only when this dimension is considered irrelevant.

When using ‘all’ or ‘everybody’, the standard practice is to ignore irrelevant entities (von Fintel 1994), but not to allow any other exceptions.
Generalized Quantifier theory

Determiners (every, some, ...) as relations between properties (sets of individuals):

If in M: Every boy walks, Every man talks, ...
[[Every]]_M = {<[[boy]]_M, [[walks]]_M>, <[[man]]_M, [[talks]]_M>, ...}

If in M: Some boy dances, Some boy doesn't dance, ...
[[Some]]_M = {<[[boy]]_M, [[dances]]_M>, <[[boy]]_M, [[doesn't dance]]_M>, ...}
Determiners \((\text{every, some, ...})\) as relations between properties \((\text{sets of individuals})\):

If in M: Every boy walks, Every man talks, ...
\([\text{Every}]_M = \{<\text{boy}_M, \text{walks}_M>, <\text{man}_M, \text{talks}_M>, \ldots\}\)

If in M: Some boy dances, Some boy doesn't dance, ...
\([\text{Some }]_M = \{<\text{boy}_M, \text{dances}_M>, <\text{boy}_M, \text{doesn't dance}_M>, \ldots\}\)

-Noun phrases as sets of properties \((\text{sets of individuals})\):

\([\text{Some boy}]_M = \{\text{dances}_M, \text{doesn't dance}_M, \text{walks}_M, \ldots\}\)
Generalized Quantifier theory

**Determiners** *(every, some, ...)* as relations between properties *(sets of individuals):*

If in M: Every boy walks, Every man talks, ...
[[Every]]_M = {<[[boy]]_M, [[walks]]_M>, <[[man]]_M, [[talks]]_M>, ...}

If in M: Some boy dances, Some boy doesn't dance, ...
[[Some]]_M = {<[[boy]]_M, [[dances]]_M>, <[[boy]]_M, [[doesn't dance]]_M>, ...}

-Noun phrases as sets of properties *(sets of individuals):*
[[Some boy]]_M = {[[dances]]_M, [[doesn't dance]]_M, [[walks]]_M, ...}

**Barwise and Cooper 1981; Keenan and Stavi 1986; Keenan 1996; 2003:**
- Which types of relations between sets are possible interpretations of natural language determiners? Constraints on possible interpretations limit the task faced by the language learner – thus help account for how the semantic system is learned with limited exposure to imperfect data.
- What natural subclasses of determiners can be distinguished?
### Some Types of Determiners in English

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lexical Dets</strong></td>
<td>every, each, all, some, a, no, several, neither, most, the, both, this, my, these, John's, ten, a few, a dozen, many, few</td>
</tr>
<tr>
<td><strong>Cardinal Dets</strong></td>
<td>exactly/approximately/more than/fewer than/at most/only ten, infinitely many, two dozen, between five and ten, just finitely many, an even/odd number of</td>
</tr>
<tr>
<td><strong>Approximative Dets</strong></td>
<td>approximately/about/nearly/around fifty, almost all/no, hardly any, practically no</td>
</tr>
<tr>
<td><strong>Definite Dets</strong></td>
<td>the, that, this, these, my, his, John's, the ten, these ten, John's ten</td>
</tr>
<tr>
<td><strong>Exception Dets</strong></td>
<td>all but ten, all but at most ten, every...but John, no...but Mary,</td>
</tr>
<tr>
<td><strong>Possessive Dets</strong></td>
<td>my, John's, no student's, either John's or Mary's, neither John's nor ..</td>
</tr>
<tr>
<td><strong>Evaluative Dets</strong></td>
<td>too many, a few too many, (not) enough, surprisingly few, ?many/?few</td>
</tr>
<tr>
<td><strong>Proportionality Dets</strong></td>
<td>exactly half the/John's, two out of three, (not) one...in ten, less than half the/John's, a third of the/John's, ten per cent of the/John's, not more than half the/John's</td>
</tr>
<tr>
<td><strong>Partitive Dets</strong></td>
<td>most/two/none/only some of the/John's, more of John's than of Mary's, not more than two of the ten</td>
</tr>
<tr>
<td><strong>Negated Dets</strong></td>
<td>not every, not all, not a (single), not more than ten, not more than half, not very many, not quite enough, not over a hundred, not one of John's</td>
</tr>
<tr>
<td><strong>Conjoined Dets</strong></td>
<td>at least two but not more than ten, most but not all, either fewer than ten or else more than a hundred, both John's and Mary's, at least a third and at most two thirds of the, neither fewer than ten nor more than a hundred,</td>
</tr>
<tr>
<td></td>
<td>every... and...</td>
</tr>
</tbody>
</table>
Boolean Closure - on all NL domains of interpretation

NOT (negation), AND (Conjunction), OR (disjunction)

V  Isn’t dancing
A  Not tall
NP Not every boy
DET Not every

Complement

\[[\text{not } C]\]_M = D_C - [[C]]
Boolean Closure - on all NL domains of interpretation
NOT (negation), AND (Conjunction), OR (disjunction)

V    Isn’t dancing    Dancing and singing
A    Not tall        Tall and intelligent
NP   Not every boy   Every boy and no girl
DET  Not every       At least 2 and at most 4
every man and woman

Complement
\[[\text{not } C]\]_M = D_c = [[C]]_M

Intersection
\[[\text{C1 and C2}]\]_M = [[C1]]_M \cap [[C2]]_M
Boolean Closure - on all NL domains of interpretation

NOT (negation), AND (Conjunction), OR (disjunction)

<table>
<thead>
<tr>
<th>V</th>
<th>Isn’t dancing</th>
<th>Dancing and singing</th>
<th>Dancing or singing</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Not tall</td>
<td>Tall and intelligent</td>
<td>Tall or intelligent</td>
</tr>
<tr>
<td>NP</td>
<td>Not every boy</td>
<td>Every boy and no girl</td>
<td>John or some girl</td>
</tr>
<tr>
<td>DET</td>
<td>Not every</td>
<td>At least 2 and at most 4 every man and woman</td>
<td>Either at least 2 or else at most 4</td>
</tr>
</tbody>
</table>

Complement
\[ [[\text{not } C]]_M = \overline{C} = [\overline{C}]_M \]

Intersection
\[ [[C_1 \text{ and } C_2]]_M = [C_1]_M \cap [C_2]_M \]

Union
\[ [[C_1 \text{ or } C_2]]_M = [C_1]_M \cup [C_2]_M \]
Generalized Quantifier theory (GQT)
Keenan and Stavi 1986:

- $\mathbf{D}^{\text{DET-BASIC}}$: \{Every\}_M, \{some\}_M, \{one\}_M, \{two\}_M, \ldots, \{John’s one\}_M, \ldots$

- $\mathbf{D}^{\text{DET}}$ – the set of all determiners of an NL – is the Boolean closure of $\mathbf{DET}^{\text{BASIC}}$
Generalized Quantifier theory (GQT)

Keenan and Stavi 1986:

\[-\mathbf{D}_{\text{DET-BASIC}}: [\text{Every}]_{\text{M}}, [\text{some}]_{\text{M}}, [\text{one}]_{\text{M}}, [\text{two}]_{\text{M}}, \ldots, [\text{John's one}]_{\text{M}}, \ldots\]

\[-\mathbf{D}_{\text{DET}} – \text{the set of all determiners of an NL} – \text{is the Boolean closure of } \mathbf{D}_{\text{DET-BASIC}}\]

i.e., we can assign a denotation to each determiner in the list by applying the intersection, union and/or complement operations (AND, OR and NOT) to denotations of basic determiners:

- at most 3 = NOT at least 4;
- exactly 3 = at least 3 AND (NOT at least 4);
- between 3 and 6 = at least 3 and (NOT at least 7)
- No = NOT one
- every man and woman = every man AND every woman
- not every, not John's, all boys and some girls, not john and every boy, ...
Generalized Quantifier theory (GQT)
Keenan and Stavi 1986:

\[ \mathcal{D}_{\text{DET-BASIC}}: \{\text{Every} \}_M, \{\text{some} \}_M, \{\text{one} \}_M, \{\text{two} \}_M, \ldots, \{\text{John’s one} \}_M, \ldots \]

\[ \mathcal{D}_{\text{DET}} \] – the set of all determiners of an NL– is the Boolean closure of \( \mathcal{D}_{\text{DET-BASIC}} \)

i.e., we can assign a denotation to each determiner in the list by applying the intersection, union and/or complement operations (AND, OR and NOT) to denotations of basic determiners:

- at most 3 = NOT at least 4;
- exactly 3 = at least 3 AND (NOT at least 4);
- between 3 and 6 = at least 3 and (NOT at least 7)
- No = NOT one
- every man and woman = every man AND every woman
- not every, not John’s, all boys and some girls, not john and every boy, ...

Boolean closure (conservativity) rules out most ways we might associate sets with determiners. The ratio between the set of all determiners and the set of conservative ones:

- For \( n = 1, 2 \);
- For \( n = 2, 2^7 = 128 \) (= 65536/512);
- For \( n = 3, 2^{37} \approx (137,000,000,000) \)
Back to Adjectives:

Both adjectives and nouns are associated with dimensions (or degree functions) that help to build sets of entities that eventually form their denotations, but:

– **What are the possible interpretations of natural language adjectives, as opposed to nouns?** Constraints on possible interpretations limit the task faced by the language learner – thus help account for how the semantic system is learned.

– **Is there a set of basic adjectives, $D_{ADJ\text{-BASIC}}$ (defined by certain formal properties of basic adjectival interpretations), such that we can build denotations for all other adjectives $D_{ADJ}$ via Boolean closure on this set?**

– **Determiners are logical constants, Adjectives aren’t** – their interpretation varies with context; so our task is even more significant – we *have* to say which basic interpretations are allowed, how they arise, and when and why any of the Boolean operations is applied.

– **What natural subclasses** of these interpretations can be distinguished?
Basic Adjectives:

1- A basic-adjective interpretation is based on a basic measurement (= measurement that’s consciously conceived as one-dimensional):

Examples: height, weight, healthy wrt cholesterol; b.p., number of chairs (“the room is empty except one chair”), number of empty chairs (“full except one chair”); number of stains (“clean except one stain”), amount of dust, amount of open doors (“closed except two doors”; ‘half closed’), IQ; grades on exams (“sufficient/good/very good in math”); ...

-Noun dimensions often are binary (e.g., wings, beak). Gradability is created due to averaging on many of them, not due to scalar measurement.
Basic Adjectives:

1-A basic-adjective interpretation is based on a single basic ((consciously conceived as) one-dimensional) measurement:
Examples: height, weight, healthy wrt cholesterol; b.p., number of chairs ("the room is empty except one chair"), number of empty chairs ("full except one chair"); number of stains ("clean except one stain"), amount of dust, amount of open doors ("closed except two doors"; ‘half closed’), IQ; grades on exams ("sufficient/good/very good in math"); ...

-Noun dimensions often are binary (e.g., wings, beak) and many. Gradability is created due to averaging on many of them, not due to measurement.

2-Unit based measure phrases are only licensed by basic Adjs (the ratio-scale ones) same with:
- literally interpreted ratio phrases;
- relatively ‘unproductive’ nominalization form s(hegth, weight) ?
- *un (*untall, *unheavy, *unwide, *undeep; ... ) ?
Basic Adjectives:

1-A basic-adjective interpretation is based on a single basic ((consciously conceived as) one-dimensional) measurement:

Examples: height, weight, healthy wrt cholesterol; b.p., number of chairs (“the room is empty except one chair”), number of empty chairs (“full except one chair”); number of stains (“clean except one stain”), amount of dust, amount of open doors (“closed except two doors”; ‘half closed’), IQ; grades on exams (“sufficient/good/very good in math”); ...

-Noun dimensions often are binary (e.g., wings, beak) and many. Gradability is created due to averaging on many of them, not due to measurement.

2-Unit based measure phrases are only licensed by basic Adjs (the ratio-scale ones)
- frequent literally interpreted ratio phrases;
- relatively ‘unproductive’ nominalization form (height, weight)
  -*un (*untall, *unheavy, *unwide, *undeep; ...)

3-Entities’ degrees must reach threshold to count as one-dim Adj
  healthy wrt cholesterol ⇔ λx. healthy-wrt-cholesterol(x) ≥ cutoff(healthy-wrt-cholesterol)
Basic Adjectives:

1-A basic-adjective interpretation is based on a single basic ((consciously conceived as) one-dimensional) measurement:
Examples: height, weight, healthy wrt cholesterol; b.p., number of chairs (“the room is empty except one chair”), number of empty chairs (“full except one chair”); number of stains (“clean except one stain”), amount of dust, amount of open doors (“closed except two doors”; ‘half closed’), IQ; grades on exams (“sufficient/good/very good in math”); ...

-Noun dimensions often are binary (e.g., wings, beak) and many. Gradability is created due to averaging on many of them, not due to measurement.

2-Unit based measure phrases are only licensed by basic Adjs (the ratio-scale ones)
-frequent literally interpreted ratio phrases;
-relatively ‘unproductive’ nominalization form (height, weight)
-*un

3-Entities’ degrees must reach threshold to count as one-dim Adj

healthy wrt cholesterol ⇔ λx. healthy-wrt-cholesterol(x) ≥ cutoff(healthy-wrt-cholesterol)

4-The selection of measurement and threshold heavily depend on context, namely on the thing the adjective is predicated of. Less so in nouns.

-Healthy dogs vs. human; full rooms vs. glasses,…
-Dutch – adjectives are virtually always interpretable, they ‘adopt’ to the subject.
Basic Adjectives:

1-A basic-adjective interpretation is based on a single basic (consciously conceived as) one-dimensional) measurement:
   Examples: height, weight, healthy wrt cholesterol; b.p., number of chairs ("the room is empty except one chair"), number of empty chairs ("full except one chair"); number of stains ("clean except one stain"), amount of dust, amount of open doors ("closed except two doors"; ‘half closed’), IQ; grades on exams ("sufficient/good/very good in math"); ...
   -Noun dimensions often are binary (e.g., wings, beak) and many. Gradability is created due to averaging on many of them, not due to measurement.

2-Unit based measure phrases are only licensed by basic Adjs (the ratio-scale ones)
   -frequent literally interpreted ratio phrases;
   -relatively ‘unproductive’ nominalization form (height, weight)
   -*un

3-Entities’ degrees must reach threshold on this measurement to count as Adj
   healthy wrt cholesterol ⇔ λx. healthy-wrt-cholesterol(x) ≥ cutoff(healthy-wrt-cholesterol)

4-The selection of measurement and threshold heavily depend on context, namely on the thing the adjective is predicated of. Less so in nouns.
   -Healthy dogs vs. human; full rooms vs. glasses,…
   -Dutch – adjectives are virtually always interpretable, they ‘adopt’ to the subject.

5-As we move towards more abstract/subjective domains conventions on measurement become less possible, but subjective ones are arguably still at work.
Boolean closure on basic Adj’ interpretations

1-Closure on NOT gives us one-dimensional antonymy: (cf. Heim 2008; Buring 2007,...)

Tall     ⇔     \( \lambda x. \text{height}(x) \geq \text{cutoff(tall)} \)
Short = NOT Tall     ⇔     \( \lambda x. \text{height}(x) < \text{cutoff(tall)} \)
Different in color    ⇔     NOT identical wrt color
                          ⇔     \( \lambda x. \text{healthy-wrt-ch’}(x) < \text{cutoff(healthy-wrt-ch’)} \)
Clean of stains       ⇔     NOT stained
                          ⇔     \( \lambda x. \text{stains}(x) < \text{cutoff(stains)} \)

Recall - At the one-dimensional level *clean* is basically negative, so why is *clean* intuitively positive? (and *safe, identical*, etc.)
Boolean closure on basic Adj’ interpretations

1-Closure on NOT gives us one-dimensional antonymy:  (cf. Heim 2008; Buring 2007,...)
   -short = NOT Tall ⇔ λx. height(x) < cutoff(tall)
   -clean wrt stains ⇔ Not stained

2-Closure on AND and OR (or ALL/SOME) gives us multi-dimensionality:
   A trivial observation: If you are ‘healthy’, you have no disease; if you are ‘sick’, you must have some disease. So
   ‘healthy’ = healthy in H1 AND H2 AND ... (ALL dimensions); ‘sick’ = sick in H1 OR H2 OR... (SOME dimension).
   -DIM_{M,w,g}(healthy) = { healthy wrt cholesterol, healthy wrt bp, hwrt sugar,...}
   -DIM_{M,w,g}(identical) = { identical wrt color, identical wrt intelligence, ...}
   -DIM_{M,w,g}(clean) = { clean of stains, clean of dust, of cake leftovers,...}

   -healthy ⇔ AND(DIM(healthy)) ⇔ λx. ∀Q∈DIM(healthy): Q(x) ≥ cutoff(Q)
      (‘healthy wrt ALL Q of DIM’)
   -clean      ⇔ AND(DIM(clean))  ⇔ λx. ∀Q∈DIM(clean): Q(x) < cutoff(Q)
             (‘clean wrt ALL Q of DIM’)

   -healthy
   -clean
Boolean closure on basic Adj’ interpretations

1-Closure on NOT gives us one-dimensional antonymy: (cf. Heim 2008; Buring 2007,...)
- short = NOT Tall ⇔ \( \lambda x. \) height(x) < cutoff(tall)
- sick wrt cholesterol = NOT healthy wrt cholesterol
- clean wrt stains ⇔ No stains

2-Closure on AND and OR (or ALL/SOME) gives us multi-dimensionality:
A trivial observation: If you are ‘healthy’, you have no disease; if you are ‘sick’, you must have some disease. So ‘healthy’ = healthy in H1 AND H2 AND ... (ALL dimensions); ‘sick’ = sick in H1 OR H2 OR... (SOME dimension).
- healthy ⇔ AND(DIM(healthy)) ⇔ \( \lambda x. \) \( \forall Q \in \text{DIM(healthy)}: Q(x) \geq \text{cutoff}(Q) \)
  (‘healthy wrt ALL Q of DIM’)
- clean ⇔ AND(DIM(clean)) ⇔ \( \lambda x. \) \( \forall Q \in \text{DIM(clean)}: Q(x) < \text{cutoff}(Q) \)
  (‘clean wrt ALL Q of DIM’)

3-Closure under NOT again gives us Multidimensional antonyms:
- Sick = NOT Healthy ⇔ \( \lambda x. \) \( \neg \forall Q \in \text{DIM(healthy)}: Q(x) \geq \text{cutoff}(Q) \)
  ⇔ \( \lambda x. \) \( \exists Q \in \text{DIM(healthy)}: Q(x) < \text{cutoff}(Q) \) (flu OR sugar OR bp ...)
- Dirty = NOT clean ⇔ \( \lambda x. \) \( \neg \forall Q \in \text{DIM(clean)}: Q(x) < \text{cutoff}(Q) \)
  ⇔ \( \lambda x. \) \( \exists Q \in \text{DIM(clean)}: Q(x) \geq \text{cutoff}(Q) \) (Stains OR dust OR leftovers...)
The double nature of negative antonymy:

- On this proposal, cultural and linguistic negativity usually match.

- The crux is that negation may also play a role at a level higher than that of a single dimension.

- Logical (rather than degree-) negation can affect the operators that bind the different dimensions of a single multidimensional adjective.

- So healthy = not sick?? But entities may be neither healthy, nor sick. Solt (2010):
  
  — Sentential negation (not tall) also creates gaps, e.g., entities may be neither tall nor not tall.
  — The use of antonyms (e.g. short) differs only in triggering larger gaps.
Multidimensional antonyms

At the multidimensional level –

**Positivity** = Universal generalizations
(healthy in all respects, clean of any dirt type)

**Negativity** = Existence of a counterexample
(not being healthy in all respects;
   not being free of any dirt-type, etc.)

- Dissociations between linguistic and cultural negativity – mainly in one-dime adjectives (late-early; expensive-cheap; old-young).
- A surprisingly tight correspondence between linguistic and ‘ethical’ negativity on the multidimensional level,
- Suggests that we use adjectives to posit our attitude towards things.
- Nouns are used to classify the world, adjectives to evaluate it.
Multidimensional antonyms

Prediction:
if antonyms are formed by a negation operator: \textbf{short} = \neg \textbf{tall},
Quantifier force will systematically vary in antonym pairs, because of negation:
- Healthy: \lambda x. \forall Q \in \text{DIM}(\text{healthy}), \text{healthy-wrt}(x,Q)
- Sick: \lambda x. \neg \forall Q \in \text{DIM}(\text{healthy}), \text{healthy-wrt}(x,Q)
= \lambda x. \exists Q \in \text{DIM}(\text{healthy}), \neg \text{healthy-wrt}(x,Q)

- \forall is the default for basic (unmarked; positive) lexical items;
- the antonyms (\exists) are derived through \neg,
- so POS >_{cog-salience} NEG

- We capture the cognitive prominence of positive antonyms (dimensions of their dominant interpretations are listed in the lexicon) vs. the complexity and negative connotation of negative ones (Horn 1989).
  - low frequency
  - slow processing
  - delayed acquisition
Multidimensional antonyms

The dimensions of Positive and Negative multidimensional Adjectives integrate via universal and existential quantification, respectively.

Positives are often conjunctions of one dimensional negatives:

- **Identical** *(similar)* = difference in no contextually relevant respects.
- **Different** = in at least one respect;

- **Clean** = dirt of **no sort** (dust, stains, etc.)
- **Dirty** = dirt of **some sort**

- Individuals with a slight cold are **not strictly speaking healthy** since they are **not all healthy**
- **One** type of sickness suffices to count as **sick**.

- **Safe / dangerous**; **Excellent / bad**, etc.
Multidimensional adjectives and measurement

Many alleged pos vs. neg features are confined to markers of basic adjectives:

<table>
<thead>
<tr>
<th>POS 😊 (贬)</th>
<th>NEG 贬</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure-phrases</td>
<td>two meters tall</td>
</tr>
</tbody>
</table>

Rare in most adjectives regardless of polarity because it is impossible at the multidimensional level

Ratio-phrases: twice as tall as > #twice as short as

Rare in most adjectives regardless of polarity because it is impossible at the multidimensional level

~twice as glad as ≅ ~twice as sad as
~twice as similar as ≅ ~twice as different as

Nominalization height, width, depth, weight #‘shortness’

Not so productive forms, characteristic of 1-dim ratio-scale (pos) adjectives; in multidim’ cases - other nominalization forms and in both polarity poles.

similarity ⪯ difference
ease ⪯ Difficulty
happiness ⪯ unhappiness

Un modification unhappy ⪯ *unsad

Productive forms, characteristic of multidim non-ratio-scale pos ∀adjectives v.v.

*utall, *unwide, *unold *unshort, *unnarrow...
SCOPE in derived comparatives

*More* takes one dimension at a time

If **healthy** = H1 and H2 and H3..., what does **healthier** mean?

1. Do we build a degree function for conjunctions? \( \text{er} > \text{and} \)
   
   Multivalued degrees of truth theories: \( \text{er}(H1 \text{ and } H2 \text{ and } H3) \)

2. Or do we interpret \( \text{er} \) in narrow scope: \( \text{and} > \text{er} \)

Boolean theories: \( H1\text{er AND } H2\text{er AND} \ldots \)?

If healthy = H1 and H2 and H3..., what does healthier mean?

1. Do we build a degree function for conjunctions? \( \text{er} > \text{and} \)
   
   Multivalued degrees of truth theories: \( \text{er}(H1 \text{ and } H2 \text{ and } H3) \)

2. Or do we interpret \( \text{er} \) in narrow scope: \( \text{and} > \text{er} \)

Boolean theories: \( H1\text{er AND } H2\text{er AND} \ldots \)?
1- Multi-valued logical theories

- Associate complex predicates (e.g., negated, conjunctive and disjunctive ones) with graded structures (a mapping of entities to numerical degrees).
- There are multiple choices for the fuzzy AND, OR, and NOT operators (Yen, 1999; Hajek, 2009).

a. \[[\neg P(x)]]_t = t_{\text{not},t}([[P(x)]]_t)

b. \[[P(x) \land Q(y)]]_t = t_{\text{and},t}([[P(x)]]_t, [[Q(y)]]_t)

c. \[[P(x) \lor Q(y)]]_t = t_{\text{or},t}([[P(x)]]_t, [[Q(y)]]_t)

A common choice: \( t_{\text{and}} = \times \) (product); and \( t_{\text{or}} = + \) (sum)

A Fuzzy natural language semantic theory:

Assuming that:
\[[\text{more } P]]_t = \lambda x_2 \in D_x. \lambda x_1 \in D_x. f_p(x_1) - f_p(x_2) > 0

\[[\text{more } P \text{ and } Q]]_t = \lambda x_2 \in D_x. \lambda x_1 \in D_x. f_{\text{P-and-Q},t}(x_1) - f_{\text{P-and-Q},t}(x_2) > 0

\[[\text{more } P \text{ or } Q]]_t = \lambda x_2 \in D_x. \lambda x_1 \in D_x. f_{\text{P-or-Q},t}(x_1) - f_{\text{P-or-Q},t}(x_2) > 0
2- Boolean theories

a. Tweety is big
b. Tweety is brown
c. Tweety is brown and big
d. Tweety is big and brown

∀t∈T: [[P (and) Q]]_t = [[P]]_t \cap [[Q]]_t
[[brown and big]]_t = \lambda x \in D_x.\text{brown}(x) \land \text{big}(x) = [[\text{brown}]]_t \cap [[\text{big}]]_t

A Boolean natural language semantic theory:

a. [[more P and Q]]_t = [[more P and more Q]]_t
   = \lambda x_2 \in D_x.\lambda x_1 \in D_x. (f_{P,t}(x_1) - f_{P,t}(x_2) > 0) \land (f_{Q,t}(x_1) - f_{Q,t}(x_2) > 0)

b. [[more P or Q]]_t = [[more P or more Q]]_t
   = \lambda x_2 \in D_x.\lambda x_1 \in D_x. (f_{P,t}(x_1) - f_{P,t}(x_2) > 0) \lor (f_{Q,t}(x_1) - f_{Q,t}(x_2) > 0)
SCOPE in derived comparatives

*More* takes one dimension at a time

If healthy = H1 and H2 and H3..., what does healthier mean?

Do we build a degree function for conjunctions?

Or do we interpret er in narrow scope: “H1er and H2er and ...”?

A study of conjunctions of one-dimensional gradable Adjectives:

Context: X and Y have equal price and different duration:

Result: X and Y are neither *more* nor *less* nor equally

“EXPENSIVE AND TIME CONSUMING”.

(about 90% of about 30 subjects)

Account: “more EXPENSIVE AND TIME-CONSUMING” =

“more EXPENSIVE and more TIME-CONSUMING”
Compound adjectives support Boolean analyses: Hebrew

### Table 2: The first conjunction condition:

Despite a 30kg weight difference (all other things being equal), Moshe is generally not judged *more fat and tall*.

<table>
<thead>
<tr>
<th>Section 3, e-f</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Moshe -100 kg; Danny -70 kg; both -1.95 cm tall</td>
<td>% ‘Yes’</td>
<td>‘No’</td>
<td>‘Yes’</td>
</tr>
<tr>
<td>e. Is Moshe more fat and tall than Danny?</td>
<td>6%</td>
<td>33</td>
<td>2</td>
</tr>
<tr>
<td>f. Is Moshe less fat and tall than Danny?</td>
<td>0%</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>g. Is it easier to determine that Moshe is fat and tall than that Danny is?</td>
<td>38%</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>h. Is it harder to determine that Moshe is fat and tall than that Danny is?</td>
<td>12%</td>
<td>30</td>
<td>4</td>
</tr>
</tbody>
</table>

### Table 3: The second conjunction condition:

When Moshe is *fatter* and Danny *balder*, generally, neither is judged to be *more fat and bald*.

<table>
<thead>
<tr>
<th>Section 4c-f</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Moshe -100 kg; Danny -70 kg; Moshe — not bald; Danny — bald</td>
<td>% ‘Yes’</td>
<td>‘No’</td>
<td>‘Yes’</td>
</tr>
<tr>
<td>c. Is any of them more <em>fat and bald</em> than the other?</td>
<td>0%</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>d. Is it easier to determine that one of them is fat and bald than that the other is?</td>
<td>3%</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td>e. Are they equally <em>fat and bald</em>?</td>
<td>9%</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td>f. Is it equally easy to determine that they are fat and bald?</td>
<td>6%</td>
<td>31</td>
<td>2</td>
</tr>
</tbody>
</table>
Compound adjectives support Boolean analyses: Hebrew

1- 74% (4g) of subjects agree that one is more fat or bald, yet-
2- less than half (16/35) were willing to provide a name indicating who is (4h).

-Thus, the 19 subjects (54%) not providing an answer to 4h interpreted more fat or bald as more fat or more bald, classifying both characters as such.
-They interpreted or as scoping over more, trying to determine whether Aharon is more fat or more bald and whether Danny is more fat or more bald; both are clearly the case as Aharon is fatter and Danny balder, rendering 4g true and at the same time making it impossible to choose one answer to 4h.

<table>
<thead>
<tr>
<th>Section 4g-j</th>
<th>%‘Danny’/ %‘Yes’</th>
<th>‘Moshe’/ ‘No’</th>
<th>‘Danny’/ ‘Yes’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moshe -100kg; Danny -70 kg; Moshe – not bald; Danny – bald</td>
<td>74%</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>g. Is one of them more fat or bald than the other?</td>
<td>63%</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>h. Who is more fat or bald?</td>
<td>68%</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>i. Is it easier to determine that one of them is fat or bald than that the other is?</td>
<td>77%</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>j. For whom is it easier to determine that?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: The first disjunction condition:
When Moshe is fatter and Danny balder, generally, both are judged to be more fat or tall.
Compound adjectives support Boolean analyses: Hebrew

Section 7
Two non-flying creatures; The first one doesn’t call; The second one is more typical of a calling creature

<table>
<thead>
<tr>
<th></th>
<th>%‘Yes’</th>
<th>‘No’</th>
<th>‘Yes’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Is the second more typical of a flying calling creature?</td>
<td>29%</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>b. Is the second more typical of a flying and calling creature?</td>
<td>37%</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>c. Is the second more typical of a flying or calling creature?</td>
<td>80%</td>
<td>7</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 7d: The ‘repetition’ condition - general
Compound adjectives support Boolean analyses: ENGLISH

Price AND Time - Siam is more TIME-CONSUMING AND EXPENSIVE than Vole is

- Yes: 0 (0%)
- No: 13 (100%)

Price AND Time - Siam is less TIME-CONSUMING AND EXPENSIVE than Vole is

- Yes: 0 (0%)
- No: 13 (100%)

Price AND Time - Siam and Vole are equally TIME-CONSUMING AND EXPENSIVE

- Yes: 0 (0%)
- No: 13 (100%)

Q2
Compound adjectives support Boolean analyses:

**Q1**

- **Price AND Time - Siam is more TIME-CONSUMING and more EXPENSIVE than Vole is**
  - Yes: 0 (0%)
  - No: 21 (100%)

- **Price AND Time - Siam is more TIME-CONSUMING AND EXPENSIVE than Vole is**
  - Yes: 0 (0%)
  - No: 21 (100%)

- **Price AND Time - Siam is less TIME-CONSUMING AND EXPENSIVE than Vole is**
  - Yes: 0 (0%)
  - No: 21 (100%)

- **Price AND Time - Siam and Vole are equally TIME-CONSUMING AND EXPENSIVE**
  - Yes: 0 (0%)
  - No: 21 (100%)
Compound adjectives support Boolean analyses: ENGLISH

Q2

Price OR Time - Siam is more TIME-CONSUMING OR EXPENSIVE than Vole is

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siam</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Vole</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

85% Yes  15% No

Price OR Time - Siam is less TIME-CONSUMING OR EXPENSIVE than Vole is

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siam</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Vole</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8% Yes  92% No

Price OR Time - Siam and Vole are equally TIME-CONSUMING OR EXPENSIVE

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siam</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Vole</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

46% Yes  54% No

Price OR Time - Siam and Vole are equally TIME-CONSUMING OR EXPENSIVE

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siam</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Vole</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

57% Yes  43% No
Derived comparatives, Preliminary results:
A-er ⇔ A1er and A2er and... ⇔ ALL n, A_n-er

healthier, more typical – \( \lambda x. \forall Q \in \text{DIM}: \text{er} (\text{healthy-wrt} (x, Q)) \).

Worse, sicker/ more sick – \( \lambda x. \exists Q \in \text{DIM}: \text{er} (\text{sick-wrt} (x, Q)) \).

- Er does not take scope over the dimension binding operation; it takes each dimension separately.

- Within category Er selects difference/ ratio measurements (so one-dim adj.s, not nouns)

- Gradability in adjectives is confined to basic single dimensional measurements; no averaging or other multidimensional grading seems to take place (at least as far as degree morphology is concerned/ shows anything). We can discuss typicality of “tall and wise people”, but not measurement degrees in “tall and wise”

- Dimensions of multidim’ Adj’s function as categorization criteria

Precisely, the opposite from nouns and their compounds!
Compound nouns support multi-valued analyses

-Hampton (1987; 1988a; 1997a,b) has analyzed ratings of goodness of example (typicality) of a list of entities in modified-nouns of the form 'Ps which are Qs' (such as, for instance, pets which are (not) birds) and in their constituents (e.g. pets and birds).

-The findings support a fuzzy semantic analysis: Nominal concepts associate with functions corresponding to the mean of entities in a variety of dimensions.

-For any item x, it is possible to predict x's typicality rating in, e.g., a modified-noun, \( f_{P \text{-and-} Q}(x) \), from x's ratings in the constituents.

-Hampton (1987; 1988a; 1997a,b)

a. \( f_{P \text{-and-} Q}(x) = W_P f_P(x) + W_Q f_Q(x) + W_{P \times Q} (f_P(x) \times f_Q(x)) \).

b. \( f_{P \text{-and-} \neg Q}(x) = W_P f_P(x) - W_Q f_Q(x) - W_{P \times Q} (f_P(x) \times f_Q(x)) \).

c. \( f_{P \text{-or-} Q}(x) = W_P f_P(x) + W_Q f_Q(x) - W_{P \times Q} (f_P(x) \times f_Q(x)) \).

-But in general, the typicality ratings in modified-nouns are better fitted by a composite-prototype representation, wherein the weight of each dimension is adjusted by a special function.
Thus, **gradability is different in nouns and adjective**

- The interpretation of adjectives like *tall* and *expensive*, for example, directly relate to conventional measurements of height and cost, rather than to a weighted mean in a set of dimensions.
- Adjectives can combine with *more* (or *er*) to create within-predicate comparisons (as in, e.g., *two meters taller*), *while* nouns and noun phrases do not
Thus, **gradability is different in nouns and adjective**

-Morphological gradability here seems to be confined to difference/ratio scale measurements (not operating on ordinal, multidimensional, averaging based ones). The dimensions of Nouns are bound at an early stage by the averaging op, hence are not accessible for grammar.

-While (conceptual) gradability of nominal compounds may be correctly described using some sort of **Fuzzy** semantics, adjectival conjunctions and disjunctions, are not systematically associated with degree functions. The **Boolean** operators *and and or*, take wide scope with respect to *more*, so that *more* modifies each conjunct/ disjunct separately (Sassoon 2011; Bale 2007).
Thus, **gradability is different in nouns and adjectives**

- The requirement for a unique dimension in the use of a comparative morpheme can only be abandoned in between-predicate comparisons.
- Conjunctive and disjunctive concepts seem to be felicitous and to receive interpretations with the connective in narrow scope in such comparisons (cf. (i)-(iii)),
- more systematic future research needs to carve out the precise set of interpretations that may be assigned to such statements.

  i. This is more a kitchen utensil than an electronic device.
  ii. This is more a piece of furniture and a game than a kitchen utensil or an electronic device.
  iii. Dan is more fat, bald and unhappy than good-looking, energetic and funny.
Modifiers

Universal modification is a cue to scale type OR standard type OR dimension binding type:

• All healthy
• Completely different, dissimilar
• Entirely bad

A Corpus study (18 As, ~1800 exception phrases):

• Healthy except dim >> sick except dim
• Not-sick except dim >> not-healthy except dim
• Healthy except dim >> not healthy except dim
• Not-sick except dim >> sick except dim

a. *Everyone* is happy except for Dan
b. *No one* is happy except for Dan
c. #*Someone* is happy except for Dan
A typology of predicates by their default dimension integration type

Conjunctive (A = A in every respect):
1a. ∨Dan is healthy except for high blood pressure  >>
   b. #Dan is not healthy except for blood pressure

Disjunctive (A = A in some respect):
2a. #Dan is sick except for (normative) blood pressure  <<
   b. ∨Dan is not sick, except for high blood pressure

Mixed (A = A in every/ some respect):
3.a ∨Dan is intelligent, except in mathematics  ≈
   b. ∨Dan is not intelligent, except in mathematics

None (A = A in many dimensions/ a high mean in A’s dimensions):
4.a # Tweety is a bird, except in size/wrt flying  ≈ 0
   b. # Tweety is not a bird, except in size/wrt flying  ≈ 0
.Exception phrases - Google Adjectives vs. nouns

36.5% of the ~1500 “adj. except” counts tested are of the type “adj. except a dim”.
Normal, Healthy, Typical, familiar, Healthier, identical, Bad, Sick, Atypical, Abnormal, Different, Unfamiliar, Worse, Dissimilar, Similar, Intelligent, Better, good

0% of the ~370 “noun except” counts tested are of the type “noun except a dim”.
bird, table, mother, capital, carrot
Exception phrases: Adjectives

36.5% of the “Adjective except” counts are of the type “Adjective except a dimension”
Nouns vs. Adjectives: An account

**Nominal** dimensions tend to combine via **averaging**, not via quantifiers. **Adjectival** dimensions tend to combine via **quantifiers**, not via averaging.
Corpus-based study: Predictions

The frequency of dimension-set uses with/without negation:

<table>
<thead>
<tr>
<th>Dimensional Uses</th>
<th>Conclusiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Conjunctivity</td>
</tr>
<tr>
<td>No. of positive dimensional uses ((P \ except \ dim))</td>
<td></td>
</tr>
<tr>
<td>No. of positive uses ((P \ except))</td>
<td></td>
</tr>
</tbody>
</table>

at least 2-3 times **higher in conjunctive adjectives** & **lower in Disjunctive adjectives** than

<table>
<thead>
<tr>
<th>Dimensional Uses</th>
<th>Disjunctivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>No. of negative dimensional uses ((not \ P \ except \ dim))</td>
<td></td>
</tr>
<tr>
<td>No. of negative uses ((not \ P \ except))</td>
<td></td>
</tr>
</tbody>
</table>
The frequency of dimensional readings of exception phrases with non-negated vs. negated As

Disjunctive | Mixed | Conjunctive
--- | --- | ---
0.06 | 0.09 | 0.29
0.30 | 0.33 | 0.55
0.62 | 0.69 | 0.91
0.91 | 1.00 | 1.14
1.20 | 1.75 | 3.85
3.85 | 4.82 | 4.84
4.84 | 6.12 | 6.87

A combined value –

Normalized conjunctivity –

Conj(A)
(Conj(A)+Disj(A))

Bad Sick Atypical Abnormal Different Worse Dissimilar Intelligent Better Good Similar Identical Healthier Familiar Healthy Typical Normal
Polarity as a predictive factor

- The 8 positive adjectives have higher normalized conjunctivity values than their 8 negative antonyms

- **Mean**: 0.7 vs. 0.246; **std**: 0.16 vs. 0.135, resp.
Polarity as a predictive factor

a. Normal, typical, healthy, familiar, healthier  Conjunctive  Positive  (n = 10)
b. Identical, similar, good, better, intelligent  Borderline-conjunctive

c. Dissimilar, worse, unfamiliar  Borderline-disjunctive

d. Bad, sick, atypical, abnormal, different  Disjunctive  Negative  (n = 8)

The conjunctivity values of positive adjectives are higher than their disjunctivity values (mean 0.51 vs. 0.25; Std: 0.22 vs. 0.21), whereas those of negative adjectives are smaller than their disjunctivity values (Mean: 0.17 vs. 0.44; Std: 0.18 vs. 0.23).

The conjunctivity values are higher in positive than in negative adjectives (Mean: 0.55 vs. 0.17; Std: 0.23 vs. 0.18), whereas the disjunctivity values are lower in positive than in negative adjectives (Mean: 0.25 vs. 0.44; Std: 0.22 vs. 0.23).
Derived comparatives
Is A-er of the same type as A?

Preliminary results: Yes...

- healthy, healthier – conjunctive \( \lambda x. \forall Q \in \text{DIM}: \text{er}(\text{healthy-wrt}(x,Q)) \).
- bad, worse – disjunctive, borderline disjunctive
- good, better – borderline

Compositionality appears to play its expected role (for future research to establish!)
Modifier distribution

Kennedy and McNally (2005); Winter and Rothstein (2005):

- Dirty is ‘partial’: Even minimally dirty entities are dirty.
- Clean is ‘total’: Only maximally clean entities are clean.
- Tall is ‘relative’: Its standard is context dependent

• Degree modifiers mark scale bounds and/or standard type:
  - completely clean/#dirty; - slightly dirty/#clean ...

• Language acquisition:
Degree modifiers are cues for acquiring standard type (Syrett 2007; Tribushinina 2010).
Standard type as a predictive factor

Frequency of co-occurrences of adjectives with degree-modifiers characterizing relative (very) vs. total (entirely, perfectly) vs. partial (partially, slightly) interpretations.

Assume the standard type is fixed once, not per a dimension.

**Total (maximum standard) => Conjunctive:**
Healthy = maximally healthy => maximally healthy in all dimensions.
\[(\text{cf. Mean}(H_1,\ldots,H_n) \geq 1)\]

**Partial (minimum standard) => Disjunctive:**
Sick = minimally (somewhat) sick => Sick in but one dimension.
\[(\text{cf. Mean}(H_1,\ldots,H_n) \leq 0)\]
Standard type as a predictive factor

A’s normalized totality = $\frac{\text{Total}(A)}{\text{Total}(A) + \text{Partial}(A)}$

$\text{Total}(A) = \frac{|\text{entirely/perfectly } A|}{|A|}$

$\text{Partial}(A) = \frac{|\text{slightly/partially } A|}{|A|}$

A’s normalized conjunctivity = $\frac{\text{Conj}(A)}{\text{Conj}(A) + \text{Disj}(A)}$

$\text{Conj}(A) = \frac{|A \text{ except Dim}|}{|A \text{ except}|}$

$\text{Disj}(A) = \frac{|\text{Neg A except Dim}|}{|\text{Neg A except}|}$

Correlation is $r = 0.308$ including the three comparatives, and $r = 0.62$ excluding the three comparatives.
Standard type as a predictive factor

Normalized totality vs. conj.

Normalized totality vs. conj.
excluding the comparatives and
bad \((r = 0.7; r^2 = 0.5):\)
Explicit binding of dimension variables

*(healthy/sick in all respects)*

Hence:

(i) dimensional readings may account for much of the distribution of degree modifiers *(especially/completely/slightly/no different).*

(ii) binding by *more* may account for the felicity of cross-polar multidimensional comparisons:

- *more similar than different*
- *more different than similar*

The number of dimensions wrt which a pair is *similar* compares to the number wrt which it is *not* (cf. Doetjes 2010).
5. Flexibility of Adjectival interpretation

a. ...working like a dog this weekend as I do every weekend. 24 hours in 2 days. Not healthy, except for my bank account
b. Patient 4 was atypical except for the high pitched voice
c. Siiick! (= cool)

In (a) *healthy* merely conveys ‘good’; with such a sense it can be used disjunctively.

In (b), a scientific context, *atypical* conveys “belongs to an atypical group”, which is, apparently, conjunctive (“exhibit all of the symptoms defining the group”).
Flexibility of Adjectival interpretation

Dimension-binding cues may explain certain context shifts:
- For example, uses of negative adjectives with cues of integration via ALL (e.g. ALL sick, sick EXCEPT) may trigger retrieval of secondary senses

- on one novel sense, sick conveys ‘cool’), and because integration via ALL is associated with positivity, these senses are positive (it is a complement to be called siiick).

Thus – modifiers are prevalent because they tell us what Boolean operations – hence which interpretations – to use (at least up to negative vs. positive encoded interpretations).

Neg adj occur with different degree modifiers – markers.
A noun-adjective dissociation? EEG

-A noun-adjective dissociation in terms of implicit versus explicit cognitive processing, respectively.

Implicit processing is typical of similarity-based categorization. It only enables awareness of its output, suggesting why grammatical operations can only access noun instances, not the mean-functions underlying their classification.

By contrast, explicit processing involves conscious reasoning about the categorization process, which is more demanding, as indicated by direct projections to frontal executive-functions, a delayed processing time-course, and difficulties in consistent use up to age ~5 (Ashby and Maddox 2005; Zelazo et al 2004).

Collaboration with E.Tribushinina, K.Timmer and M.Gulian (Utrecht-Leiden) has already began with a neurolinguistic experiment that studies the temporal course of adjectival versus nominal processing, by measuring electrical activity along the scalp (EEG).

A solution to the category vs. property puzzle related with the noun-adjective word class distinction?
Conclusions

A hypothesis worth further research:

**Quantification over dimensions** is a component of interpretation of many adjectives, but not nouns, which explains intuitions underlying negative antonymy, modifier selection, comparison and context-dependency.

**Ongoing research:**
- Corpora
- EEG
- Surveys
THANK YOU!
Comments are most welcomed
galitadar@gmail.com
Two types of categorization systems

Preliminary evidence for the thesis that nouns and adjectives trigger processing by different cognitive systems.
Two types of categorization systems

Ashby and Maddox (2005), *Annual Review of Psychology* 56, 149-78:

The huge literature concerning classification and concept learning appears to contain inconsistent results.

A consistent picture is revealed when studies are divided by categorization task.

Rule vs. similarity-based categorization tasks have different neural and developmental correlates.
Rule vs. Similarity-based categorization

Ashby and Maddox (2005), Annual Review of Psychology 56, 149-78:

Classification in **rule-based categories** depends on a **single dimension** or a simple enough **conjunction** or **disjunction** of dimensions.

Conversely, in **similarity-based categories**, information about instances’ degrees in **multiple dimensions** is integrated at an early processing stage, typically by **averaging**.
Adjectives vs. nouns: The Rule vs. Similarity hypothesis

Adjectives are rule-based:
Classification depends on a single dimension or a simple enough conjunction or disjunction of dimensions.

Nouns are similarity-based:
Information about instances’ degrees in multiple dimensions is integrated at an early processing stage, typically by averaging.
Rule vs. Similarity: The cognitive distinction
Ashby and Maddox (2005), Annual Review of Psychology 56, 149-78:

(i) Rules are easy to reason about explicitly.
(verbatim, declarative memory)

Conversely, similarity-based criteria are hardly accessible through introspection.
Processing is reflexive, automatic, ...
(implicit procedural memory)
Why are nouns and adjectives intuitively seen as denoting categories and properties, resp.?

Both nouns and adjectives denote object sets.

But –

On the one hand, more features of the denoted objects are encoded in the interpretation of nouns than of adjectives –

so nouns are more identified with the objects

On the other hand, only adjectival processing allows awareness of the dimension

so adjectives are more identified with predication of single properties
Rule vs. Similarity: The neural distinction
Ashby and Maddox (2005), Annual Review of Psychology 56, 149-78:

(ii) Rule-based (‘Boolean’) processing tasks
Employ an explicit hypothesis testing system that uses working memory and executive functions, which recruit primarily frontal-striatal regions.
The anterior cingulate, the prefrontal cortex and the head of the caduate nucleus

Conversely, Similarity (‘mean-based’) tasks involve implicit procedural learning.
The tail of the caduate nucleus; no direct projection to prefrontal cortex = non awareness

Details about dissociations: pp 163-165
Neural correlates

Bookheimer 2002 and references therein; Martin 2003: 66-67

Noun processing tasks: in naming tasks with object categories which are typically nominal, temporal, rather than frontal activation is typical.

Adjective processing tasks: Left frontal activation for similarity & abstract/concrete judgments (BA 45,47)
Rule vs. Similarity: The developmental distinction

Ashby and Maddox (2005), Annual Review of Psychology 56, 149-78:

(iii) Children younger than 3 years old consistently perform similarity-based tasks.

Conversely, due to late maturation of the prefrontal cortex, by age 3 children tend to still have difficulties in consistently using rules (Kyle 1979), which only get resolved at age 5 (Frye et al 1995; Zelazo et al 1996, 2004).
Rule vs. Similarity: The developmental distinction

For ex.:
Given a pile of cards with red and blue triangles, 3 year olds succeed in tasks such as “if red, put here, otherwise put there”.

Yet, they fail when circles are added and the instructions demand paying selective attention -
- to one of two dimensions: -”if circle,...; if triangle,...”
-Or to their conjunction:
-“if a red triangle,...”
Rule vs. Similarity: The developmental distinction

For ex.:
Given a pile of cards with red and blue triangles, 3 year olds succeed in tasks such as “if red, put here, otherwise put there”.

Yet, they fail when circles are added and the instructions demand paying selective attention -
- to one of two dimensions: -”if circle,...; if triangle,...”
-Or to their conjunction: -“if a red triangle,...”

These are precisely the capacities required for the interpretation of adjectives.
Adjectives and nouns as typically triggering processing by the rule vs. similarity systems

Adjectives have a respect argument, which can be saturated:
1) Tweety is healthy with respect to blood pressure

Alternatively, it can be bound:
2) Tweety is healthy in every / some / most respect(s)

And it can be implicitly saturated or bound:
3) Tweety is healthy

⇒ Speakers and listeners have to pay selective attention to one of multiple dimensions, or to conjunctions or disjunctions of dimensions.

⇒ Adjectival dimensions are typically processed with the rule-based system (they combine via Boolean operations)
Nouns do not have a respect argument

It can be saturated:
1) Tweety is healthy with respect to blood pressure
   #a bird with respect to size/ flying

It can be bound:
2) Tweety is healthy in every / some / most respect(s)
   #a bird in every / some / most respects / #generally a bird

It can be implicitly saturated / bound:
3) Tweety is healthy

Apparently ‘exceptional’ nouns that appear to be licensing a dimensional argument (e.g., health, similarity, an Italian, a genius, etc.,) virtually always have corresponding adjectival entries (e.g., healthy, similar, Italian, etc.)

⇒ Nominal dimensions are not handled each separately, they are integrated via averaging
Transformation operations

**Context:** We discover that birdhood depends on ten genes (categorization criteria):

Tan is a bird wrt to gene 1-6 but not a bird wrt genes 7-10.
Developmental correlates (2/7)

Consistent with findings from noun /adjective acquisition:

- Acquisition of adjectives (relatively frequent and consistent use) is significantly delayed
- This occurs in children 3-5 years old.

Syrett, Kennedy and Lidz, 2009:
At age 3, children already gain control of meaning components of one-dimensional adjectives
(e.g., of comparison classes, standard types, etc.)

But –
adjectives are not consistently used:
Developmental correlates (3/7)


Nouns are learnt earlier, faster and with fewer errors:

1.5 year olds learn nominal labels of artificial categories (“this is a dax”), while 3-year olds have difficulty with adjectival labels (“this is a dax one”; “dax, not red”).

Errors in understanding are frequent with adjectives at age 3-5, (consisting in confusion between dimensions, e.g., dark / loud).

Keil 1979: Children (up to age 10) often base categorization on similarity.

Tribushinina to appear: children have difficulties combining world knowledge and context

Fewer languages have the Adj. word class.
Developmental correlates (6/7)

Polinsky 2005:
Unlike competent speakers (n=4), adults incomplete learners, whose acquisition of Russian was interrupted at age 4 (n=3), 5 (n=1), and 6 (n=1) perform significantly better with nouns than with adjectives:
Nouns (and verbs) >> Adjectives

- Recognition time is longer for adjectives than for nouns only in incomplete learners,
- Translation to a second language is less accurate, and is based on words from other classes, etc.
(these word class effects are stronger than frequency effects; Polinsky 2005: 423)
4. Developmental correlates: A connectionist network

Gassar and Smith 1998 (exp. 1,3; p302):

- The network learns ‘nominal’, i.e. (5-dimensional) similarity-based invented categories faster than ‘adjectival’, i.e. one-dimensional invented categories.

- The difficulty with ‘adjectives’ occurs when ‘distracting’ dimensions are available (black vs. white is difficult iff attention to size, shape, etc. has to be suppressed).
5. Developmental correlates: The acquisition of conjunctions / disjunctions

Gedalyovich, 2003:
The semantic acquisition of conjunctions and disjunctions (coordination constructions) is also delayed to age 5 (Hebrew).

- Same time course as reported for adjectives, which is in accordance with the thesis that adjectival dimensions are processed as rules (conjunctions / disjunctions) and consistent use of rules develops late (at age 5).