PLIN3005/PLING229
Advanced Semantic Theory B

Lecture Notes

Yasutada Sudo
University College London
y.sudo@ucl.ac.uk

Spring 2015

Contents

Preface

This document is the main reading for PLIN3005/PLING229 Advanced Semantic Theory B, Spring 2015. The main topics of the course this year are **degree semantics** and **plurality** (Last year, we discussed **events** and **plurality**). Correspondingly, the present lecture notes have two parts. Each chapter is for one class, and we will cover Part I on degree semantics before the reading week, and Part II on plurality after the reading week.

The material in Part I heavily draws on Rick Nouwen's lecture notes from his graduate course taught at the Massachusetts Institute of Technology in Fall 2011 (cited here as Nouwen 2011), and Chapter 3 of Marcin Morzycki's manuscript on modification (cited here as Morzycki 2014), which is to be published from the Cambridge University Press. Nouwen's lecture notes are probably not available to you, but I will share them with you upon request. Morzycki's book is available on his website (as of December 2014), and is highly recommended. It covers various topics about 'modification' (which he thinks is an ill-defined term), and is accessible, if you know Heim & Kratzer (1998). Also it is very nicely written and makes you read (which is very rare in academic writing, sadly).

As for plurality, there seems to be no textbooks to the topic that is comprehensive and at the same time accessible to students at this level, unfortunately. I find Lucas Champollion's lecture notes (cited here as Champollion 2014) very useful, but they are probably a bit too advanced, especially the second half, but I nonetheless recommend them as further reading. They are available on his website (as of December 2014). Other useful book-long references on plurality include Schwarzschild (1996), Landman (2000) and Winter (2001a). Schwarzschild (1996) and Winter (2001a) are revised versions of their Ph.D. dissertations (from the University of Massachusetts, Amherst, and Universiteit Utrecht, respectively), and contain a number of original proposals, some of which we will discuss in class, but for this reason, they are not meant to be textbooks. Landman (2000) is a book about event semantics and plurality that grew out of advanced lectures. Although this could be usable as a textbook, perhaps in an advanced graduate course, but it is very technical and also delves into event semantics, which we have not time to talk about this term. But if you want to know more about plurality than what we can cover in the course, it is surely a good read.

For some of the sub-topics, you can also find survey/handbook articles. I list some of them below.

- Kennedy (2011) and Sorensen (2013) on vagueness

- Schwarzschild (2008) on degree constructions
- Lasersohn (2011) on mass nouns and plural nouns
- Doetjes (2012) on mass/count from a cross-linguistic perspective
- Nouwen (to appear) and Scha & Winter (to appear) on plurality (distributivity, cumulativity)

Chapter 0

Review of Advance Semantic Theory A

This chapter is a quick review of PLIN3004/PLING218 Advanced Semantic Theory A.

0.1 Model-Theoretic Semantics

- The theory of semantics developed there is **model-theoretic semantics**.
- Model-theoretic semantics aims at explaining native speakers' semantic intuitions about various types of sentences, in particular **truth-conditional intuitions**: When given a declarative sentence and a situation, they can tell whether the sentence is true or false.
- In order to account for the fact that native speakers have truth-conditional intuitions about infinitely many grammatical sentences, we assume that English and other natural languages obey the **(Local) Compositionality Principle**.

> **The (Local) Compositionality Principle**
> The meaning of a syntactically complex phrase A is determined solely by the meaning of its immediate daughters and the syntactic structure.

- Every grammatical phrase in (a 'fragment' of) English is assigned a model-theoretic object as its **denotation** relative to a **model**. The meanings of syntactically complex phrases are derived by combining meanings of their parts, as the Compositionality Principle states.

0.2 Models

- A model $\mathcal{M}$ is meant to model a particular state of affairs. Meaning in model-theoretic semantics is always relative to some model, reflecting the intuition that truth/falsity is relative to a state of affairs.

- Formally, a model $\mathcal{M}$ is a pair $\langle \mathcal{D}, \mathcal{I} \rangle$, where $\mathcal{D}$ is a non-empty set of individuals/entities and $\mathcal{I}$ is a function that assigns meanings to constants.
- The denotation of a given (gramamtical) phrase α relative to assignment a and model $\mathcal{M}$ is written as $[\![\alpha]\!]^{a,\mathcal{M}}$. The assignment is used to interpret traces and pronouns (see below).
- Model-theoretic objects are either of the following three kinds:
  - Individuals/Entities
  - Truth-values (0 or 1)
  - Functions of various kinds
- Model-theoretic objects are grouped according to their **type**s.

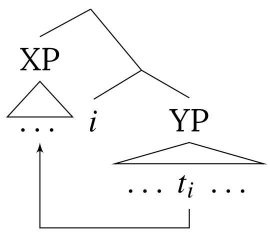
(0.1) *Types*
  a. e is a type.
  b. t is a type.
  c. If σ and τ are both types, $\langle \sigma, \tau \rangle$ is also a type.
  d. Nothing else is a type.

(0.2) *Domains*
  a. D_e is the set of individuals, $\mathcal{D}$.
  b. D_t is the set of truth-values, $\{0,1\}$.
  c. $D_{\langle \sigma, \tau \rangle}$ is the set of functions whose domain is D_σ and whose range is D_τ.

- Notation: We sometimes write et for $\langle e, t \rangle$.

0.3 The Syntax-Semantics Interface

- We assume that the syntax generates hierarchically organised syntactic objects called LFs (or 'Logical Forms'). We assume that branching is at most binary.
- LFs are fully disambiguated with respect to lexical and structural ambiguities, including quantifier scope.
- In particular we assume that Quantifier Raising has resolved type-mismatches that arise with quantifiers in non-subject position.
- We also assume that a syntactic movement, overt or covert, creates a binding index node and a trace with the same index, where indices are simply natural numbers. This is depicted in the following schematic diagram where $i \in \mathbb{N}$.



- The semantic is composed of two major components.

- The **lexicon** is a list of meanings of morphemes (it may contain other information necessary for morphology, syntax and phonology, but we will not be concerned with it).
- The **compositional rules** are instructions as to how to combine different kinds of meanings.

- Our semantic theory is **type-driven**, meaning that the semantic types (as opposed to, say, syntactic labels) determine how the semantic composition proceeds.
- The task of a semanticist is to enrich the lexicon and/or compositional rules so as to cover more phrases and constructions of various natural languages.

0.4 Semantic System

In Term 1, we covered the following types of items.

- Proper names denote individuals/entities.

(0.3) For any assignment a and for any model $\mathcal{M}$,
a. $[\![\textbf{Nathan}]\!]^{a,\mathcal{M}} = \mathcal{I}(\textbf{Nathan}) = \text{Nathan}$
b. $[\![\textbf{Andrea}]\!]^{a,\mathcal{M}} = \mathcal{I}(\textbf{Andrea}) = \text{Andrea}$

- Intransitive predicates denote functions of type $\langle e,t\rangle$.

(0.4) For any assignment a and for any model $\mathcal{M}$,
a. $[\![\textbf{smokes}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x \text{ smokes in } \mathcal{M}$
b. $[\![\textbf{linguist}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x \text{ is a linguist in } \mathcal{M}$
c. $[\![\textbf{British}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x \text{ is British in } \mathcal{M}$

(More formally, $[\![\textbf{smokes}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x \in \mathcal{I}(\textbf{smokes})$)

- Transitive predicates denote functions of type $\langle e,\langle e,t\rangle\rangle$.

(0.5) For any assignment a and for any model $\mathcal{M}$,
a. $[\![\textbf{likes}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ y \text{ likes } x \text{ in } \mathcal{M}$
b. $[\![\textbf{part}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ y \text{ is part of } x \text{ in } \mathcal{M}$
c. $[\![\textbf{fond}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ y \text{ is fond of } x \text{ in } \mathcal{M}$
d. $[\![\textbf{from}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ y \text{ is from } x \text{ in } \mathcal{M}$

(More formally, $[\![\textbf{likes}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\langle x,y\rangle \in \mathcal{I}(\textbf{likes})$)

- Semantically vacuous items denote identity functions.

(0.6) For any assignment a and for any model $\mathcal{M}$,
a. $[\![\textbf{a}_{\textbf{predicative}}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P$
b. $[\![\textbf{is}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P$
c. $[\![\textbf{of}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$
d. $[\![\textbf{who}_{\textbf{relative}}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P$

e. $[\![\textbf{such}_{\textbf{relative}}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ P$

- Quantificational determiners denote functions of type $\langle et, \langle et, t\rangle\rangle$.

(0.7) For any assignment a and for any model $\mathcal{M}$,

a. $[\![\textbf{every}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda Q \in D_{\langle e,t \rangle}.$
for every $x \in D_e$ such that $P(x) = 1$, $Q(x) = 1$

b. $[\![\textbf{no}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda Q \in D_{\langle e,t \rangle}.$
for no $x \in D_e$ such that $P(x) = 1$, $Q(x) = 1$

c. $[\![\textbf{some}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda Q \in D_{\langle e,t \rangle}.$
for some $x \in D_e$ such that $P(x) = 1$, $Q(x) = 1$

- These meanings combine via various **compositional rules**.
- For branching structures, we have three compositional rules:

Functional Application (FA)
For any assignment a and for any model $\mathcal{M}$, if **A** is a branching node with children **B** and **C** such that $[\![\textbf{C}]\!]^{a,\mathcal{M}} \in \text{dom}([\![\textbf{B}]\!]^{a,\mathcal{M}})$, then $[\![\textbf{A}]\!]^{a,\mathcal{M}} = [\![\textbf{B}]\!]^{a,\mathcal{M}}([\![\textbf{C}]\!]^{a,\mathcal{M}})$.

Predicate Modification (PM)
For any assignment a and for any model $\mathcal{M}$, if **A** is a branching node with children **B** and **C** such that $[\![\textbf{B}]\!]^{a,\mathcal{M}}$ and $[\![\textbf{C}]\!]^{a,\mathcal{M}}$ are both of type $\langle e,t\rangle$, then $[\![\textbf{A}]\!]^{a,\mathcal{M}} = [\lambda x \in D_e.\ [\![\textbf{B}]\!]^{a,\mathcal{M}}(x) = [\![\textbf{C}]\!]^{a,\mathcal{M}}(x) = 1]$ for some variable x of type e.

Predicate Abstraction (PA)
For any assignment a, for any model $\mathcal{M}$, and for any index $i \in \mathbb{N}$,

$$\left[\!\!\left[\begin{array}{c} \wedge \\ i \quad \textbf{B} \end{array} \right]\!\!\right]^{a,\mathcal{M}} = [\lambda x \in D_e.\ [\![\textbf{B}]\!]^{a[i\to x],\mathcal{M}}] \text{ for some variable } x \text{ of type } e.$$

- Predicate Modification accounts for NP-modification by adjectives, PPs, and relative clauses.
- Predicate Abstraction accounts for structures with movement.
- Assignments
  - Assignments are functions from indices (natural numbers) to individuals.
  - Assignment modification $a[i \to x]$ is the assignment that is just like a except that $a(i) = x$.
- For non-branching structures, we have the following three compositional rules.

Non-Branching Node Rule (NB)

For any assignment a and for any model $\mathcal{M}$, $\left[\!\!\left[\begin{array}{c} \textbf{A} \\ | \\ \textbf{B} \end{array} \right]\!\!\right]^{a,\mathcal{M}} = [\![\textbf{B}]\!]^{a,\mathcal{M}}$.

Lexical Terminal Node Rule (TN)
For any assignment a and for any model $\mathcal{M}$, if **A** is a terminal node and is not a trace or pronoun, then $[\![\mathbf{A}]\!]^{a,\mathcal{M}}$ is in the lexicon.

Traces and Pronouns Rule (TP)
For any assignment a and for any model $\mathcal{M}$, if **A** is a trace or a pronoun with an index $i \in \mathbb{N}$, then $[\![\mathbf{A}]\!]^{a,\mathcal{M}} = a(i)$.

Part I

Degree Semantics

Chapter 1

Vagueness and Degrees

1.1 Eubulides's Paradoxes: Neither True Nor False

- We have been assuming that declarative sentences denote 0 or 1 relative to a situation, i.e. they are either true or false with respect to a model.
- Such a semantics is called **bivalent** or two-valued semantics, and has been the standard assumption since Aristotle.
- However, natural language actually has a number of phenomena where bivalency seems to be not enough. We will discuss some of the problems of bivalent semantics pointed out by Eubulides. As we will see, these problems have not convinced everyone to renounce bivalency, however.
- Eubulides of Miletus (circa 405–330 BC) was a student of Euclid of Megara (not Euclid of Alexandria, the 'Father of Geometry'), who in turn was a student of Socrates. He was a contemporary of Aristotle.
- Eubulides is famous for his paradoxes that pose problems for Aristotle's logic, and natural language semantics in general (see Seuren 2005 for more on Eubulides' philosophy). It is not at all an exaggeration to say that much, if not all, of contemporary semantics still concerns these problems.
- The following four paradoxes are of particular interest for us, as they suggest that bivalency might not be enough for natural language semantics (1–3 are from Kneale & Kneale's 1962, cited in Seuren 2005; 4 is from Hyde 2014).
  1. **The Liar (*pseudomenos*) paradox**
     "A man says he is lying. Is what he says true or false?"
  2. **The Horns (*keratinês*) paradox**
     "What you have not lost you still have. But you have not lost your horns. So you still have horns."
  3. **The Bald Man (*phalakros*) paradox**
     "Would you say that a man was bald if he had only one hair? Yes. Would you say that a man was bald if he had only two hairs? Yes. Would you..., etc. Then where do you draw the line?"

4. **The Heap (*sôritês*) paradox**
   "Would you describe a single grain of wheat as a heap? No. Would you describe two grains of wheat as a heap? No. ... You must admit the presence of a heap sooner or later, so where do you draw the line?"

- These four paradoxes led some scholars to explore semantic systems with more than two truth-values, although whether such an elaborate system is necessary and/or adequate is controversial.
- The latter two paradoxes (the Bald Man and Heap paradoxes) relate to the main topic of the present chapter, but first, let us briefly touch on the first two.

1.1.1 The Liar Paradox

- The Liar paradox illustrates the problem of **self-reference**.
- Another version of the paradox: Consider the sentence in (1.1).

  (1.1) This sentence is false.

  The sentence (1.1) cannot be true because if it's true, it has to be false, which is a contradiction. But it cannot be false either, because if it's false it has to be true, which is also a contradiction. So this sentence is neither true nor false.
- One analysis of the Liar paradox postulates a third truth-value which represents the state of being neither true nor false. But there are a number of other theories of the Liar paradox that maintain bivalency.
- If you are interested in this topic, please see Beall & Glanzberg (2014) and references therein.
- The Liar paradox has been discussed almost exclusively by philosophers and has not been paid much attention in theoretical linguistics so far (which of course does not mean that it is not a problem for linguists!).

1.1.2 The Horns Paradox

- The Horns paradox has to do with the problem of **presupposition**.
- The paradox stems from the premise "What you have not lost you still have" which only means "What you *have had and have not lost* you still have". This implicit bit ("have had and") of the meaning of the word *lost* is called a presupposition.
- Unlike the problem of self-reference, presupposition has been extensively studied by both philosophers and linguistics, and has convinced many that bivalency has to be relaxed to some extent.[1]

[1] In a number of recent papers (Schlenker 2008, 2009, in particular), Philippe Schlenker has explored a possibility of keeping bivalency by delegating the issue of presupposition to pragmatics.

- One commonly accepted intuition behind the theories of presupposition is that sentences whose presuppositions are not met sound neither true nor false.[2] For example, if Nathan does not have a unicycle, (1.2) is definitely not true, but also is not straightforwardly false either.

  (1.2) Nathan lost his unicycle.

- This intuition led some scholars to postulate a third truth-value (Peters 1979, Beaver & Krahmer 2001, George 2008), and some others to postulate even a fourth one (Karttunen & Peters 1979, Cooper 1983, Sudo 2012). Such semantic systems are called trivalent.
- Another group of analyses make use of partial functions that can be undefined for certain inputs (Heim 1983, Beaver 2001), although this idea is closely related to trivalent theories. For historical reasons (largely due to Heim's 1983 influential work), an analysis using partial functions has been considered to be the standard today.
- For different theories of presuppositions, see the overview articles by Beaver & Geurts (2013) and Beaver (2001:Part I).

1.2 Vagueness: The Bald Man and Heap Paradoxes

- The Bald Man and Heap paradoxes illustrate the issue of **vagueness**.
  - **The Bald Man (*phalakros*) paradox**
    "Would you say that a man was bald if he had only one hair? Yes. Would you say that a man was bald if he had only two hairs? Yes. Would you..., etc. Then where do you draw the line?"
  - **The Heap (*sôritês*) paradox**
    "Would you describe a single grain of wheat as a heap? No. Would you describe two grains of wheat as a heap? No. ... You must admit the presence of a heap sooner or later, so where do you draw the line?"

  Side note: The term the **sorites paradox** (*sorites* means 'heap' in Classical Greek) is often used to refer to the Heap paradox and sometimes even to the kind of paradox involving vagueness. Also sometimes the Heap paradox talks about heaps of sand, rather than heaps of wheat.

- Predicates like *bald* and *heap* that give rise to paradoxes of the above kind are called **vague predicates**. One of the essential properties of vague predicates is

[2] Recent works such as von Fintel (2004) and Abrusán & Szendrői (2013) investigate the nature of this intuition, in an attempt to account for the fact that certain sentences seem to be more or less straightforwardly false, when the presupposition is false, e.g. (i).

(i) Daniel sold Nathan's unicycle to Jamie.

Relatedly, in my dissertation (Sudo 2012), I claimed that certain types of presupposition are completely independent from the truth and falsity of the sentence, an idea entertained also by Karttunen & Peters's (1979) seminal work.

that they are insensitive to small changes (e.g. one hair or one grain of wheat). This property is called **tolerance**.

- Another reincarnation of vagueness is **borderline cases**.
  - Some people are definitely bald, and some people are definitely not bald. But there are people who are kind of bald but not quite.
  - Similarly, there are things that are clearly heaps of wheat, and there are things that are clearly not heaps of wheat, e.g. two grains of wheat. But there are also things that are too small to be called heaps but are too big to be non-heaps.
- Besides *bald* and *heap*, vague predicates include *tall* (*tall* also gives rise to a paradox like the Bald Man and Heap paradoxes; see the Exercises at the end).
- *Tall* has borderline cases. For example (Nathan is about 183 cm):

  (1.3) Nathan is tall.

  If someone tells me (1.3), my initial reaction would be, "Well, it depends". It's not straightforwardly true, and it's also not straight forwardly false. So the predicate denoted by *tall* does not straightforwardly apply to Nathan.
- It seems that vague predicates generally exhibit **context-sensitivity** of a particular kind. That is, in some contexts, (1.3) is more or less true, e.g. if Nathan and I go to Tokyo together, my parents might truthfully say (1.3). But in other contexts, (1.3) becomes false, e.g. if Nathan quits linguistics and joins an NBA team, he'll be short. In yet other contexts it is neither straightforwardly true nor straightforwardly false, e.g. if we go to Amsterdam, he'll be about the average height for men. We'll come back to this aspect of meaning later on.

1.2.1 The Problem of Vagueness

- Can our semantic system account for the vagueness of predicates like *tall*? We have assumed so far that each one-place predicate denotes a function of type $\langle e,t \rangle$ which returns 0 or 1 for each individual, e.g.

  (1.4) For any model $\mathcal{M}$ and for any assignment a,
  a. $[\![\textbf{smokes}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ smokes in $\mathcal{M}$
  b. $[\![\textbf{student}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a student in $\mathcal{M}$
  c. $[\![\textbf{British}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is British in $\mathcal{M}$

  So if we extend this analysis to vague predicates like *tall*, we'll have something like (1.5) (recall this from Term 1).

  (1.5) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is tall in $\mathcal{M}$

  The problem of this denotation is that this is always either true or false for a given individual, which means that there *is* a clear boundary! So there is no place for borderline cases. Everyone is either tall or not tall. Similarly, everyone is either bald or not bald and everything is either a heap or not a heap.

- It is instructive to see the same thing with sets. Recall that type-$\langle e,t\rangle$ functions characterise sets of individuals. For example, the ones in (1.4) characterise the sets in (1.6), and (1.5) characterises (1.7).

(1.6) a. $\{ x \in D_e \mid x$ smokes in $\mathcal{M} \}$
b. $\{ x \in D_e \mid x$ is a student in $\mathcal{M} \}$
c. $\{ x \in D_e \mid x$ is British in $\mathcal{M} \}$

(1.7) $\{ x \in D_e \mid x$ is tall in $\mathcal{M} \}$

Since the sets clearly state which things belong to the set and which things don't in a given model, again, there is a clear boundary. So there is no room for vagueness.

- As the issue of vagueness is a deep one and has several facets, we cannot provide full answers to it in the present course. We will present one standard semantic basis for a theory of vagueness, called **degree semantics**. We will lay out the foundations of degree semantics in the next section, and come back to the issue of vagueness in the subsequent section.

1.2.2 Vagueness and Other Types of Uncertainty

- Before moving on, it is important to be clear about what vagueness is not.
- Ambiguity
  - An ambiguous sentence has two or more distinct meanings/truth-conditions.
  - Two types of ambiguity:[3]
    - Lexical ambiguity: The sentence is ambiguous due to a word with two different meanings.

      (1.8) Nathan went to the bank.

    - Structural ambiguity: The sentence is ambiguous due to

      (1.9) Daniel saw a man with binoculars.

      (Structural ambiguity might involve scope ambiguity, but this is a theoretically loaded position)
  - It should be easy to see that ambiguity and vagueness are different phenomenon.
    - One big difference is that ambiguity does not give rise to borderline cases. If Nathan went to the bank as a financial institution, (1.8) is true under one reading and false under the other. Similarly for (1.9).
    - VP-ellipsis test (originally pointed out by Lakoff 1970) for ambiguity: An elided expression needs to receive the same interpretation as its antecedent (or you have a rhetorical effect, called *syllepsis* or *zeugma*). For instance, (1.10a)

[3] There can be more kinds of ambiguity. For instance, as we will see later, plurality sometimes gives rise to ambiguity (with so-called *mixed predicates*). One might or might not want to call it a structural ambiguity. Another such example is sentences containing reciprocals (Winter 2001b, Dalrymple, Kanazawa, Kim, McHombo & Peters 1998). Also, under certain recent views of scalar implicatures (Chierchia 2006, Chierchia, Fox & Spector 2012, Fox 2006), sentences that can have scalar implicatures are ambiguous between readings with and without them.

is four-way ambiguous, due to the two-way ambiguity for each sentence. With VP-ellipsis, as in (1.10b), the sentence is only two-way ambiguous. That is, both sentences are about the financial institution or both are about the river embankment.

(1.10)a. Ad went to the bank. Hans did not go to the bank.
b. Ad went to the bank. Hans did not.

The following example illustrates the same point with a structural ambiguity (modelled after Kennedy's (2011)). Each sentence in (1.11a) is ambiguous between an interpretation where the subject is the agent of *eat* and one where it is the patient of *eat*. (1.11b) is, again, only two-way ambiguous.

(1.11) a. The fish is ready to eat, but the chicken isn't ready to eat.
b. The fish is ready to eat, but the chicken isn't.

Vagueness is not a matter of ambiguity, as evidenced by the fact that VP-ellipsis has no effects on it.

(1.12)a. Nathan is tall, but Daniel is not tall.
b. Nathan is tall, but Daniel is not.

- Imprecision:
  - More often than not, sentences like (1.13) are used imprecisely.

    (1.13)Nathan is 180 cm tall.

    This sentence is judged as true even if Nathan's height is not exactly 180.00 cm but reasonably close to it.
  - Lasersohn (1999) called such wiggle room for the truth-conditions **pragmatic slack**.
  - It's not that obvious that vagueness and imprecision are distinct phenomena, but there are several differences (for more on this topic, see Pinkal 1995, Lasersohn 1999, Kennedy & McNally 2005, Kennedy 2007, 2011, Sauerland & Stateva 2007, 2011, ooijROOIJVanvan Rooij 2011, Burnett 2012, Solt 2015).
    - Matushansky (2002) points out that *seem* (with a non-infinitival complement) is compatible with vague predicates but not with (merely) imprecise predicates.

      (1.14)a. Nathan seems tall.
      b. *Nathan seems 180 cm tall.
    - There are many expressions that either increase or decrease pragmatic slack (called *slack regulators*).

      (1.15)a. Nathan is precisely 180 cm tall.
      b. Nathan is approximately 180 cm tall.

    These do not work with vague predicates.[4]

[4] As Patrick asked in class, it is important to make sure that (1.16) is syntactically okay. Sacha's examples *precisely/approximately correct* suggest that *precisely* and *approximately* are indeed syntactically compatible with adjectives. Notice also that *approximately* can modify quantifiers, e.g. *approximately everyone*, but *precisely* cannot, **precisely everyone*, so there indeed are restrictions on these adverbs.

(1.16)a. *Nathan is precisely tall.
b. *Nathan is approximately tall.

◦ Certain types of imprecision are sensitive to 'round numbers', e.g. (1.17a) has more pragmatic slack than (1.17b) (Krifka 2002, 2007)

(1.17) a. Nathan is 180 cm tall.
b. Nathan is 181 cm tall.

The first sentence often judged as true even if Nathan is 181 cm, but the second sentence is judged as false if Nathan is 182 cm tall. Vagueness does not seem to be sensitive to this type of factor, e.g. *tall* remains equally vague for people who are 180 cm tall and people who are 181 cm tall.

1.3 Degree Semantics

- The idea of degree semantics is, as its name suggests, that sentences like (1.18) refers to entities called **degrees**.

(1.18)Nathan is tall.

According to degree semantics, (1.18) is true iff the degree to which Nathan is tall exceeds the 'standard degree of tallness'. So the meaning of the sentence refers to two degrees and says one is greater than the other.

- A number of questions arise immediately. First of all, what are degrees? This is a very important question, but for our purposes of giving model-theoretic denotations to sentences like (1.18), we do not need to make a lot of ontological commitments about degrees. We could treat them as just unanalysable primitives for the purposes of this course.[5]

- But there are certain crucial assumptions that we need to make about degrees. One is that they form **scales**.
  - A scale is a set of degrees that is totally ordered, i.e. any two degrees on a single scale are commensurable.
  - It is also often assumed that a scale is dense, i.e. between any two degrees on a single scale, there is a degree between them.

More formally,[6]

[5]One alternative is to define degrees as 'equivalent classes of individuals'. This idea is explored by Cresswell (1976). Since we already have individuals (and also sets thereof) in our model, under this conception of degrees, degrees come for free. A related idea due to Anderson & Morzycki (to appear) is that degrees are special types of entities called kinds. Kinds are needed for independent reasons (see papers in Carlson & Pelletier 1995 and Mari, Beyssade & Del Prete 2012), and their idea is to extend the ontology of kinds to include kinds of states, which for them are degrees. An interesting aspect of their theory that it accounts for the fact that across languages, it is common to use the same morphological items to talk about kinds and degrees (and also manners). They claim that this is because they all refer to kinds of some sort. Another approach to degrees championed by Moltmann (2007, 2009) says that degrees refer to so-called 'tropes'. This is ontologically more loaded, but Moltmann offers a number of arguments for postulating them.

[6]As Sacha pointed out in class, we want to use $<$ here, instead of $\leq$ to exclude trivial cases.

(1.19) A set S containing at least two degrees with an ordering relation $\prec$ is a scale iff

a. for all $d, d' \in S$, $d \prec d'$ or $d' \prec d$
(for any two degrees in S, either one is greater than the other)

b. for all $d, d' \in S$ such that $d \prec d'$, there is d'' such that $d \prec d'' \prec d'$
(For any two degrees in S, you can find another degree in between)

One might wonder if density is really needed. Indeed, it is not trivial to empirically motivate it.[7] But whether scales need to be dense or not, it is clear that we want them to be fine grained enough. For example, we don't want a scale with just two degrees. If we have density for every scale we can make sure that they are all fine-grained enough. We will assume density for the rest of this course.

- A degree semantic analysis of (1.18) refers to two degrees on the same scale, namely the scale of tallness.
  - One degree is Nathan's height. Let's call it d_N.
  - The other degree is a bit more abstract and is the standard degree of tallness. Let's call it d_s.

  The sentence (1.18) is true iff d_N exceeds d_s on the scale of tallness.

- With different predicates, different scales are involved.

  (1.20) Yasu is rich.

  This is true iff the degree to which Yasu is rich exceeds the standard of richness on the scale or richness.

  (1.21) Winston Churchill was bald.

  This is true iff the degree to which Winston Churchill was bald exceeds the standard of baldness on the scale of baldness.

- Thus the idea is that each scale is associated with some standard degree, and simple sentences like the ones above refer to it. We will discuss the nature of the standard in the next section, and how to derive these truth-conditions compositionally next week.

- Degree semantics is very useful in giving semantic analyses of various types of constructions involving gradable predicates.

- It is, however, not immediately clear whether sentences containing nouns like *heap* also refer to degrees. It is rather trivial to give a degree semantic analysis of these predicates, but the question is whether there is any evidence for it (we'll talk about adjectives next week). We will put aside this issue in this course, but it would be a very interesting topic to investigate (for your thesis, for example). See Morzycki (2014:§6.3.2) and references therein.[8]

[7] See Fox & Hackl (2006), Sauerland & Stateva (2007), Nouwen (2008), Abrusán & Spector (2011), Cummins, Sauerland & Solt (2012) for analyses of various linguistic phenomena that crucially make use of the density of degrees.

[8] As Laura pointed out, the fact that the nominal comparative construction like *more (of) a man* suggests that nouns should have a degree semantics. We'll talk about related points about adjectives next time.

1.4 Two Views on Vagueness

- Degree semantics itself does not provide a solution to the issues of vagueness. Recall that there are two important issues to solve.
  - How can we understand the Bald Man and Heap paradoxes (and their variants)?
  - How can we represent borderline cases in our model-theoretic semantics?
- In this section we discuss two popular views on how to think about vagueness. Both views claim that the Bald Man and Heap paradoxes are only appear to be paradoxical.
- Caveat: The discussion in rather open-ended, however. See the overview articles by Kennedy (2011) Sorensen (2013), eemterDEEMTERVanvan Deemter (2010), ooijROOIJVanvan Rooij (2011) and Solt (2015), and references there for other theories (there are many! E.g. fuzzy logic and supervaluation are especially well-discussed; Shapiro 2011 is an accessible overview article on vagueness in logic) and more in-depth discussion on pros and cons of each approach (you might find Van Rooij's chapter a little challenging).
- Caveat 2: The two approaches are inherently independent from degree semantics.

1.4.1 Epistemicism

- According to the degree semantic analysis of *bald*, there is still a clear boundary between people who are bald and people who are not bald. That is, those whose baldness exceeds the standard of baldness are bald and everybody else is not bald.
- A question that immediately arises is: What is the standard of baldness?
- The epistemic analysis of vagueness championed by Williamson (1994, 1997) states that the issues of vagueness stem from our ignorance about the standards of vague predicates. Thus, there is actually a sharp boundary for a vague predicate but it is not known to us and in fact is unknowable.
- This might sound unrealistic, but there is something intuitive about it, too. When someone says *Winston Churchill is bald* and we agree with them, none of us indeed cannot specify the standard of baldness explicitly.
- Under this view, the paradoxes cease to be problematic, since the inductive hypothesis that if a man with n hairs is bald, a man with $n + 1$ hairs is also bald is false for a particular n. That is, although we do not know where the standard lies exactly, but there is one. So at some point, we have a situation where a man with $s - 1$ hairs is bald, but a man with s hairs is not bald.
- Borderline cases can be thought of as uncertainty regarding the standard. For extreme cases we are pretty sure that the subject is above or below the standard, but as one gets closer to the standard, it gets exceedingly hard to be sure.

1.4.2 Contextualism

- Raffman (1994, 1996) and Fara (2000) put forward a different view, according to which the inherent context sensitivity of vague predicates gives rise to the apparent paradoxes (other proponents include Kamp 1981 and Shapiro (2006) among many others).
- As mentioned above, what counts as the standard of a give vague predicate changes in different contexts. The idea of contextualism is that the paradoxes involves a series of changes in contexts, but the induction hypothesis is not true with respect to a single context.
- Generally speaking, you do not distinguish the baldness of a man with n hairs and the baldness of a man with $n+1$ hairs, for any number n in a single context ('tolerance'). But this does not guarantee that you do not distinguish them from the baldness of a man with $n+1000$ hairs *in the same context*. Indeed, you do not distinguish the baldness of a man with $n+1$ hairs and the baldness of a man with $n+2$ hairs, and you might not also distinguish them from the baldness of a man with n hairs, but importantly, you are now in a different context with different sets of men with different numbers of hairs. And if you repeat this process, you might end up in a context where you do not distinguish the baldness of a man with $n+9999$ hairs and the baldness of a man with $n+10000$ hairs, but you do distinguish them from the baldness of a man with n hairs.
- The idea of contextualism can be phrased as follows in the framework of degree semantics: The standard of a vague predicate changes according to the context, e.g. according to which two objects you are comparing.
- However, this view does not straightforwardly provide a way to represent borderline cases. Generally speaking, one can have a vague predicate with borderline cases within a particular context. For example, in a context where you compare the heights of five different men, the guy in the middle is neither tall nor not tall. Thus, it needs to be augmented with a way to represent borderline cases.
- It would take us too far to do so in precise terms here, but it should be pointed out that a mere trivalent semantics will not be enough. That is, one could give up on bivalency, and assign the following kind of meaning to vague predicates.

  $$(1.22)\llbracket \textbf{tall} \rrbracket^{a,\mathcal{M}} = \lambda x \in D_e. \begin{cases} 0 & x \text{ is not tall in } \mathcal{M} \\ 1 & x \text{ is tall in } \mathcal{M} \\ 2 & x \text{ is neither tall nor not tall in } \mathcal{M} \end{cases}$$

  This function classifies the individuals into three groups:

  - Those who are assigned 0 are not-tall people.
  - Those who are assigned 1 are tall people.
  - Those who are assigned 2 are borderline cases.
- This treatment is definitely a first step toward a solution but is insufficient in itself. The reason is because we now have clear boundaries between the three groups, while intuitively, what counts as a borderline case should be vague itself.

This vagueness of borderline cases is called **second-order vagueness**. And one can actually talk about vagueness for second-order vagueness, i.e. third-order vagueness, and so on (so, there shouldn't be a clear boundary at any level). The cover term for all of these is **higher-order vagueness**.

- One widely entertained way to deal with higher-order vagueness is *supervaluation semantics* developed by Fine (1975). We will not go into it here.

1.5 Further Readings

The early proponents of degree semantics include: Seuren (1973), Cresswell (1976), von Stechow (1984), Heim (1985), Bierwisch (1989), followed by many such as Kennedy (1999, 2007).

There is a major alternative called *supervaluation semantics* or sometimes *delineation semantics* proposed by Kamp (1975), Klein (1980, 1991) that does not postulate degrees. But it turns out that degree semantics is more useful for purposes of analysing various constructions in natural language, as we will demonstrate later in the course. But this of course does not mean that supervaluation/delineation semantics is wrong. See in particular Doetjes, Constantinescu & Součková (2011) for recent discussion on this.

For vagueness, the three overview articles mentioned above (Kennedy 2011, Sorensen 2013, Van Rooij 2011) are good places to start. The literature is copious and spans across philosophy, logic, and linguistics. What is mentioned here is only the first step towards this interesting issue.

For pragmatic slack, see Lasersohn (1999) and for pragmatic effects of round numbers, see Krifka (2002, 2007).

1.6 Exercises

i) *Tall* is a vague predicate. Construct a paradox like the Bald Man and Heap paradoxes using *tall*.

ii) The versions of the Bald Man and Heap paradoxes involve adding a thing at a time, a hair or grain of wheat. We construct different versions of the paradoxes by subtracting a thing at a time, e.g.

   - *Bald*: Suppose there is a man with lots of hair. He is not bald. Plucking one hair from his head will not make him bald. Plucking two hairs from his head will not make him bald. Plucking three hairs from his head will not make him bald. ... So if plucking n hairs from his head will not make him bald, then plucking $n + 1$ will not make him bald. But for some number k, plucking k hairs from his head should make him bald!!
   - *Heap*: Take a heap of wheat. After removing one grain of wheat, you still have a heap of wheat. After removing one more grain of wheat, you still

> have a heap of wheat. ... So removing one grain at a time will keep the heap. But at some point, the heap should cease to be a heap!

Come up with one vague preposition and one vague verb, and construct paradoxes like these ones.

Chapter 2

Compositional Semantics with Degrees

2.1 Gradable Predicates

- Review of last week:
  - Vague predicates are those predicates that give rise to a 'sorites paradox'.
  - Vague predicates have borderline cases.
  - Degree semantic analysis of vague predicates: e.g., (2.1) is true iff the degree to which Nathan is tall exceeds (or is equal to) the standard of tallness on the scale of tallness.

    (2.1) Nathan is tall.

- Vague predicates are generally **gradable predicates**, i.e. predicates that refer to extents to which things satisfy the predicate.
- We'll mostly discuss gradable adjectives such as *tall*, which most of the research on this topic focuses on.
- Gradable adjectives can be modified by a number of adverbs called **degree modifiers**:
  1. Gradable adjectives take the comparative form with *-er/more*.

     (2.2) Nathan is **taller** than Yasu is.

  2. Gradable adjectives can be intensified with adverbs like *very*, *extremely*, *totally*, *so*, *too*, etc.

     (2.3) a. Nathan is **very tall**.
     b. Nathan is **too tall**.

  3. Gradable adjectives combine with *how* to form questions.

     (2.4) **How tall** is Nathan?

- Compare these to non-gradable adjectives, e.g. *prime*:

(2.5) a. 3 is prime.
b. #3 is more prime than 4.
c. #3 is very prime.
d. #3 is too prime.
e. #How prime is 3?

- NB: Non-gradable adjectives can sometimes be 'coerced' to gradable ones.

(2.6) a. Eric is more Japanese than Yasu is.
b. Eric is very Japanese.
c. Eric is too Japanese.

(Question: Are there non-gradable adjectives that have no gradable uses?)

- The unmodified form, as in (2.7), is called the **positive form**.

(2.7) Nathan is tall.

- Some more gradable adjectives:

(2.8) *Open*
a. This door is more open than that one.
b. This door is very open.
c. This door is too open.
d. How open is that door?

(2.9) *Dangerous*
a. This area is more dangerous than that area.
b. This are is very dangerous.
c. This area is too dangerous.
d. How dangerous is this area?

- As we will see, the modifiers like *-er/more*, *very*, *how*, etc. can be given compositional semantic analyses in degree semantics.

2.2 The Scale Structures of Gradable Adjectives

- There are a number of degree modifiers, but some of them can only modify a subset of gradable adjectives (an observation originally due to Bolinger 1972).
- For example, *completely* can modify *open*, but not *tall*.

(2.10) a. *Nathan is completely tall.
b. The door is completely open.

- Kennedy & McNally (2005) and Rotstein & Winter (2004) claim that the acceptability of these modifiers tracks the scale structure of the gradable adjective.
- We assume that gradable adjectives are associated with a scale, e.g. *tall* is associated with a scale of tallness.[1]

[1]Certain adjectives are compatible with multiple scales. For example, *long* can be about spatial

- Recall: a scale is a (uncountably infinite) set S of degrees with an ordering relation $\prec$ such that:
  - for any two distinct degrees $d, d' \in S$, $d \prec d'$ or $d' \prec d$; and (totality)
  - for any two distinct degrees $d, d' \in S$ such that $d \prec d'$, there is another degree $d'' \in S$ such that $d \prec d'' \prec d'$. (density)
- The definition does not say anything about the end points. There might or might not be minimal and maximal elements.
- Four possible scale structures:
  1. **Totally open scale**: e.g. *tall*
     A scale without a lower or upper limit.
     Isomorphic to $(0, 1)$, i.e. $\{ r \in \mathbb{R} \mid 0 < r < 1 \}$ ○————○
  2. **Totally closed scale**: e.g. *open*
     A scale with both lower and upper limit.
     Isomorphic to $[0, 1]$, i.e. $\{ r \in \mathbb{R} \mid 0 \leqslant r \leqslant 1 \}$ •————•
  3. **Upper closed scale**: e.g. *safe*
     A scale with only the upper limit.
     Isomorphic to $(0, 1]$, i.e. $\{ r \in \mathbb{R} \mid 0 < r \leqslant 1 \}$: ○————•
  4. **Lower closed scale**: e.g. *dangerous*
     A scale with only the lower limit.
     Isomorphic to $[0, 1)$, i.e. $\{ r \in \mathbb{R} \mid 0 \leqslant r < 1 \}$: •————○

  (• means that the end point is on the scale. ○ means that there is no end point)
- *Completely* can only modify gradable adjectives with scales that have a maximum degree, i.e. totally closed and upper closed scales.[2]
  - (2.11) means that the degree to which the door is open is the maximal degree. This is fine since the scale of openness (for a door) has a maximal degree.

    (2.11) The door is completely open.

  - Because the scale of tallness has no maximal degree, (2.12) is unacceptable.

    (2.12) *Nathan is completely tall.

  The following contrast indicates that *safe* has a maximal degree, while *dangerous* doesn't.

  (2.13) a. This area is completely safe.
  b. *This area is completely dangerous.

- *Slightly* can only modify gradable adjectives with scales that have a minimum value, i.e. 2 and 4 above.

extension or temporal extension. But this is largely a type of lexical ambiguity. More interesting are those gradable adjectives that are, most often than not, about multiple scales at the same time, e.g. *intelligent*, *healthy*, *similar*, *talented*, *normal*. See Bierwisch (1988, 1989), Kamp (1975), Klein (1980), Sassoon (2013), and Morzycki (2014:§3.7.3) for more on such multidimensional adjectives.

[2] Modifiers like *completely* and *totally* have *intensificational uses*, which are similar to *very*. This seems to be a common phenomenon cross-linguistically, but why this is so is not very well understood (see works by Magda Kaufmann, Eric McCready, Ryan Bochnak, and Andrea Beltrama for some ideas).

(2.14) a. *Nathan is slightly tall.
b. The door is slightly open.
c. *This area is slightly safe.
d. This area is slightly dangerous.

This is because *slightly* means the relevant degree is a bit above the minimum degree of the scale.[3]

- According to (2.14), the scale of *tall* has no lower bound, but you might wonder about the status of 0 cm. We need to assume that it is not on the scale of tallness.
- *Half* can only modify gradable adjectives with totally closed scales, i.e. 2.

(2.15) a. *Nathan is half tall.
b. The door is half open.
c. *This area is half safe.
d. *This area is half dangerous.

This is because *half* expresses a proportion, and proportions only make sense on a closed scale.

- We'll analyse the meanings of these adverbs in Section 2.5.

2.3 Relative vs. Absolute Adjectives

- Kennedy & McNally (2005) and Kennedy (2007) make an important point about the relation between the scale structure and the interpretation of the positive form (i.e. the unmarked form).
  - Totally open scales give rise to vagueness.
  - The other three types of scales have non-vague uses.

  Kennedy & McNally call adjectives with totally open scales **relative adjectives** and adjectives with scales closed at least on one end **absolute adjectives**.
- We discussed a couple of examples of relative adjectives last time.

(2.16) Nathan is tall.

- Here are some examples of absolute adjectives.

(2.17) *Totally closed scales*
a. The door is open.
b. The door is closed.
c. This restaurant is empty.
d. This restaurant is full.

(2.18) *Upper closed scales*

[3] But Nouwen 2011 remarks that *slightly* with a totally closed scale is often degraded, e.g. ??*The hard drive is slightly full.* See also Sassoon (2012) and Sassoon & Zevakhina (2012) for quantitative data, and an alternative analysis where *slightly* is analysed as a 'granularity shifter'.

a. This area is safe.
b. This rod is straight.

(2.19) *Lower closed scales*
a. This area is dangerous.
b. This rod is bent.

One important characteristic of absolute adjectives that sets them apart from relative adjectives is that their standards are relatively clear.

- The door is open iff there is an aperture.
- The door is closed iff there is no aperture.
- The rod is straight iff it is parallel to a line.
- The rod is bent iff it is not parallel to a line.

The clear standards make these adjectives non-vague.

- But this does not mean that the standard is context insensitive. In order to see this, consider (2.20) with a gradable adjective with a totally closed scale (see McNally 2011 for related discussion).

(2.20) a. This beer glass is full.
b. This wine glass is full.

Depending on the subject, the standard is taken to be different (even if you have the same glass). For beer, *full* means full to the rim, but for wine, the *full* means something like 1/2 of the glass or even less.

- Also, the standards for *safe* and *dangerous* are less clear. It is in fact not impossible to construct a sorites paradox and borderline cases for these adjectives, e.g.:

(2.21) An area with zero criminals and 10,000 inhabitants (in Tokyo) is safe. Add one yakuza. It is still safe. Add another yakuza. It is still safe. Add yet another yakuza. It is still safe. So adding one yakuza at a time doesn't seem to make the area not safe. But at some point, the area should become dangerous!! Where is the boundary?

- Furthermore, it is possible to construct a sorites paradox and borderline cases for any of the above absolute adjectives. Here is a case of *empty* (which has a totally closed scale).

(2.22) A sushi restaurant with 50 seats is empty when there is 0 customer. Add one customer. It is still empty. Add another customer. It is still empty. So adding one customer at a time doesn't make the restaurant not empty. But at some point, it should become empty. When is it?

- Kennedy (2007) claims that these vague uses of absolute adjectives are due to *imprecision*. Recall that language can be used imprecisely (see also Pinkal 1995).

(2.23) a. Nathan is 180 cm tall.
b. Everyone is asleep in this town.

And imprecision generally gives rise to vagueness: It's easy to see that the sentences in (2.23) have borderline cases.

(recall from last week that it is not uncontroversial that these two things are distinct phenomena or that they are unrelated; see Fara 2000, Burnett 2012, 2014 and Solt 2015 for discussion).

- So according to Kennedy & McNally, what differentiates relative and absolute adjectives is that the latter, but not the former, *can* be used precisely.

(2.24) a. Precisely speaking, the restaurant is empty.
b. #Precisely speaking, Nathan is tall.

And when used precisely, the positive forms of absolute adjectives do not give rise to vagueness.

- It is also remarkable that the standards of absolute adjectives are always taken to be a closed end-point of the scale (modulo imprecision) (Kennedy & McNally 2005). So the following generalisations hold.
  - For fully open scales, the standard is contextually determined.
  - For scales closed on one end, the standard is the sole closed point.
  - For fully closed scales, the standard is either one of the end points (the minimum for *awake* and *open*; the maximum for *asleep* and *straight*).

For fully closed scales, there doesn't seem to be a general rule for which end-point is taken to be the standard. If so, this information needs to be somehow encoded in the lexical entry of each gradable adjective with a fully closed scale. But see Kennedy & McNally (2005) and Kennedy (2007) for a theory that predicts the scale structure and the standard for de-verbal adjectives.

- Kennedy (2007) proposes that the above generalisations are due to the fact that there is a principle that favours more contribution of the conventional meanings.

(2.25) *Interpretive Economy*: Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions. (Kennedy 2007:49)

The idea is that that the end-point of a scale is a lexical aspect of the adjective (it's specified as an end-point), while a mid-point is not. So you prefer truth-conditions that refer to an end-point, whenever possible (but see McNally (2011) for criticisms).

2.4 Antonyms

- *Tall* and *short* are related in that when one of them is true, the other one cannot true. For example, the sentences in (2.26) cannot be simultaneously true (although they can be simultaneously not true; such pairs are called *contraries*):

(2.26) a. Daniel is tall.
b. Daniel is short.

Pairs of gradable adjectives like these are called **antonyms**.

- It is often the case that one of the antonym pairs is 'marked' in the following sense (Seuren 1978, Kennedy 2001, Sassoon 2010, Morzycki 2014, Rett 2014).
  - The marked one does not combine with measure phrases in the positive form.

(2.27) a. Daniel is 180 cm tall.
b. *Daniel is 180 cm short.

NB: Many gradable adjectives do not combine with measure phrases:[4]

(2.28) a. *Daniel is 70 kg heavy.
b. *Daniel is 70 kg light.

(2.29) a. *The train is 80 km/h fast.
b. *The train is 80 km/h slow.

And for some antonym pairs, both gradable adjectives are compatible with measure phrases.

(2.30) a. The clock is 3 min late.
b. The clock is 3 min early.

Side remark: Interestingly, languages differ here. In German and Mandarin Chinese, all of the unmarked members of (2.27)-(2.29), as well as (2.30), are acceptable. In Dutch, *fast* but not *heavy* is compatible with measure phrases. In Russian none of the above examples are acceptable. In Japanese, (2.27)-(2.29) are unacceptable, while (2.30) is acceptable. However, all of these languages allow 'differential measure phrases' in comparatives (e.g. *Nathan is 3 cm taller than Yasu (is)*).
Here are some data from Schwarzschild (2005:210):

(2.31) pesante quasi due tonnellante
heavy almost two tons
'weighs almost two tons' Italian (attributed to Zamparelli 2000)

(2.32) is 2 boeken rijk
is 2 books rich
'owns two books' Dutch (attributed to Corver 1990)

See Schwarzschild (2005), Winter (2005) and Breakstone (2012) for discussion on this.

- Marked members resist nominalisation :

(2.33) *Long-short*
a. length
b. #shortness

(2.34) *Wide-narrow*
a. Width
b. #Narrowness

[4] As Laura pointed out after class, 3 *min fast/slow* is acceptable. This suggests that the restrictions are semantic in nature, rather than morpho-syntactic.

- The equative form of the marked member entails the positive form for both things that are compared (a property sometimes called *evaluativity* or more literally *positive-entailingness*).

(2.35) Ad is as tall as Andrew.
  a. $\nRightarrow$ Ad is tall.
  b. $\nRightarrow$ Andrew is tall.

(2.36) Ad is as short as Andrew.
  a. $\Rightarrow$ Ad is short.
  b. $\Rightarrow$ Andrew is short.

- Similarly, in a *how*-question, the marked member gives rise to an entailment to the positive form.

(2.37) How tall is Hans?
  $\nRightarrow$ Hans is tall.

(2.38) How short is Hans?
  $\Rightarrow$ Hans is short.

- Generally, antonyms have the same kind of scale structure.
  - If one member of the pair has a totally open scale, the other one does too, e.g. *tall-short*.
  - If one member of the pair has a totally closed scale, the other one does too, e.g. *full-empty*.
  - If one member of the pair has an upper closed scale, the other one has a lower closed scale, e.g. *wet-dry*.
  - If one member of the pair has a lower closed scale, the other one has an upper closed scale, *wet-dry*.
- This suggests that antonyms involve the same scale with the opposite ordering relation. More precisely:
  - Suppose a gradable adjective A is associated with a scale $S_A = \langle D, \prec_A \rangle$.
  - Then its antonym B is associated with a scale $S_B = \langle D, \prec_B \rangle$ such that for any $d, d' \in D$, $d \prec_A d'$ iff $d' \prec_B d$.

2.5 Compositional Semantics with Degrees

- We are now ready to give a compositional semantics for gradable adjectives.
- The key idea is that they refer to **degrees**.
- This requires the model to be enriched with degrees.
  - A model $\mathcal{M}$ is a triple $\langle \mathcal{D}, \mathcal{S}, \mathcal{I} \rangle$ such that
  - $\mathcal{D}$ is a set of individuals; and
  - $\mathcal{I}$ is an interpretation function.

– $\mathcal{S}$ is a set of scales;
– Each scale $s \in S$ is a pair, $\langle G_s, <_s \rangle$ that forms a scale according to the definition of scales in Section 2.2; and
– The intersection of all such G's is the set of degrees.

- We also introduce a new type d for degrees.
- There are several approaches, but for the moment, we consider one of the standard approaches that analyses gradable adjectives as relations between individuals and degrees, i.e. they are functions of type $\langle d, \langle e, t \rangle \rangle$ (Cresswell 1976, von Stechow 1984, among many others).

(2.39) For any model $\mathcal{M}$ and for any assignment a,
$$[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d . \lambda x \in D_e . \left[\begin{array}{l} \text{the degree to which } x \text{ is tall in } \mathcal{M} \text{ is greater} \\ \text{than or equal to } d \text{ on the scale of tallness in } \mathcal{M} \end{array} \right]$$

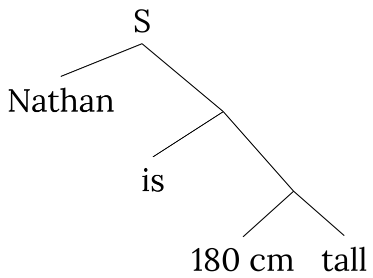
This is often written as (2.40) in the literature:

(2.40) For any model $\mathcal{M}$ and for any assignment a,
$$[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d . \lambda x \in D_e .\ x \text{ is } d\text{-tall in } \mathcal{M}$$

Or sometimes:

(2.41) For any model $\mathcal{M}$ and for any assignment a,
$$[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d . \lambda x \in D_e .\ \textbf{height}(x) \geq_{\text{tall}} d$$

- A degree modifier is analysed as operating on the d-slot of gradable adjectives. The simplest case is this:

(2.42)


We assume that 180 *cm* is a 'proper name' for a degree (following Klein 1980; but see Schwarzschild 2005 for a different view which analyses measure phrases as predicates of intervals), i.e.

(2.43) For any model $\mathcal{M}$ and assignment a,
$[\![\textbf{180 cm}]\!]^{a,\mathcal{M}}$ = the degree of 180 cm

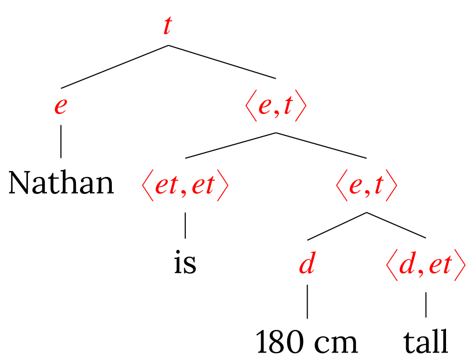
This is of type d, so it can combine with *tall* to yield the following function of type $\langle e, t \rangle$.

(2.44) For any model $\mathcal{M}$ and assignment a,
$$[\![\textbf{180 cm tall}]\!]^{a,\mathcal{M}} = [\![\textbf{tall}]\!]^{a,\mathcal{M}}([\![\textbf{180 cm}]\!]^{a,\mathcal{M}}) = \lambda x \in D_e .\ x \text{ is 180 cm tall in } \mathcal{M}$$

Consequently, (2.42) is true iff Nathan is 180 cm tall in $\mathcal{M}$.

The types of each subtree is annotated in the following diagram:

(2.45)



- One important caveat here is that (2.45) is an '*at-least* interpretation', i.e. the sentence is true iff Nathan is 180 cm *or higher* in $\mathcal{M}$.
  - This is because the semantics of $[\![\textbf{tall}]\!]^{a,\mathcal{M}}$ is inherently downward monotonic with respect to the ordering of the scale of tallness $\prec_{\text{tall}}$. That is, is Nathan is d-tall in $\mathcal{M}$, then for any smaller degree $d' \prec_{\text{tall}} d$, Nathan is d'-tall in $\mathcal{M}$ as well.
  - You might think this is inadequate given that (2.42) sounds like saying that Nathan is exactly 6' tall (modulo imprecision), rather than at least 6' tall.
  - A way out here is that the exactly reading of (2.42) is a scalar implicature, generated in competition with *Nathan is 181 cm tall*, *Nathan is 182 cm tall*, etc.
- Remaining issue: Why can't all gradable adjectives combine with measure phrases (in English)? And why is there cross-linguistic variation?

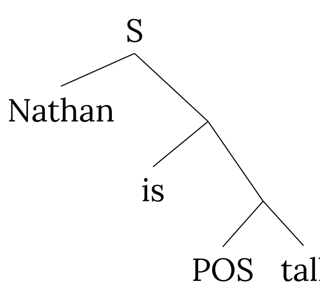
(2.46)a. Nathan is 180 cm tall.
b. *Nathan is 180 cm short.

(2.47)a. *This water is 40°C hot.
b. *This water is 40°C cold.

2.5.1 Positive Form

- What about the positive sentence in (2.48), which does not seem to involve a degree modifier? – There is an indivisible degree modifier called POS (Bartsch & Vennemann 1975, Cresswell 1976, von Stechow 1984, Kennedy 1999).[5]
- The structure of positive sentences looks like (2.48).

(2.48)



[5]This is arguably a foible of this approach. As far as I know, no language has an overt realisation of POS. There are theories that basically encode it in the meaning of the adjective itself, such as Klein (1980, 1991), Doetjes et al. (2011), but we will not discuss these alternatives here for reasons of time.

Recall that our analysis of this positive sentence refers to the standard degree of tallness, i.e. (2.48) is true iff the degree to which Nathan is tall is greater than the standard of tallness. How is the standard determined?

- We know that the standard shift in different contexts. For example, if we are in Tokyo, the standard of tallness is lower than if we are in Amsterdam. Given that the model is meant to represent a conversational situation against which the truth of a given sentence is assessed, we can have the model set the standard in the following way:

(2.49) $[\![\textbf{POS tall}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$'s tallness $\mathcal{M} \geq_{\text{tall}}$ the standard of tallness in $\mathcal{M}$

- However, there is reason to believe that this is not the right approach. The same adjective can have different standards with different subjects, even if they are evaluated in the same context.

(2.50) Klaus is tall. Martina (Klaus's 8-year old daughter) is tall, too.

The standard of tallness for Klaus is evidently much higher than for Martina, his daughter, because Martina is compared to other 8-year olds. So we do not want to fix the standard once and for all within a model.

- A more promising idea is that POS implicitly refers to a set of things that are compared, which is called a **comparison set**. For Klaus, it's taken to be the set of men, and for Martina, it's the set of 8-year-olds (or 8-year-old girls).
- We discussed essentially the same idea in Term 1, when we talked about the seeming non-intersectivity of adjectives like *tall*, *big* and *expensive*. For example, *tall* seems to be sensitive to which noun it modifies in the sense that the standard is clearly different in the following two cases.

(2.51) a. tall man
b. tall 8-year-old

For (2.51a), the comparison set is the set of (relevant) men, and for (2.51b), it is the set of (relevant) 8-year-olds.

We also noted, however, that the noun does not determine the comparison set, although it is usually taken to be the default comparison set. Rather, this default choice can be overridden by contextual factors, e.g.

(2.52) a. Klaus built a tall snowman.
b. Martina built a tall snowman.

For (2.52a), the default reading is salient, but for (2.52b), a reading is easily available where the snowman's height is compared to kids.

- Furthermore, it is possible to specify the comparison set with phrases like *for*-phrases (*for*-phrases in English have several different functions; see Bylinina 2013):

(2.53) a. Yasu is tall for a Japanese.
b. Martina is tall for a 8-year-old.

- So we assume that POS refers to a comparison set, either implicitly or explicitly (by a *for*-phrase). For the implicit case, let's suppose that there is a null pronominal element C denoting a type-$\langle e,t\rangle$ function that characterises the comparison set.
  - C is a pronoun, so according to our analysis of pronouns, it should have an index.
  - But it denotes a function of type $\langle e,t\rangle$, rather than an individual.
  - We have been assuming that a is a function whose domain is natural numbers $\mathbb{N}$ and whose range is the set of individuals, but from now on, we assume that a maps a pair of a number and type to an element of that type (this idea is discussed in Heim & Kratzer 1998 too).

    (2.54)For any model $\mathcal{M}$ and for any assignment a,
    $[\![\mathbf{he}_{\langle 23,e\rangle}]\!]^{a,\mathcal{M}} = a(\langle 23,e\rangle)$, which is a member of D_e

    Now, the null pronominal C has an index of the form $\langle n,\langle e,t\rangle\rangle$.

    (2.55)For any model $\mathcal{M}$ and for any assignment a,
    $[\![C_{\langle 11,\langle e,t\rangle\rangle}]\!]^{a,\mathcal{M}} = a(\langle 11,\langle e,t\rangle\rangle)$, which is a member of $D_{\langle e,t\rangle}$

- Using C, we analyse POS as follows:

  (2.56)For any model $\mathcal{M}$, for any assignment a, and for any natural number n,
  $$[\![[\mathbf{POS}\ C_{\langle n,\langle e,t\rangle\rangle}]\ \mathbf{tall}]\!]^{a,\mathcal{M}} = \lambda x \in D_e. \left[\begin{array}{l} x\text{'s tallness in } \mathcal{M} \succeq_{\text{tall}} \text{the standard of tallness} \\ \text{for } \{\, y \in D_e \mid a(\langle n,\langle e,t\rangle\rangle)(y) = 1 \,\} \text{ in } \mathcal{M} \end{array}\right]$$

  - We don't say much about how the value of POS is determine (and similarly we don't say how the value of a (free) pronoun is determined). This belongs to pragmatics.
  - Assuming appropriate values for the two comparison sets, we can analyse (2.50):

    (2.57)Klaus is POS $C_{\langle 18,\langle e,t\rangle\rangle}$ tall. Martina is POS $C_{\langle 2,\langle e,t\rangle\rangle}$.

    For example,

    (2.58)a. $[\![C_{\langle 18,\langle e,t\rangle\rangle}]\!]^{a,\mathcal{M}}$
    $= a(\langle 18,\langle e,t\rangle\rangle)$
    $= \lambda x \in D_e.\ x$ is a grown-up man who lives in London in $\mathcal{M}$

    b. $[\![C_{\langle 2,\langle e,t\rangle\rangle}]\!]^{a,\mathcal{M}}$
    $= a(\langle 2,\langle e,t\rangle\rangle)$
    $= \lambda x \in D_e.\ x$ is a 8 year-old who lives in London in $\mathcal{M}$

- Given the above analysis of the comparison set, the semantics of POS should look like (2.59). The first argument f denotes the comparison set, the second argument G is the denotation of the gradable adjective.

  (2.59)For any model $\mathcal{M}$ and for any assignment a,
  $[\![\mathbf{POS}]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle e,t\rangle}.\lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.$

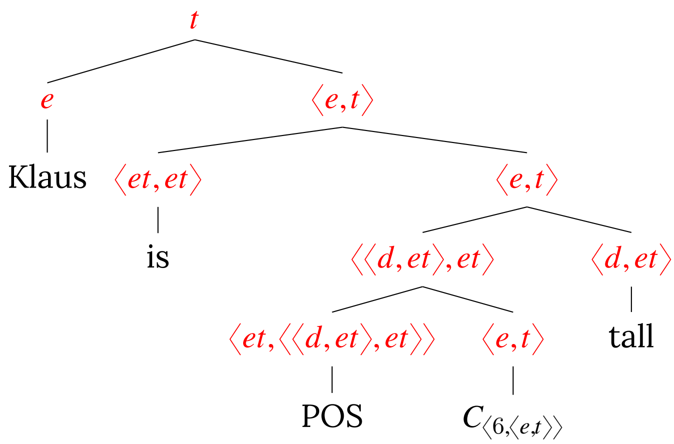
for the standard degree d_s of the scale associated with G with respect to the comparison set $\{ y \in D_e \mid f(y) = 1 \}$, $G(d_s)(x) = 1$
(i.e. the degree to which x is G exceeds d_s)

This is a bit too long, so we denote 'the standard degree d_s of the scale associated with G with respect to the comparison set $\{ y \in D_e \mid f(y) = 1 \}$' by $\textbf{standard}(G)(f)$ (which is a degree), which allows us to simplify (2.59) as (2.60):

(2.60) For any model $\mathcal{M}$ and for any assignment a,
$$[\![\textbf{POS}]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle e,t\rangle}.\lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\ G\big(\textbf{standard}(G)(f)\big)(x) = 1$$

The types look like (2.61).

(2.61)



Notice that POS takes C and also the gradable adjective *tall*. This is because POS needs to know which scale the standard should be on, and in order to do so, it has to take *tall* as its argument.

- For absolute adjectives, the standard is an end-point. But recall that they are still context-sensitive, as illustrated by (2.20). This can be understood as due to different comparison sets: The comparison set is the set of different states of the glass in question, and untypical cases, i.e. a completely full wine glass, may be contextually excluded (see Toledo & Sassoon 2011 for related ideas).

(2.20) a. This beer glass is full.
b. This wine glass is full.

- There is one more important remark about the standards of scales with maximum degrees. According to our analysis, for *a is G* to be true, the degree to which a is G exceeds or is equal to the standard of G (with respect to some comparison set). But if the standard is the minimum degree, which is the case for adjectives like *open* and *bent*), there will be a problem: the sentence is predicted to be trivial. So we assume that the standard for such adjectives is a bit above the minimum.[6]

- The *for*-phrase can be analysed as an overt version of C, i.e. it denotes a type-$\langle e,t\rangle$ function, e.g.

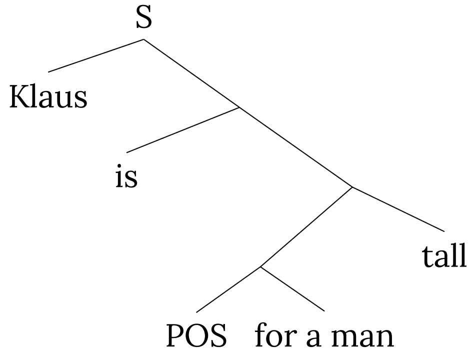
[6] But there is something unsatisfying about this idea. See Kennedy & McNally (2005) and Kennedy (2007) for deeper discussion on this point.

(2.62) $[\![\textbf{for a man}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a man in $\mathcal{M}$

But it needs to occupy a position different from the surface position.

(2.63) a. Klaus is tall for a man.

b.



We could assume that *for a man* undergoes obligatory 'extraposition' to the right periphery at some point in the derivation, but this movement has no semantic consequences. We also do not go into the internal composition of the *for*-phrase. See Fults (2006), Bylinina (2013) for more on this.

2.5.2 Other Degree Modifiers

- *Very* can be seen as a **standard booster**, i.e. it shifts the standard to a higher degree.
- We could just say that it appears instead of POS and adds some extra degree to the standard:

(2.64) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{very}\ C_{\langle 2,\langle e,t\rangle\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ the degree to which x is tall exceeds $\textbf{standard}([\![\textbf{tall}]\!]^{a,\mathcal{M}})(a(\langle 2,\langle e,t\rangle\rangle)) + \delta$
(where $d + \delta$ is some degree that is higher on the scale than d)

- Klein (1980) suggests a very interesting alternative idea: *very* operates on the comparison set, namely it requires the members of the comparison set to satisfy the positive form. The idea is that the comparison set of *very tall* consists of individuals that are tall.

(2.65) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{very}\ C_{\langle 2,\langle e,t\rangle\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ for the standard d_s of the scale of tallness with respect to $\{\, y \in D_e \mid [\![\textbf{POS}\ C_{\langle 2,\langle e,t\rangle\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}(y) = 1 \,\}$, $[\![\textbf{tall}]\!]^{a,\mathcal{M}}(d_s)(x) = 1$

See Kennedy & McNally (2005) and Kennedy (2007) for discussion on the restrictions on the use of *very*.

- *Completely* refers to the maximum degree on the scale.

(2.66) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{completely}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.$ for the maximum degree $d_{\max}$ on the scale associated with G, $G(d_{\max})(x) = 1$

We assume that if G's scale does not have a maximum degree, the result is infelicitous (which you can say is due to a presupposition of *completely*).

- *Slightly*, on the other hand, refers to the minimum degree on the scale.

  (2.67) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{slightly}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.$ for the minimum degree d_{min} on the scale associated with G, $G(d_{\text{min}} + \varepsilon)(x) = 1$

  Again, if the scale does not have a minimum degree, it is infelicitous.

- A proportional degree modifier like *half* refers to both the maximum and minimum degrees.

  (2.68) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{half}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.$ for the mid-point d between the maximal degree d_{max} and the minimal degree d_{min} on the scale associated with G, $G(d)(x) = 1$

2.6 Further Readings

There is a recent surge in the literature on the topic of the semantics of gradable adjectives, including a number of interesting experimental works. It is impossible to list all individual papers here (nor do I know all of them), but papers by Galit Sassoon and Stephanie Solt are particularly relevant.

Toledo & Sassoon (2011) claim that the distinction between relative vs. absolute adjectives is a matter of how comparison classes are determined. Burnett (2012, 2014) proposes a similar idea. Their basic idea is that relative adjectives have *extensional* comparison classes, while absolute adjectives have *intensional* comparison classes. Sassoon (2012) and Sassoon & Zevakhina (2012) put forward an alternative analysis of modifiers like *slightly* and *completely* where they are analysed in terms of precision.

There is some recent psycholinguistic work. For the acquisition of gradable predicates, see Barner & Snedeker (2008), Syrett (2007), Syrett, Kennedy & Lidz (2010) and Tribushinina (2013).

2.7 Exercises

1. Find two adjectives that are *not* gradable (or have non-gradable uses). Provide a few examples to support your answer.
2. Classify the following adjectives with respect to their scale structure. For each adjective, support your answer with a few examples.

   a) famous   b) deep   c) certain   d) familiar

3. For each of the following degree modifiers, discuss what kind of distributional restrictions it has, by raising concrete examples. Also discuss whether the restrictions can be stated in terms of scale structure.

   a) perfectly   b) somewhat   c) extremely

4. Come up with one gradable adjective whose scale is totally closed and whose standard is the minimum degree. Come up with one gradable adjective whose scale is totally closed and whose standard is the maximal degree.

Chapter 3

Comparative Constructions

3.1 Comparative Constructions

- Comparative constructions in English and other languages are well studied in degree semantics. We mostly focus on English but crosslinguistic variation is very interesting in this domain (see Section 3.6 for some discussion).
- There are two main types of comparative sentences in English:

(3.1) a. Nathan is taller than Daniel. (**Phrasal comparative**)
b. Nathan is taller than Daniel is. (**Clausal comparative**)

A phrasal comparative involves a DP (or some other non-clausal material) as the complement of *than*, while a clausal comparative involves something that looks like a clause. Notice that (3.1b) has a missing item after *is* in the *than*-clause. In this example, this seems to be (almost) obligatory.

(3.2) *Nathan is taller than Daniel is tall.

But the following is fine:

(3.3) This desk is wider than the bed is long.

One way to understand (3.1b) is that it is underlyingly (3.2) but undergoes obligatory ellipsis of the adjective (see Kennedy 1999 for a different analysis that does not postulate ellipsis but an invisible operator). According to this analysis, (3.2) and (3.3) are structurally isomorphic.

- There is a lot of debate in the literature about whether phrasal and clausal comparative are syntactically related.
  - Phrasal comparatives are underlyingly clausal but just have more missing parts (Bresnan 1973, Hackl 2000, Lechner 2001, 2004, 2008, Bhatt & Takahashi 2011)
  - Phrasal comparatives cannot be reduced to phrasal comparatives (Hankamer 1973, Hoeksema 1983, Pinkal 1990, Kennedy 1999, Pancheva 2006)
- Some arguments for the existence of phrasal comparatives:
  - Accusative case:

(3.4) a. Nathan is taller than her.
b. *Nathan is taller than her is.

– Anaphor binding:

(3.5) a. No one is taller than himself.
b. *No one is taller than himself is.

– Wh-movement:

(3.6) a. Who is Nathan taller than *t*?
b. *Who is Nathan taller than *t* is?

– Scopal difference:

(3.7) a. Nathan is taller than nobody.
b. *Nathan is taller than nobody is.

(Why (3.7b) is bad is an interesting question. We'll come back to this next week.)

– There are languages that seem to only have phrasal comparatives, as we will see in Section 3.6.

These differences between phrasal and clausal comparatives are unexpected if phrasal comparatives are underlyingly clausal.

- Although the debate is not settled completely yet, we'll develop separate analyses for phrasal and clausal comparatives.
- Digression: The following type of sentence can be used to talk about comparisons but they need not involve comparative forms of the gradable adjectives (and the comparative version degrades somewhat).

(3.8) a. Compared to Andrew, Nathan is tall.
b. ??Compared to Andrew, Nathan is taller.

This construction is different from canonical comparatives in that it exhibits vagueness, as illustrated by the following example from Kennedy (2010) cited in Nouwen (2011) (Some facts: the radius of Uranus is 25,362 km, the radius of Venus is 6,052 km, and the radius of Neptune is 24,622 km).

(3.9) a. Uranus is big, compared to Venus.
b. Uranus is bigger than Venus.

(3.10) a. #Uranus is big, compared to Neptune.
b. Uranus is bigger than Neptune.

For the semantics of this construction, see Beck, Oda & Sugisaki (2004), Kennedy (2010) and Fults (2006, 2010).

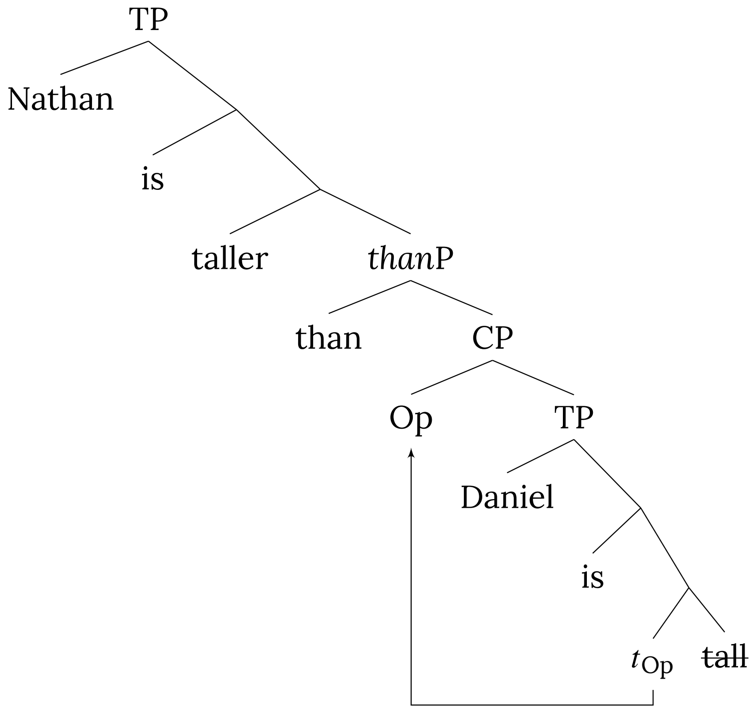
3.2 The Syntax of Clausal Comparatives

- Let us analyse the following simple sentence:

(3.11) Nathan is taller than Daniel is.

- The standard analysis of clausal comparatives postulates two phonologically null items in the *than*-clause:
  - An invisible occurrence of the gradable adjective *tall*
  - An operator-movement. Let's call this operator *Op*

(3.12)



- In the so-called 'subcomparative deletion' construction, there is no invisible adjective:

  (3.13) The desk is wider than the bed is long.

  The standard analysis says (3.11) and (3.13) have isomorphic structures, and the semantics works in exactly the same way.

- Evidence for the operator movement:
  - In some languages you see a wh-phrase:

    (3.14) Ja lublju Ivana bol'še **čem** [jego ljubit Maša].
    I love Ivan.acc more what.instr [him loves Masha.nom]
    'I love Ivan more than Masha does.' Russian (Pancheva 2007)

  - The operator-movement is island sensitive in the same way as *wh*-movement and other A-bar movements (Bresnan 1973, 1975, Chomsky 1977), although there are some exceptions. As a baseline, (3.15) shows that both the operator movement and *wh*-movement are unbounded.

    (3.15) a. Which language does Jamie think that Daniel speaks *t*?
    b. Nathan is taller than Jamie thinks Daniel is *t* ~~tall~~.

    The following show that these movements are sensitive to the same island constraints.

(3.16) *Complex NP Island*
a. *Which language did Jamie meet [a man who speaks t]?
b. *Nathan is taller than Jamie met [a man who is t ~~tall~~].

(3.17) *Adjunct Island*
a. *Which language will Ad be excited [if someone speaks t]?
b. *Nathan is smarter than Ad will be excited [if someone is t ~~smart~~].

But there is one crucial difference: The operator movement violates the so-called the Left-Branch Condition.

(3.18) *How is Daniel t tall?

See Kennedy & Merchant (2000) for more on this.

3.3 The Semantics of Clausal Comparatives

- Take a gradable adjective and combine it with *more* or *-er*, whichever is appropriate. The resulting comparative adjective is generally not vague, even if the positive form is vague.

  (3.19) a. Nathan is tall.
  b. Nathan is taller than Daniel is.

  NB: (3.19b) has a 'imprecise' use, perhaps unexpectedly: You might say it's false if Nathan is 182.5 cm tall and Daniel is 182 cm tall. But whenever it is used precisely (which you can force to some extend by using phrases like *strictly speaking*), (3.19b) is not vague.

- Recall our analysis of (3.19a):

  (3.20) Nathan is [POS C] tall.

  (3.20) is true iff the degree to which Nathan is tall is greater than or equal to the standard of tallness with respect to the degree C on the scale of tallness.

- We analyse the truth-conditions of (3.19b) to be (following Seuren 1973, 1984, Gajewski 2008, Schwarzschild 2008):[1]

  (3.21) There is a degree to which Nathan is tall and to which Daniel is not tall.

- We will assume the same type-$\langle d, et\rangle$ semantics for *tall*.

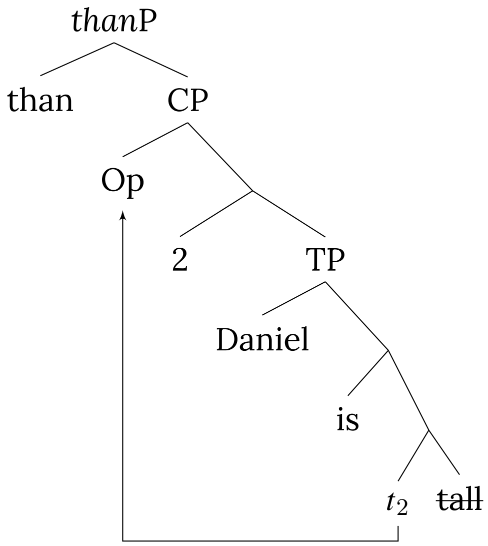
  (3.22) $[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.\lambda x \in D_e.\ x$ is d-tall in $\mathcal{M}$

[1] A popular alternative analysis says (3.21) is true iff the degree to which Nathan is tall exceeds the degree to which Daniel is tall (Cresswell 1976, von Stechow 1984, Heim 1985, Rullmann 1995, Schwarzschild & Wilkinson 2002, Heim 2006). We will not take this approach for technical reasons (in our ontology of degrees, it is not easy to refer to the maximal degree to which Daniel is tall in a compositional fashion).

3.3.1 ‘Than’-Clause

- Let’s start with the *than*-clause. We assume the structure in (3.23).

(3.23)



- We analyse (3.23) to be denoting a function of type $\langle d,t\rangle$. Specifically:

(3.24) $[\![\textbf{than Op Daniel } t_{\textbf{Op}} \textbf{ tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.$ Daniel is d-tall in $\mathcal{M}$

Recall that ‘Daniel is d-tall in $\mathcal{M}$’ means the degree to which Daniel is tall is equal to or exceeds d. So if Daniel is 180 cm tall, it maps any degree on the tallness scale that is equal to or smaller than 180 cm to 1.

- How do we derive this compositionally? We have:

(3.25) For any model $\mathcal{M}$, and any assignment a,

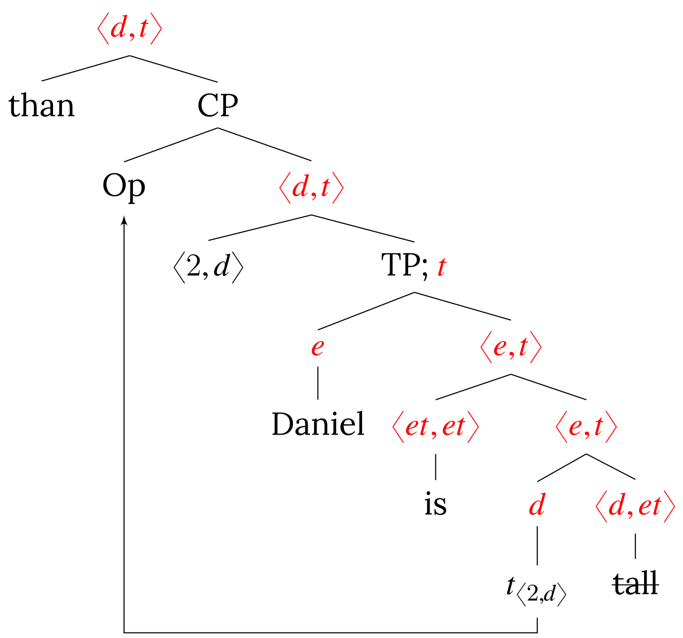
a. $[\![\textbf{Daniel}]\!]^{a,\mathcal{M}} =$ Daniel

b. $[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.\lambda x \in D_e.\ x$ is d-tall in $\mathcal{M}$

c. $[\![\textbf{is}]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle e,t\rangle}.\ f$

We want the trace of the operator to be of type d. Then we’ll have the following types (recall our new convention: an index is of the form $\langle n,\tau\rangle$ where n is a number and τ is a type).

(3.26)



- Here's a new version of Predicate Abstraction that refers to complex indices.

(3.27) **Predicate Abstraction (PA)**
For any assignment a, for any model $\mathcal{M}$, and for any index $\langle n,\tau\rangle \in \mathbb{N} \times \text{Type}$ (where Type is the set of types),

$$\left[\!\!\left[\ [\ \langle n,\tau\rangle \quad \mathbf{A}\ ]\ \right]\!\!\right]^{a,\mathcal{M}} = \left[\lambda x \in D_\tau.\ [\![\mathbf{A}]\!]^{a[\langle n,\tau\rangle \to x],\mathcal{M}}\right]$$

for some variable x of type τ.

- Then the type $\langle d,t\rangle$-subtree denotes the following function:

(3.28) $[\![\langle 2,d\rangle$ **[Daniel is [**$t_{\langle 2,d\rangle}$ **tall]]**$]\!]^{a,\mathcal{M}}$
$= \lambda d \in D_d.\ [\![$**Daniel is [**$t_{\langle 2,d\rangle}$ **tall]**$]\!]^{a[\langle 2,d\rangle \to d],\mathcal{M}}$
$= \lambda d \in D_d.\ [\![$**tall**$]\!]^{a[\langle 2,d\rangle \to d],\mathcal{M}}(d)(\text{Daniel})$
$= \lambda d \in D_d.$ Daniel is d-tall in $\mathcal{M}$

This is what we want. So let's assume that *Op* and *than* are semantically vacuous, i.e. they denote identity functions.

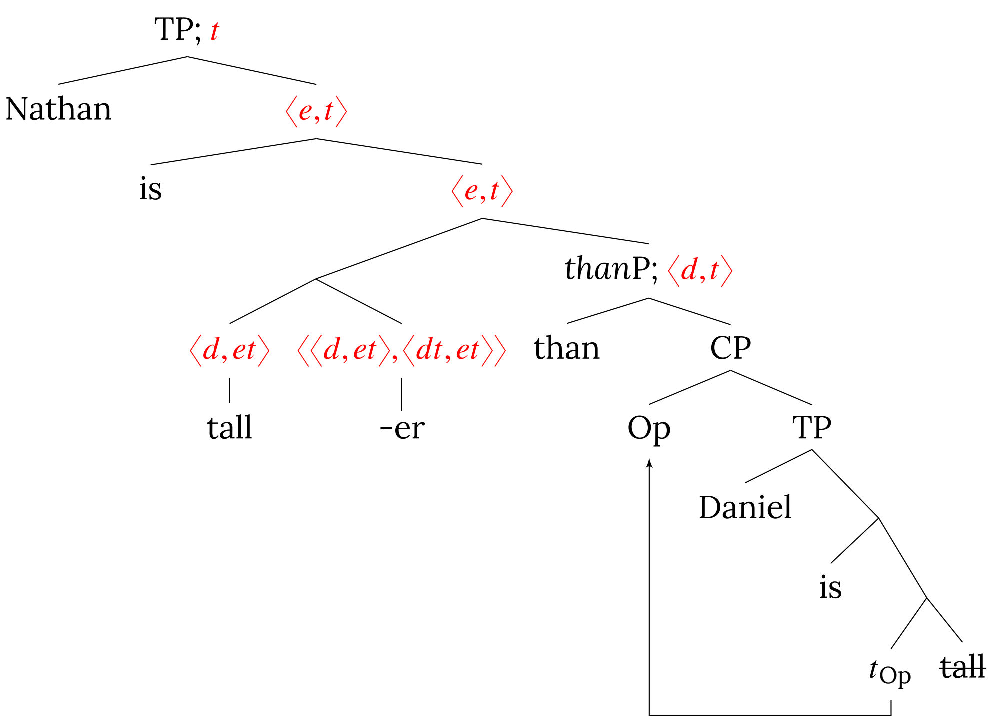
(3.29) For any model $\mathcal{M}$ and for any assignment a,
$[\![$**Op**$]\!]^{a,\mathcal{M}} = [\![$**than**$]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle d,t\rangle}.\ f$

Then we have (3.24), i.e. the *than*-clause characterises the set of degrees to which Daniel is tall.

3.3.2 Matrix Clause

- The types will work out as follows.

(3.30)



We want (3.30) to be true iff there is a degree d such that Nathan is d-tall but Daniel is not d-tall.

- In order to derive these truth-conditions, we analyse the comparative morpheme *-er* (and *more*) to be an existential quantifier over degrees plus negation for the *than*-clause.

(3.31) For any model $\mathcal{M}$ and for any assignment a,

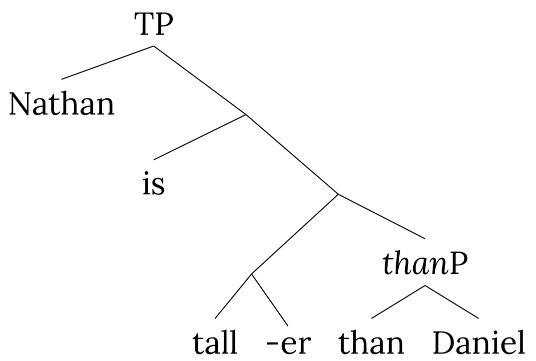
$$[\![\textbf{-er}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e.\left[\begin{array}{l}\text{there is a degree } d \in D_d\\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 0\end{array}\right]$$

- Bottom-up computation:
  - $[\![\textbf{tall -er}]\!]^{a,\mathcal{M}} = [\![\textbf{-er}]\!]^{a,\mathcal{M}}([\![\textbf{tall}]\!]^{a,\mathcal{M}})$
    $= \lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e.\left[\begin{array}{l}\text{there is a degree } d \in D_d\\ \text{such that } [\![\textbf{tall}]\!]^{a,\mathcal{M}}(d)(x) = 1 \text{ and } f(d) = 0\end{array}\right]$
  - $[\![\textbf{than Op } \langle 2,d\rangle \textbf{ Daniel } t_{\langle 2,d\rangle} \textbf{ tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.$ Daniel is d-tall in $\mathcal{M}$
  - $[\![\textbf{[tall -er] [than Op } \langle 2,d\rangle \textbf{ Daniel } t_{\langle 2,d\rangle} \textbf{ tall]]}]\!]^{a,\mathcal{M}}$
    $= \lambda x \in D_e.\left[\begin{array}{l}\text{there is a degree } d \text{ such that}\\ [\![\textbf{tall}]\!]^{a,\mathcal{M}}(d)(x) = 1 \text{ and Daniel is not } d\text{-tall in } \mathcal{M}\end{array}\right]$
    $= [\![\textbf{is [tall-er [than Op } \langle 2,d\rangle \textbf{ Daniel } t_{\langle 2,d\rangle} \textbf{ tall]]}]\!]^{a,\mathcal{M}}$
  - $[\![\textbf{Nathan is taller than Op } \langle 2,d\rangle \textbf{ Daniel } t_{\langle 2,d\rangle} \textbf{ tall}]\!]^{a,\mathcal{M}} = 1$ iff
    there is a degree d such that Nathan is d-tall in $\mathcal{M}$ and Daniel is not d-tall in $\mathcal{M}$

3.4 Phrasal Comparatives

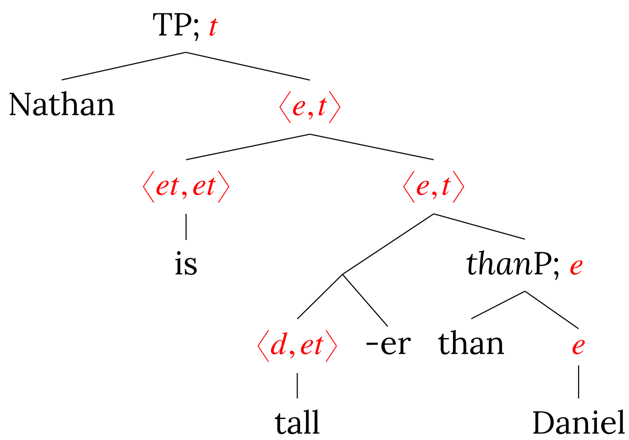
- We assume the following syntax for phrasal comparatives:

(3.32)



- Assuming again that *than* is semantically vacuous:

(3.33)



(3.34) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{than}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$$

- This requires the type of *-er/more* to be $\langle\langle d,et\rangle,\langle e,et\rangle\rangle$. So $[\![\textbf{-er}]\!]^{a,\mathcal{M}}$ is going to look like:

(3.35) $$[\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\lambda y \in D_e.\cdots$$

What is '$\cdots$'? The idea is the same as before: there is a degree d such that the matrix subject y is d-much G but the *than*-phrase x is not. So,

(3.36) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\lambda y \in D_e.\left[\begin{array}{l}\text{there is a degree } d \text{ such that}\\ G(d)(y) = 1 \text{ and } G(d)(x) = 0\end{array}\right]$$

So *taller* means:

(3.37) $$[\![\textbf{tall-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}}$$

$$= [\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}}([\![\textbf{tall}]\!]^{a,\mathcal{M}})$$

$$= \lambda x \in D_e.\lambda y \in D_e.\left[\begin{array}{l}\text{there is a degree } d \text{ such that}\\ [\![\textbf{tall}]\!]^{a,\mathcal{M}}(d)(y) = 1 \text{ and } [\![\textbf{tall}]\!]^{a,\mathcal{M}}(d)(x) = 0\end{array}\right]$$

(Notice that $[\![\textbf{tall}]\!]^{a,\mathcal{M}}$ is used twice: once for the matrix subject, once for the *than*-phrase).

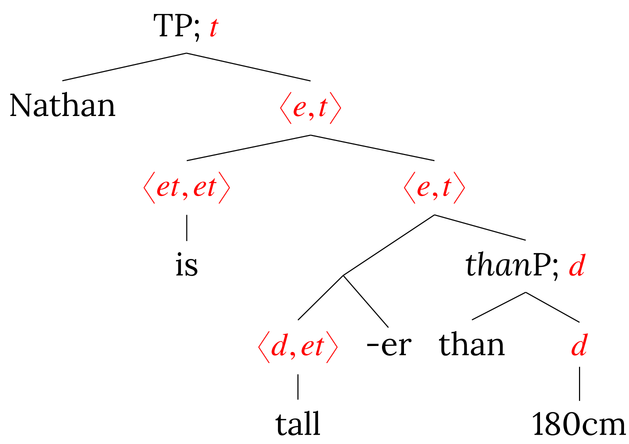
- Sample top-down computation (some steps are omitted):

$$
\begin{aligned}
&[\![\textbf{Nathan is [tall -er}_{\textbf{phrasal}}\textbf{] than Daniel}]\!]^{a,\mathcal{M}}\\
=&[\![\textbf{is [tall -er}_{\textbf{phrasal}}\textbf{] than Daniel}]\!]^{a,\mathcal{M}}(\text{Nathan})\\
=&[\![\textbf{[tall -er}_{\textbf{phrasal}}\textbf{] than Daniel}]\!]^{a,\mathcal{M}}(\text{Nathan})\\
=&[\![\textbf{tall -er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}}([\![\textbf{than Daniel}]\!]^{a,\mathcal{M}})(\text{Nathan})\\
=&[\![\textbf{tall -er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}}(\text{Daniel})(\text{Nathan})\\
=&1 \text{iff there is a degree } d \text{ such that Nathan is } d\text{-tall and Daniel is not } d\text{-tall in } \mathcal{M}
\end{aligned}
$$

- For phrasal comparatives involving a degree *than*-phrase such as (3.38), we need a slightly different semantics for *-er* and *than*.

(3.38) Nathan is taller than 180 cm.

(3.39)



(3.40) For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{-er}_{\textbf{deg.phrasal}}]\!]^{a,\mathcal{M}}$

$$= \lambda G \in D_{\langle d,et\rangle}.\lambda d \in D_d.\lambda y \in D_e.\left[\begin{array}{l}\text{there is a degree } d' \text{ such that}\\ G(d')(y) = 1 \text{ and } d \prec_{s(G)} d'\end{array}\right]$$

(where $s(G)$ is the scale associated with the gradable adjective G)

b. $[\![\textbf{than}_{\textbf{deg.phrasal}}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.\ d$

3.5 Comparatives with Multiple Adjectives

- Comparative sentences involving two gradable adjectives like (3.41) are called **subcomparatives**.

(3.41) The desk is wider than the bed is long.

Our analysis of clausal comparatives naturally extends to (3.41).

(3.42) There is a degree d such that the desk is d-wide and the bed is not d-long.

- Subcomparatives are only felicitous with ‘commensurable scales’:

(3.43) #Nathan is smarter than the desk is wide.

As is intuitively the case, the scale of smartness and the scale of width cannot be directly compared (but see below for metalinguistic comparatives).

This restriction, however, is not directly predicted by our analysis. Assuming that degrees of smartness and degrees of width are not on the same scale, (3.43) will be trivially true:

(3.44) There is a degree d such that Nathan is d-smart and the desk is not d-wide.

We can think of the restriction as a ‘presupposition’ that degrees d existentially quantified by the comparative morpheme needs to be on both of the scales. This rules out trivial cases like (3.43).

- Furthermore, subcomparatives are infelicitous with certain antonyms, a phenomenon Kennedy (1999) dubbed **cross-polar anomaly**.

(3.45) a. *Daniel is shorter than Nathan is tall.
b. *Nathan is taller than Daniel is short.

The unacceptability of these sentences is not predicted by our analysis. See Kennedy (1999, 2001) and Büring (2007) for analyses.

- **Comparatives of deviation** compares deviations from the standards. It is a feature of this construction that only analytic comparatives (*more* + adjective) give rise to this reading.

(3.46) a. San Francisco Bay is more shallow than Monterey Bay is deep.
b. *San Francisco Bay is shallower than Monterey Bay is deep. (Morzycki 2014:175)

See Kennedy (1999) and Morzycki (2014:§4.3.7) for discussion.

- **Metalinguistic comparatives** compare the ‘appropriateness’ of the words.

(3.47) a. It’s more chilly than cold.
b. Nathan is more a semanticist than a philosopher.

Metalinguistic comparatives are never possible with analytic comparatives.

(3.48) a. George is more dumb than crazy.
b. *George is dumber than crazy. (Morzycki 2014:172)

See Morzycki (2011) for an analysis.

- Another type of inter-adjective comparison is **indirect comparison** (which Luke asked about in class). The following are examples due to Bale (2006) cited by Morzycki (2014:176) (Esme and Seymour are Alan Bale’s children).

(3.49) a. Let me tell you how pretty Esme is. She’s prettier than Einstein was clever.

b. Although Seymour was both happy and angry, he was still happier than he was angry.
c. Seymour is taller for a man than he is wide for a man.

Unlike comparison of deviation and metalingusitic comparatives, indirect comparisons are possible with synthetic comparatives (Adj+-*er*).

See Bale (2008) and other works cited in Morzycki (2014:§4.3.8).

3.6 Crosslinguistic Variation

3.6.1 Languages that Distinguish Clausal and Phrasal Comparatives

Some languages draw morphological distinctions between phrasal and clausal comparatives.

(3.50)*Greek* (adapted from Merchant 2009:135f)

a. I Maria pezi kithara kalitera **apo** [ton Gianni].
the.nom Maria plays guitar better than.phr [the.acc Giannis]
'Maria plays the guitar better than Giannis.' (Phrasal)

b. I Maria pezi kithara kalitera **ap'oti** [pezi kithara o Giannis].
the.nom Maria plays guitar better than.cl [plays guitar the.nom Giannis]
'Maria plays the guitar better than Giannis does.' (Clausal)

(3.51)*Russian* (Pancheva 2007)

a. Ja lublju Ivana bol'še [Maši].
I love Ivan.acc more [Masha.gen]
'I love Ivan more than Masha.' (Phrasal)

b. Ja lublju Ivana bol'še **čem** [jego ljubit Maša].
I love Ivan.acc more what-instr [him loves Masha.nom]
'I love Ivan more than Masha does.' (Clausal)

(3.52)*Hungarian* (Wunderlich 2001)

a. Anna érdekes-ebb volt [Péter-nél].
Anna interesting-more was [Peter-adess]
'Anna was more interesting than Peter.' (Phrasal)

b. Anna érdekes-ebb, **mint** [a-milyen érdekes Péter volt].
Anna interesting-more than [rel-what.kind interesting Peter was]
'Anna is more interesting than Peter was.' (Clausal)

However, in these languages at least, the difference appears only on *than*, while in our analysis, there are two differences, *than* and *-er*:

(3.53)For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{than}_{\textbf{clausal}}]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle d,t\rangle}.\ f$

b. $[\![\textbf{than}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$

c. $[\![\textbf{than}_{\textbf{deg.phrasal}}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.\ d$

(3.54)For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e. \begin{bmatrix} \text{there is a degree } d \in D_d \\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 0 \end{bmatrix}$

b. $[\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\lambda y \in D_e. \begin{bmatrix} \text{there is a degree } d \text{ such that} \\ G(d)(y) = 1 \text{ and } G(d)(x) = 0 \end{bmatrix}$

c. $[\![\textbf{-er}_{\textbf{deg.phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda d \in D_d.\lambda y \in D_e. \begin{bmatrix} \text{there is a degree } d' \text{ such that} \\ G(d')(y) = 1 \text{ and } d \prec_{s(G)} d' \end{bmatrix}$

Also, there does not seem to be a language that morphologically distinguishes individual-phrasal and degree phrasal comparatives.

3.6.2 Languages without Clausal Comparatives

Hindi-Urdu (Bhatt & Takahashi 2011) and Turkish (Hofstetter 2009) are considered to only have phrasal comparatives. Korean seems to lack clausal comparatives, too. The basic ingredients of Korean comparatives are:

(3.55)-*pota* ‘than’
te ‘more’
tel ‘less’

The comparative marker *te* is optional. Here's a grammatical example with a phrasal comparative:

(3.56)John-un [Bill]-pota (te) ttokttokha-ta.
John-top [Bill]-than (more) smart-decl
‘John is smarter than Bill’

A simple clausal comparative is ungrammatical.

(3.57)*John-un [Bill-i (ttokttokha-ta)]-pota (te) ttokttokha-ta.
John-top [Bill-nom smart-decl]-pota (more) smart-decl
‘(intended) John is smarter than Bill is’

In complex comparatives, the use of the ‘nominaliser’ *kes* is obligatory, in which case the embedded predicate takes an adnominal suffix (glossed as ‘rel’ here).

(3.58)a. John-un [Bill-i ttokttokha-n **kes**]-pota te ttokttokha-ta.
John-top [Bill-nom smart-rel kes]-than more smart-decl
‘John is smarter than Bill is’

b. Bill-un [John-i [*pro* ttokttokha-ta ko] sayngkak-hayss-ten **kes**
Bill-top [John-nom [*pro* smart-decl C] think-past-rel kes
]-pota te ttokttokha-ta
]-than more smart-decl

'Bill is smarter than John thought he was'

So generally, *pota* selects for nominal complements.

There is one remaining puzzle, however: Subcomparatives are generally unacceptable, even with *kes*:

(3.59) John-un [i chimtay-ka ki-n kes]-pota (khi-ka) (te)
John-top [this bed-nom long-rel kes]-than (height-nom) (more)
khu-ta.
big-decl
'(intended) John is taller than this bed is long'

But this is not syntactically bad, because the sentence becomes acceptable if the two adjectives are identical.

(3.60) a. John-un [Bill-i ttokttokha-n kes]-pota te ttokttokha-ta.
John-top [Bill-nom smart-rel kes]-than more smart-decl
'John is smarter than Bill is'

b. John-un [Bill-i khi-ka khu-n kes]-pota te khi-ka
John-top [Bill-nom height-nom big-rel kes]-than more height-nom
khu-ta.
big-decl
'John is taller than Bill is'

Japanese is sometimes also considered a language without clausal comparatives (Beck et al. 2004, Oda 2008, Kennedy 2009, Sudo 2014), but this claim is not uncontroversial (Hayashishita 2009, Shimoyama 2012).

3.6.3 Languages without Comparative Constructions

There are languages that seem to lack dedicated comparatives altogether. However, it is not the case that these languages cannot express comparison. A common strategy in these languages is to use so-called *conjoined comparative*. Here are some data taken from Morzycki (2014:181).

(3.61) jo i ben, jo eu nag.
house this big house that small
'This house is bigger than that house.' Amele (Roberts 1987: 135)

(3.62) Tata'hkes-ew, nenah teh kan.
strong-3sg I and not
'He is stronger than me.' Menomini (Bloomfield 1962: 506)

Here's some more data from Beck, Krasikova, Fleischer, Gergel, Hofstetter, Savelsberg, Vanderelst & Villalta (2009:18–).

(3.63) Mary na lata, to Frank na kwadogi.
Mary top tall, but Frank top short
'Mary is taller than Frank.' Motu

Another strategy is to use a verb that means something akin to *exceed*, as demonstrated by the following Thai example from (Morzycki 2014:182).

(3.64) kǎw sǔung kwaà kon túk kon.
he tall exceed man each man
'He is taller than anyone.' Thai (Warotamasikkhadit 1972: 71)

Here is an example from Beck et al. (2009:21).

See Stassen (1984, 1985, 2006) and Beck et al. (2009) for more on the typology of comparative constructions.

3.7 Further Readings

There is a lot of work on comparatives in degree semantics. Classical works include: Seuren (1973), Cresswell (1976), Seuren (1984), von Stechow (1984), Heim (1985). Bresnan (1973) and Bresnan (1975) are the first papers on the syntax of comparatives. Schwarzschild (2008), Beck (2011), and Morzycki (2014:Ch.4) are accessible introductions to the degree semantic analysis of comparatives and related constructions.

There are several alternative analyses of comparatives. As mentioned in fn.1, a popular approach deploys the maximality-operator, and analyses *-er* as expressing the 'greater-than' relation (Cresswell 1976, von Stechow 1984, Heim 1985, Rullmann 1995, Schwarzschild & Wilkinson 2002, Heim 2006). For example, for *Nathan is taller than Daniel is*, the matrix clause denotes Nathan's maximal height, the *than*-clause denotes Daniel's maximal height, and the sentence says the former exceeds the latter.

Kennedy (1999) develops an analysis where gradable adjectives are analysed as type-$\langle e,d\rangle$ functions, rather than type-$\langle d,et\rangle$ functions.

Klein (1980, 1982, 1991) and Doetjes et al. (2011) develop analyses of comparatives using the delineation semantics for gradable adjectives. Their idea is to quantify over degree modifiers in comparatives. For instance, *Nathan is taller than Daniel is* means, essentially, that there is a degree modifier M such that *Nathan is M tall* is true but *Daniel is M tall is false*.

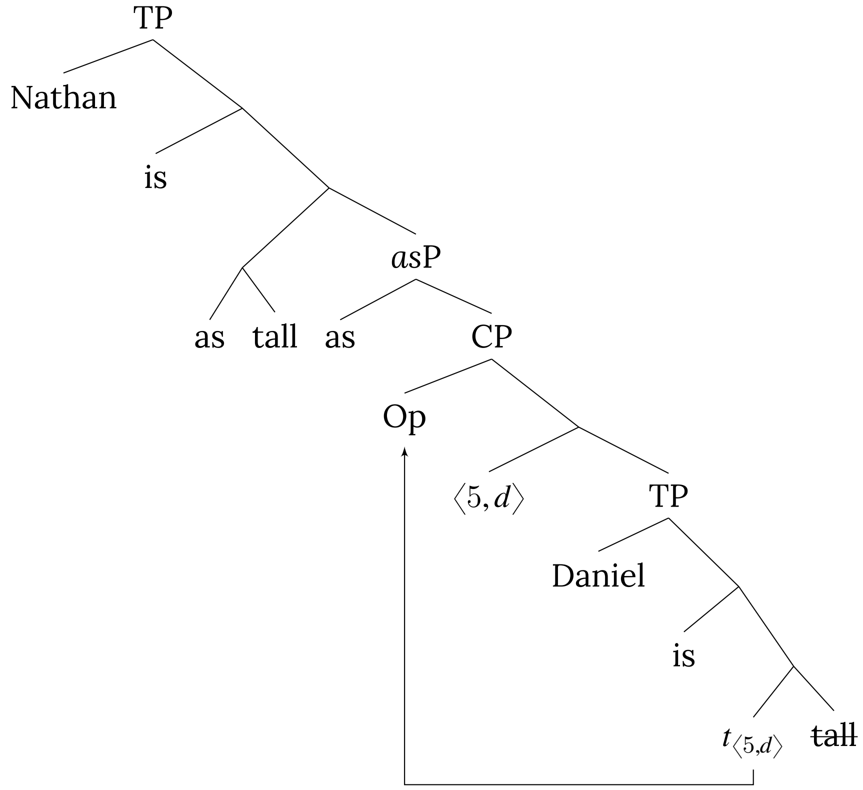
3.8 Exercise

This exercise is about (clausal) **equatives** of the form:

(3.65) Nathan is as tall as Daniel is.

We assume the following syntax for (3.65), which is isomorphic to the analysis of clausal comparatives developed above.

(3.66)



(i) Let us assume that *as* is semantically vacuous, just as *than* and *Op*. Compute the meaning of *as*P top-down by completing the following:

$$\begin{aligned} & [\![\textbf{as Op } \langle 5, d \rangle \textbf{ Daniel is } t_{\langle 5,d \rangle} \textbf{ tall}]\!]^{a,\mathcal{M}} \\ = & [\![\textbf{as}]\!]^{a,\mathcal{M}}([\![\textbf{Op } \langle 5, d \rangle \textbf{ Daniel is } t_{\langle 5,d \rangle} \textbf{ tall}]\!]^{a,\mathcal{M}}) \\ = & \cdots \end{aligned}$$

(ii) The only significant difference from the comparative construction is the meaning of the first occurrence of *as*, which occupies the position of *-er/more*. Consider the following lexical entry for *as*.

(3.67) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{as}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et \rangle}.\lambda f \in D_{\langle d,t \rangle}.\lambda x \in D_e. \left[\begin{array}{l} \text{there is a degree } d \in D_d \\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 1 \end{array} \right]$$

This analysis does not capture the correct truth-conditions. Explain with examples why it is inadequate.

(iii) OPTIONAL PROBLEM: Come up with an analysis of *as* that captures the intuitively available truth-conditions of (3.65) (Please explain your answer in words). Hint: the type should be the same as (3.67).

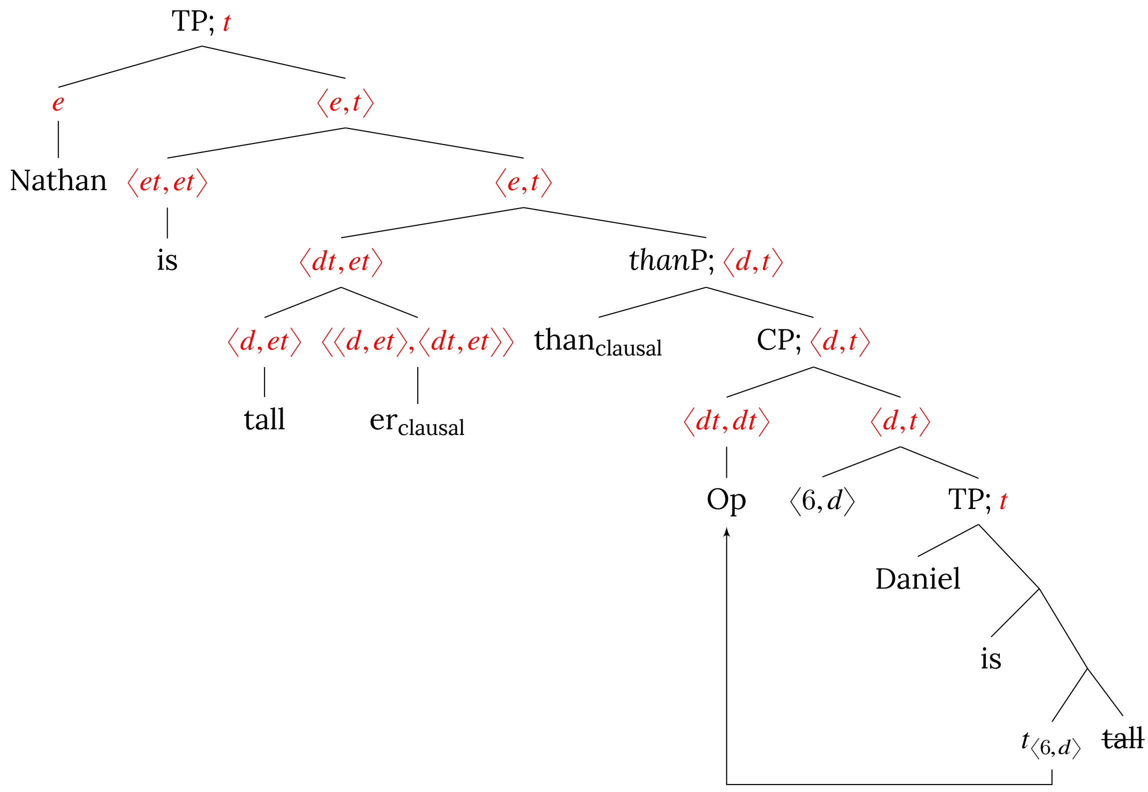
Chapter 4

Quantifiers in Comparatives

4.1 Review: The ‘A-not-A’ Theory of Comparatives

- The analysis of clausal comparatives:

(4.1) Nathan is taller than Daniel is.



  - The *than*-clause denotes (a function that characterises) the set of degrees to which Daniel is tall.[1]
  - $[\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}$ existentially quantifies over degrees and says there is a degree d such that Nathan is d-tall, but it is not a member of the *than*-clause, i.e. Daniel

[1]Recall: a function f of type $\langle\sigma,t\rangle$ characterises the set of type-σ elements $\{x \in D_\sigma \mid f(x) = 1\}$. A function of type $\langle e,t\rangle$ characterises a set of individuals and a function of type $\langle d,t\rangle$ characterises a set of degrees. $f(d) = 1$ means d is in the set, $f(d) = 0$ means d is not in the set.

is not d-tall.

- The lexical entries of the key items are in (4.2):

(4.2) For any model $\mathcal{M}$ and for any assignment a,
a. $[\![\textbf{tall}]\!]^{a,\mathcal{M}} = \lambda d \in D_d.\lambda x \in D_e.\ x$ is d-tall in $\mathcal{M}$
b. $[\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}$
$= \lambda G \in D_{\langle d,et\rangle}.\lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e. \left[\begin{array}{l}\text{there is a degree } d \in D_d \\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 0\end{array}\right]$
c. $[\![\textbf{Op}]\!]^{a,\mathcal{M}} = [\![\textbf{than}_{\textbf{clausal}}]\!]^{a,\mathcal{M}} = \lambda f \in D_{\langle d,t\rangle}.\ f$

- The *than*P denotes:

$$
\begin{aligned}
&[\![\textbf{than}_{\textbf{clausal}}\ \textbf{Op}\ \langle 6,d\rangle\ \textbf{Daniel is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}} \\
&= [\![\textbf{than}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}([\![\textbf{Op}]\!]^{a,\mathcal{M}}([\![\langle 6,d\rangle\ \textbf{Daniel is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}})) \\
&= [\![\langle 6,d\rangle\ \textbf{Daniel is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}} \\
&= \lambda d \in D_d.\ [\![\textbf{Daniel is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}} \\
&= \lambda d \in D_d.\ [\![\textbf{is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}([\![\textbf{Daniel}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}) \\
&= \lambda d \in D_d.\ [\![\textbf{is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}(\text{Daniel}) \\
&= \lambda d \in D_d.\ [\![\textbf{is}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}})(\text{Daniel}) \\
&= \lambda d \in D_d.\ [\![t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}(\text{Daniel}) \\
&= \lambda d \in D_d.\ [\![\textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}})(\text{Daniel}) \\
&= \lambda d \in D_d.\ [\![\textbf{tall}]\!]^{a[\langle 6,d\rangle \to d],\mathcal{M}}(d)(\text{Daniel}) \\
&= \lambda d \in D_d.\ \text{Daniel is } d\text{-tall in } \mathcal{M}
\end{aligned}
$$

- The rest of the sentence combines with this, yielding the following truth-conditions:

$$\begin{aligned}
&[\![\textbf{Nathan is tall -er}_{\textbf{clausal}}\textbf{ thanP}]\!]^{a,\mathcal{M}}\\
&= [\![\textbf{is tall -er}_{\textbf{clausal}}\textbf{ thanP}]\!]^{a,\mathcal{M}}([\![\textbf{Nathan}]\!]^{a,\mathcal{M}})\\
&= [\![\textbf{is tall -er}_{\textbf{clausal}}\textbf{ thanP}]\!]^{a,\mathcal{M}}(\text{Nathan})\\
&= [\![\textbf{is}]\!]^{a,\mathcal{M}}([\![\textbf{tall -er}_{\textbf{clausal}}\textbf{ thanP}]\!]^{a,\mathcal{M}})(\text{Nathan})\\
&= [\![\textbf{tall -er}_{\textbf{clausal}}\textbf{ thanP}]\!]^{a,\mathcal{M}}(\text{Nathan})\\
&= [\![\textbf{tall -er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}([\![\textbf{thanP}]\!]^{a,\mathcal{M}})(\text{Nathan})\\
&= [\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}([\![\textbf{tall}]\!]^{a,\mathcal{M}})([\![\textbf{thanP}]\!]^{a,\mathcal{M}})(\text{Nathan})\\
&= 1 \text{ iff there is a degree } d \in D_d \text{ such that}\\
&\qquad [\![\textbf{tall}]\!]^{a,\mathcal{M}}(d)(\text{Nathan}) = 1 \text{ and } [\![\textbf{thanP}]\!]^{a,\mathcal{M}}(d) = 0\\
&\quad \text{iff there is a degree } d \in D_d \text{ such that}\\
&\qquad \text{Nathan is } d\text{-tall in } \mathcal{M} \text{ and}\\
&\qquad [\lambda d' \in D_d.\ \text{Daniel is } d'\text{-tall in } \mathcal{M}](\text{d})\text{=0}\\
&\quad \text{iff there is a degree } d \in D_d \text{ such that}\\
&\qquad \text{Nathan is } d\text{-tall in } \mathcal{M} \text{ and}\\
&\qquad \text{Daniel is not } d\text{-tall in } \mathcal{M}
\end{aligned}$$

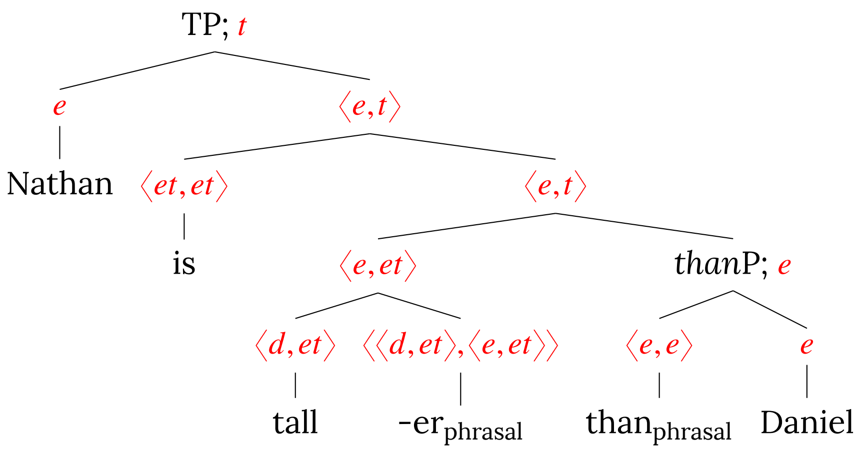
- This analysis is called 'A-not-A' (by Schwarzschild 2008) and was originally proposed by Seuren (1973, 1984).
- One nice feature of the analysis is that it straightforwardly accounts for the fact that Negative Polarity Items (NPIs) are licensed in *than*-clauses, because there is negation (which makes the *than*P a downward entailing context).

(4.3) a. Nathan is taller than **anybody** else.
b. Nathan is healthier than he **ever** was.

(4.4) a. John's laziness was stronger than his willingness to **lift a finger**.
b. He went further than I had **the slightest** intention of going.
c. My urge to steal was stronger than I **could help**.
d. This is more serious than I would have believed **at all** possible.
(Seuren 1973:534)

- The analysis can be extended to phrasal comparatives:

(4.5) Nathan is taller than Daniel.



(4.6) For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{than}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$

b. $[\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\lambda y \in D_e. \left[\begin{array}{l}\text{there is a degree } d \text{ such that}\\ G(d)(y) = 1 \text{ and } G(d)(x) = 0\end{array}\right]$

- For phrasal comparatives with degrees like (4.7), we postulate yet another lexical entry of *-er* (see last week's lecture notes).

(4.7) Nathan is taller than 180 cm.

4.2 The Problem of Quantifiers in 'Than'-Clauses

- Review: Generalised Quantifier Theory
  - Quantificational Ds denote functions of type $\langle et,\langle et,t\rangle\rangle$. In set talk, they are relations between two sets.

    (4.8) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{every}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}. \left[\begin{array}{l}\text{for every } x \in D_e \text{ such that } P(x) = 1,\\ Q(x) = 1\end{array}\right]$$
$$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ P \subseteq Q)$$

$$[\![\textbf{some}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}. \left[\begin{array}{l}\text{there is some } x \in D_e \text{ such that}\\ P(x) = 1 \text{ and } Q(x) = 1\end{array}\right]$$
$$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ P \cap Q \neq \varnothing)$$

$$[\![\textbf{no}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}. \left[\begin{array}{l}\text{there is no } x \in D_e \text{ such that}\\ P(x) = 1 \text{ and } Q(x) = 1\end{array}\right]$$
$$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ P \cap Q = \varnothing)$$

$$[\![\textbf{exactly two}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}. \left[\begin{array}{l}\text{there are exactly two } x \in D_e \text{ such that}\\ P(x) = 1 \text{ and } Q(x) = 1\end{array}\right]$$
$$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ |P \cap Q| = 2)$$

– Quantificational DPs denote functions of type $\langle et,t\rangle$ (aka **generalised quantifiers**), rather than individuals.

(4.9) For any model $\mathcal{M}$ and for any assignment a,

$$
\begin{aligned}
[\![\textbf{every girl}]\!]^{a,\mathcal{M}} &= \lambda Q \in D_{\langle e,t\rangle}.[\text{for every girl } x \text{ in } \mathcal{M}, Q(x) = 1] \\
&\quad (\approx \lambda Q \in D_{\langle e,t\rangle}.\ \{\, x \mid x \text{ is a girl in } \mathcal{M}\,\} \subseteq Q) \\
[\![\textbf{some girl}]\!]^{a,\mathcal{M}} &= \lambda Q \in D_{\langle e,t\rangle}.[\text{there is some girl } x \text{ in } \mathcal{M} \text{ such that } Q(x) = 1] \\
&\quad (\approx \lambda Q \in D_{\langle e,t\rangle}.\ \{\, x \mid x \text{ is a girl in } \mathcal{M}\,\} \cap Q \neq \varnothing) \\
[\![\textbf{no girl}]\!]^{a,\mathcal{M}} &= \lambda Q \in D_{\langle e,t\rangle}.[\text{there is no girl } x \text{ in } \mathcal{M} \text{ such that } Q(x) = 1] \\
&\quad (\approx \lambda Q \in D_{\langle e,t\rangle}.\ \{\, x \mid x \text{ is a girl in } \mathcal{M}\,\} \cap Q = \varnothing) \\
[\![\textbf{exactly two girls}]\!]^{a,\mathcal{M}} &= \lambda Q \in D_{\langle e,t\rangle}.[\text{there are exactly two girls } x \text{ in } \mathcal{M} \text{ such that } Q(x) = 1] \\
&\quad (\approx \lambda Q \in D_{\langle e,t\rangle}.\ |\{\, x \mid x \text{ is a girl in } \mathcal{M}\,\} \cap Q| = 2)
\end{aligned}
$$

- **The problem:** Our analysis above makes wrong predictions for than-clauses with certain quantifiers.
- For instance, consider (4.10).

  (4.10)Nathan is taller than every girl is.

  The intuitively available reading of (4.10) is paraphrased by:

  (4.11)For every girl g, Nathan is taller than g.

  But what we predict is (4.12), which is unavailable as a reading of the sentence.

  (4.12)Nathan is taller than the shortest girl.

  Let's go through the computation to see why this is the prediction.
- The *than*P denotes (the function that characterises) the set of degrees to which every girl is tall.

$$
\begin{aligned}
&[\![\textbf{than}_{\textbf{clausal}}\ \textbf{Op}\ \langle 6,d\rangle\ \textbf{every girl is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}\\
&= [\![\textbf{than}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}([\![\textbf{Op}]\!]^{a,\mathcal{M}}([\![\langle 6,d\rangle\ \textbf{every girl is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}))\\
&= [\![\langle 6,d\rangle\ \textbf{every girl is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}\\
&= \lambda d \in D_d.\ [\![\textbf{every girl is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{is}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{is}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}(d))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda d' \in D_d.\lambda x \in D_e.x \text{ is } d'\text{-tall in } \mathcal{M}](d))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ [\![\textbf{every}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ \left[\lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{for every } y \in D_e \text{ such that}\\ P(y)=1,\ Q(y)=1\end{array}\right]\right]\\
&\qquad\qquad ([\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } y \in D_e \text{ such that } [\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}(y)=1,\\ [\lambda x \in D_e.\ x \text{ is } d\text{-tall in } \mathcal{M}](y)=1\end{array}\right]\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } y \in D_e \text{ such that } [\lambda z \in D_e.\ z \text{ is a girl in } \mathcal{M}](y)=1,\\ [\lambda x \in D_e.\ x \text{ is } d\text{-tall in } \mathcal{M}](y)=1\end{array}\right]\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } y \in D_e \text{ such that } y \text{ is a girl in } \mathcal{M},\\ y \text{ is } d\text{-tall in } \mathcal{M}\end{array}\right]
\end{aligned}
$$

This characterises the set of degrees to which every girl is tall. What does this mean?

To be concrete, suppose that we have five girls $g_1, g_2, \dots g_5$, and their heights are, 165 cm, 170 cm, 173 cm, 175 cm and 180 cm, respectively. What is the set of degrees to which every girl is tall? Recall that according to our semantics of *tall*, if somebody is 165 cm tall, they are also 164 cm tall, 163 cm tall, etc. (or in other words, *tall* is downward monotonic with respect to its degree argument).

Then, in our context, every girl's height is more than 100 cm, so 100 cm is in the set. Likewise, every girl's height is more than 160 cm, so 160 cm is in the set too. However 166 cm is not in the set, because there is one girl, namely g_1, who is not that tall. So the set of degrees that the *than*P characterises is:

$$\{\, d \in D_d \mid d \text{ is on the scale of tallness and } d \preceq_{\text{tall}} 165 \text{ cm} \,\}$$

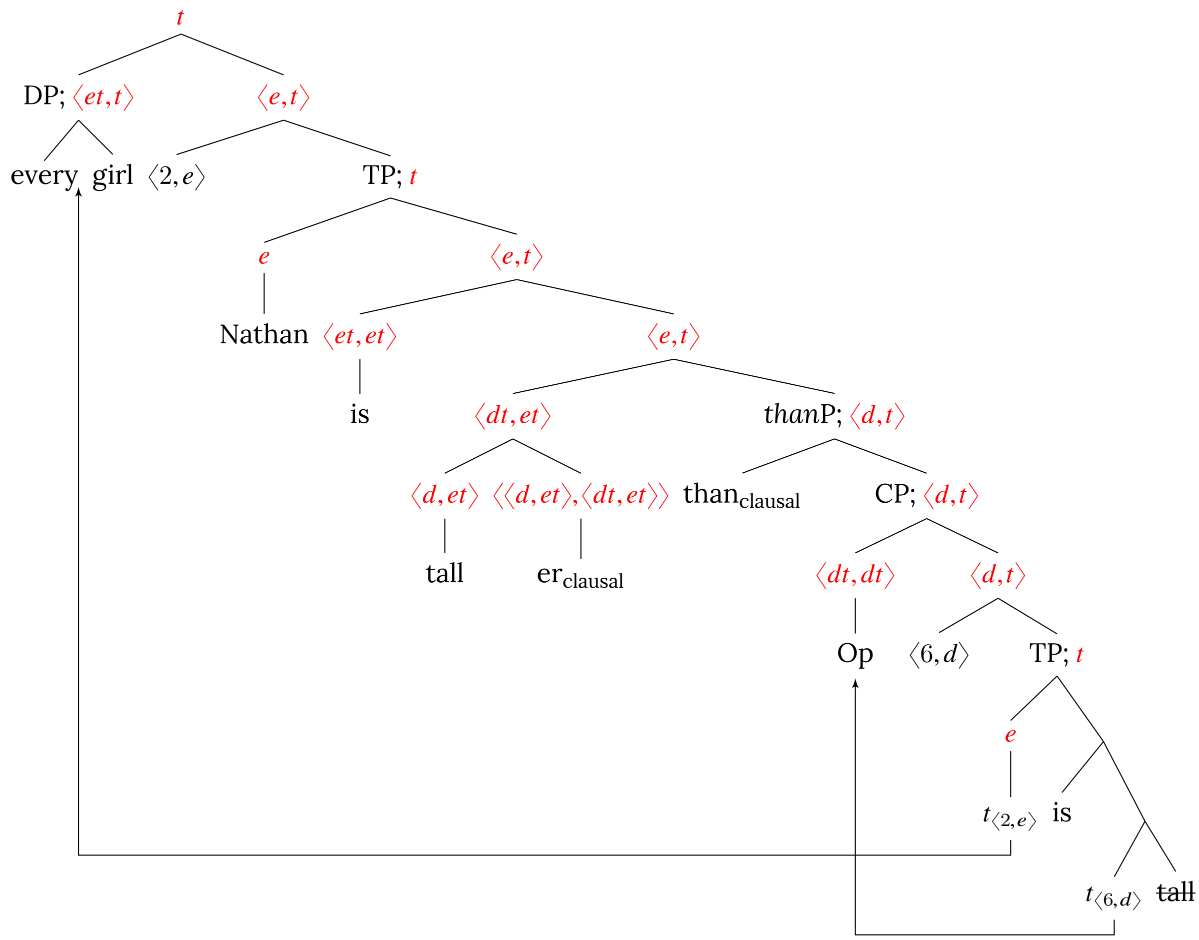
That is, this is the set of degrees d that the shortest girl is d-tall.

The rest of the sentence computes as before, and says: there is a degree d such that Nathan is d-tall and d is not in the set of degrees characterised by the *than*P. So the sentence is predicted to be true iff Nathan is taller than the height of the shortest girl.

4.3 Wide Scope?—No!

- We can derive the correct truth-conditions, if the quantifier moves out of the *than*-clause via Quantifier Raising (QR).

  (4.13)Nathan is taller than every girl is.



  This effectively means: For every girl g, Nathan is taller than g is tall, which is the right meaning.

- However, there are reasons to believe that this is not the right approach.

1. Firstly, the QR depicted above presumably violates locality constraints. Above all, it's extraction out of a finite clause, which is an island for quantifiers. For example, (4.14) cannot mean: for every professor, there is a different student who thinks that they are smart.

   (4.14)A different student thinks that every professor is smart.

Also, *wh*-movement is not possible from the *than*-clause:

(4.15) *Which girl is Nathan taller than t_{wh} is?

2. Secondly, we do not rule out the problematic reading. In particular, notice that the QR to the matrix clause is not necessary for interpretive purposes. Because the quantifier is in subject position in the *than*-clause, it can be interpreted there. Then, even if the QR is possible, we predict the sentence to be ambiguous, contrary to fact.

3. Thirdly, if the QR can apply to every quantifier, the following sentence should have a coherent reading.

   (4.16) *Nathan is taller than no girl is.

   This should be able to mean: There is no girl such that Nathan is taller than her.

   Contrast this with the phrasal version, (4.17), which is felicitous and means exactly this.

   (4.17) Nathan is taller than no girl.

   This observation suggests that QR out of *than*P is indeed possible in phrasal comparatives, but not in clausal comparatives.

4. Fourthly, QR won't give you the right meaning for sentences like (4.18), as pointed out by Schwarzschild & Wilkinson (2002).

   (4.18) Nathan is taller than I predicted most girls were.

   Since we have not covered the meaning of intensional predicates like *predict*, we will not discuss the derivation, but you can still appreciate the problem with your intuitions.

   The nub of the problem is this: if the quantifier takes scope in the matrix clause, the sentence should be about a particular set of girls, because it would mean: For most girls, Nathan is taller than I predicted they were. But (4.18) is (also) true in contexts where I didn't make predications about particular girls, e.g. my prediction is: Most girls less than 175 cm. My prediction is not about a particular majority of girls.

5. (Fifthly, an analogous problem can be created with intensional predicates, which are essentially quantifiers themselves, but unlike quantificational DPs, presumably do not undergo QR. For instance, according to our analysis, (4.19) means "Nathan is taller than the minimum permitted height", while the intuitively available reading says "Nathan is taller than the maximum required height".

   (4.19) Nathan is taller than he is required to be.

   Since intensional predicates do not QR, we cannot solve this problem with QR.)

4.4 Negation in the Than-Clause

The discussion above suggests that the quantifier needs to stay in the *than*-clause. Let us review the problem of (4.10) again.

(4.10) Nathan is taller than every girl is.

The problem is that the *than*-clause denotes (the characteristic function of) the set of degrees to which every girl is tall, which is equivalent to the set of degrees that the shortest girl is tall. So the comparison is effectively only between Nathan and the shortest girl.

Another way of looking at this problem is that the negation encoded in *-er* takes scope over *every girl*, yielding the truth-conditions in (4.20).

(4.20) There is a degree d such that Nathan is d-tall and NOT(every girl is d-tall).

Putting the question of composition aside for the moment, it turns out that the correct truth-conditions are predicted if the negation takes scope below *every girl*:

(4.21) There is a degree d such that Nathan is d-tall and every girl is NOT d-tall.

But how can we have the quantifier outscope the negation, if we cannot QR the quantifier?

4.4.1 Invisible Negation

The proposal is to put the negation in the *than*-clause, rather than in the lexical entry of *-er*. Of course, this negation is invisible in our examples, but this is not completely *ad hoc*. In a number of languages, *than*-clauses can (or sometimes must) contain negation without changing the meaning.

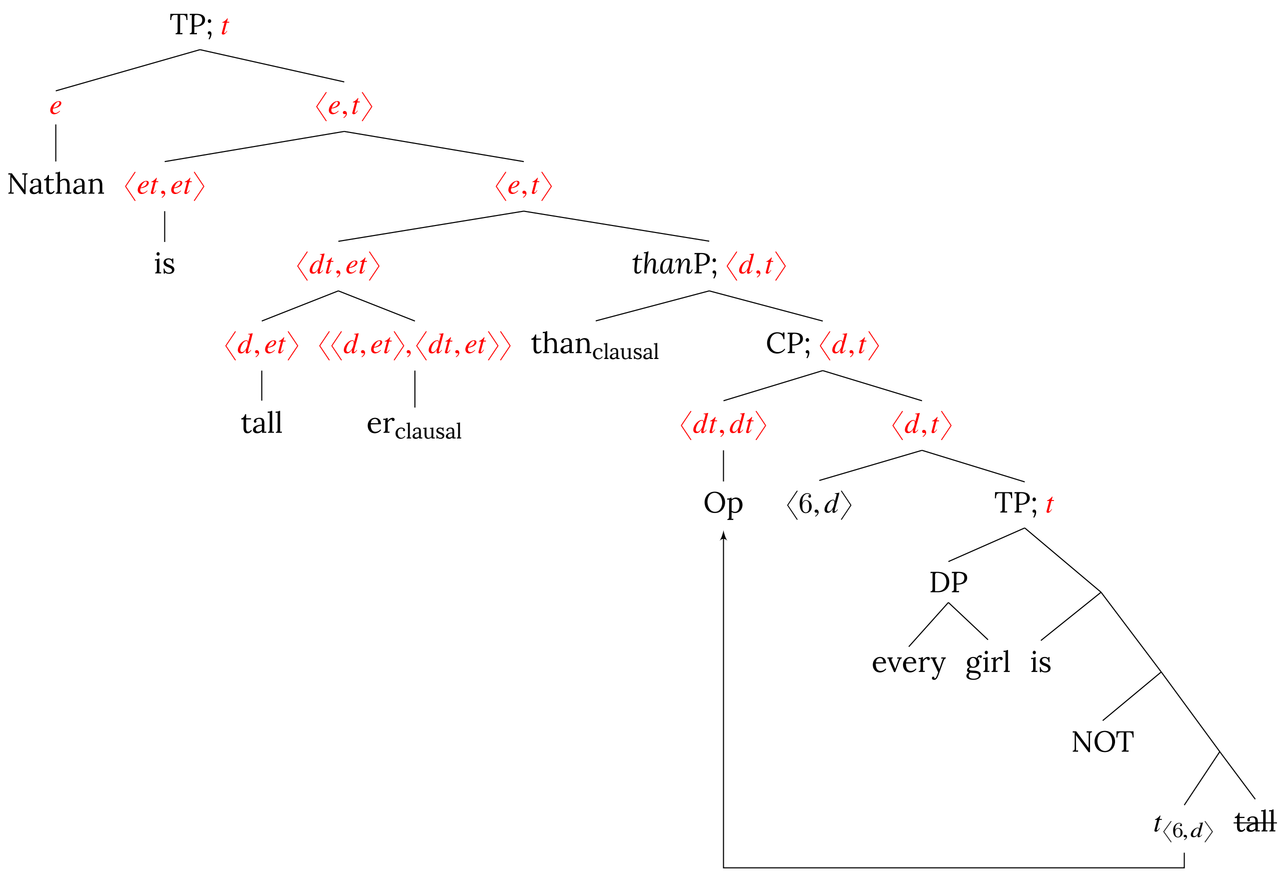
(4.22) Gianni è più alta di quanto (**non**) lo sia Maria.
Gianni is more tall of how.much **neg** it is Maria
'Gianni is taller than Maria is.' (Italian; Jacopo Romoli, p.c., 9 Feb 2015)

(4.23) Jean est plus grand que je **ne** pensais.
Jean is more tall than I **neg** thought
'Jean is taller than I thought.' (French; Seuren 1973:535)

(4.24) She did a better job than what I **never** thought she would.
(Cockney English; Seuren 1973:535)

If this negation is actually not semantically vacuous and takes scope below the quantifier, we derive the correct meaning. Let us see how this works step by step. Firstly, the LF looks like (4.25).

(4.25)Nathan is taller than every girl is.



Let's assume the following lexical entries:

(4.26)For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}$

$$= \lambda G \in D_{\langle d,et\rangle}.\lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e.\left[\begin{array}{l}\text{there is a degree } d \in D_d \\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 1\end{array}\right]$$

b. $[\![\textbf{NOT}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ \lambda x \in D_e.\ P(x) = 0$

The negation is separated from *-er* now (the old entry said $f(d) = 0$). Also the negation in (4.26b) works in simple, non-gradable sentences like *John is not British* and *John did not run*.

Then the *than*-clause denotes (the characteristic function of) the set of degrees to which every girl is not tall. Here's the computation:

$$
\begin{aligned}
&[\![\textbf{than}_{\textbf{clausal}}\ \textbf{Op}\ \langle 6,d\rangle\ \textbf{every girl is NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}\\
&= [\![\textbf{than}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}([\![\textbf{Op}]\!]^{a,\mathcal{M}}([\![\langle 6,d\rangle\ \textbf{every girl is NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}))\\
&= [\![\langle 6,d\rangle\ \textbf{every girl is NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a,\mathcal{M}}\\
&= \lambda d \in D_d.\ [\![\textbf{every girl is NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{is NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{is}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}\ t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}\ \textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![t_{\langle 6,d\rangle}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{tall}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}(d)))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda d' \in D_d.\lambda x \in D_e.x \text{ is } d'\text{-tall in } \mathcal{M}](d)))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{NOT}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}]))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda P \in D_{\langle e,t\rangle}.\ \lambda y \in D_e.\ P(y)=0]([\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}]))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}(\lambda y \in D_e.[\lambda x \in D_e.x \text{ is } d\text{-tall in } \mathcal{M}](y))\\
&= \lambda d \in D_d.\ [\![\textbf{every girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\lambda y \in D_e.y \text{ is not } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ [\![\textbf{every}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}([\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})([\lambda y \in D_e.y \text{ is not } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ \left[\lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{for every } x \in D_e \text{ such that}\\ P(x)=1,\ Q(x)=1\end{array}\right]\right]\\
&\qquad\qquad ([\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}})([\lambda y \in D_e.y \text{ is not } d\text{-tall in } \mathcal{M}])\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } x \in D_e \text{ such that } [\![\textbf{girl}]\!]^{a[\langle 6,d\rangle\to d],\mathcal{M}}(x)=1,\\ [\lambda y \in D_e.y \text{ is not } d\text{-tall in } \mathcal{M}](x)=1\end{array}\right]\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } x \in D_e \text{ such that } [\lambda z \in D_e.\ z \text{ is a girl in } \mathcal{M}](x)=1,\\ [\lambda y \in D_e.\ y \text{ is not } d\text{-tall in } \mathcal{M}](x)=1\end{array}\right]\\
&= \lambda d \in D_d.\ \left[\begin{array}{l}\text{for every } x \in D_e \text{ such that } x \text{ is a girl in } \mathcal{M},\\ x \text{ is not } d\text{-tall in } \mathcal{M}\end{array}\right]
\end{aligned}
$$

This function of type $\langle d,t\rangle$ characterises the set of degrees to which every girl is not tall. To see what kind of set this is, suppose again that there are five girls and their heights are 165 cm, 170 cm, 173 cm, 175 cm and 180 cm. Then every girl is (at least) 165 cm tall, some girls are 175 cm tall or taller, etc. But no girl is 190 cm tall. So every girl is such that she is NOT 190 cm tall. So 175 cm is not in the set, but 190 cm is. So the set of degrees characterised by the *than*P is:

$$\{\ d \in D_d \mid d \text{ is on the scale of tallness and } d \succ_{\text{tall}} 180 \text{ cm}\ \}$$

This is the set of degrees exceeding the tallest girl's height.

The rest of the sentence says: There is a degree d such that Nathan is d-tall and d is in the set of degrees characterised by the *than*P (Notice that it now says d is in the set, because *-er* doesn't contain negation anymore). That is to say, there is a degree d such that Nathan is d-tall and d exceeds the tallest girl's height on the scale of tallness. This is the correct truth-conditions.

The negation-in-the-than-clause analysis predicts the right truth-conditions for (4.27) as well.

(4.27) Nathan is taller than some girl is.

This is left for an exercise.

A nice feature of the present analysis is that it gives an account of (4.28) with an auxiliary assumption about trivially true sentences.

(4.28) *Nathan is taller than no girl is.

If there is a hidden negation taking scope under *no girl*, the *than*-clause characterises the set of degrees d such that no girl is not d-tall, which is the set of degrees to which every girl is d-tall. The whole sentence therefore says, there is a degree d to which Nathan is d-tall and every girl is d-tall. Notice that this is trivially true, since everybody by assumption has some degree of tallness. Assuming that such a trivially true sentence is ruled out (see Gajewski 2002 for an interesting idea related to this assumption), the unacceptability of (4.28) is explained.

Furthermore, this explanation applies to (almost) all downward monotonic quantifiers.[2]

(4.29) a. *Nathan is taller than not every girl is.
b. *Nathan is taller than few girls are.
c. *Nathan is taller than fewer than 10 girls are.
d. *Nathan is taller than neither Daniel nor Andrew is.

These sentences are trivial for the same reason above (more on this in the exercise). (Remaining puzzle: compare (4.29c) to "??Nathan is taller than at most 9 girls are"; see Gajewski 2008 for related discussion).

However, as Laura pointed out in class, one problem is that we predict that addition another negation to the above sentences should make the above sentences non-trivial. For instance, consider (4.30).

(4.30) *Nathan is taller than no girl is not.

The predicted truth-conditions of this sentence is:

(4.31) There is a degree d such that Nathan is d-tall and no girl is d-tall.

This reading is not available.

Besides, there are a number of other problems.

[2] Recall that a generalised quantifier Q is downward monotonic if for any $A, B \in D_{\langle e,t\rangle}$ such that for every $x \in D_e$ whenever $A(x) = 1$, $B(x) = 1$, whenever $Q(B) = 1$, then $Q(A) = 1$.

4.4.2 Problems

1. Firstly, we took out the negation from *-er* in the clausal comparative and put it in the *than*-clause, but this is not possible with phrasal comparatives.

   (4.32) Nathan is taller than Daniel.

   So *-er* for phrasal comparatives still needs negation, while *-er* for clausal comparatives do not contain negation.

   (4.33) a. $[\![\textbf{-er}_{\textbf{clausal}}]\!]^{a,\mathcal{M}}$
   $$= \lambda G \in D_{\langle d,et\rangle}.\lambda f \in D_{\langle d,t\rangle}.\lambda x \in D_e.\left[\begin{array}{l}\text{there is a degree } d \in D_d \\ \text{such that } G(d)(x) = 1 \text{ and } f(d) = 1\end{array}\right]$$
   b. $$[\![\textbf{-er}_{\textbf{phrasal}}]\!]^{a,\mathcal{M}} = \lambda G \in D_{\langle d,et\rangle}.\lambda x \in D_e.\lambda y \in D_e.\left[\begin{array}{l}\text{there is a degree } d \text{ such that} \\ G(d)(y) = 1 \text{ and } G(d)(x) = 0\end{array}\right]$$

   This is not an empirical problem, but it suggests that something is amiss.

2. Secondly, the analysis still predicts the incorrect truth-conditions for (4.10), provided that the quantifier can take scope below negation.

   In non-comparative sentences like (4.34), *every girl* can take narrow scope.

   (4.34) Every girl didn't meet Nathan.
   a. Reading 1: every girl is such that she didn't meet Nathan.
   b. Reading 2: Not every girl met Nathan.

   If the negation takes scope above the quantifier in our comparative sentence (4.10), the *than*P characterises the set of degrees d such that not every girl is d-tall. What is this set? In our scenario above, 165 cm and every degree below it is such that every girl is that tall. Every other degree belongs to the set of degrees that not every girl has, which is every degree above 165 cm. Consequently, the sentence is true iff Nathan is taller than 165 cm, the height of the shortest girl. Unsurprisingly, this is the same wrong prediction that we started out with.

   In order to block this, it needs to be stipulated that the universal quantifier cannot take scope below NOT, although it can in simple sentences like (4.34). This is not impossible, but lacks cogency.

   Furthermore, NOT should be able to take scope above certain quantifiers. For instance, consider (4.35).

   (4.35) Nathan is taller than any girl is.

   The correct reading is only predicted with NOT outscoping *any girl* (which is an existential quantifier). This might not be so surprising given that *any* is an NPI. But this shows that NOT does not always take narrow scope.

   A small note: the wide scope NOT requires a type-$\langle t,t\rangle$ version of the entry (which is the negation $\neg$ in Propositional Logic):

   $$[\![\textbf{NOT}]\!]^{a,\mathcal{M}} = \lambda v \in D_t.\begin{cases}1 & \text{if } v = 0\\ 0 & \text{if } v = 1\end{cases}$$

3. Thirdly, the analysis makes the wrong prediction for non-monotonic quantifiers. For instance, consider:

   (4.36) Nathan is taller than exactly two girls are.

   The *than*P characterises the set of degrees d such that exactly two girls are not d-tall. To see what this means, consider the five girls mentioned above (whose heights are: 165 cm, 170 cm, 173 cm, 175 cm and 180 cm). For these girls, the set contains any degree greater than 170 cm and smaller than 173 cm. And the whole sentence says there is a degree d such that Nathan is d-tall and d is in this set. This will be true as soon as Nathan is taller than 170 cm. So the sentence is true even if Nathan is taller than all the girls. Effectively, the sentence is predicted to be synonymous with 'Nathan is taller than the shortest two girls are'.

   Also, having the negation above the quantifier does not help in this case. With the wide scope NOT, the *than*P denotes the set of degrees d such that it is not the case that exactly two girls are d-tall. In the above scenario, the degrees greater than 173 cm and up to 175 cm are the degrees to which exactly two girls are tall. Therefore, the predicted truth-conditions are: (4.36) is true iff Nathan is tall to a degree that doesn't fall in this range. Notice that this is a trivially true statement. Nathan has some height, so there is definitely a small degree far below 173 cm to which Nathan is tall. Thus, this doesn't capture the attested reading either.

   The same problem arises with all non-monotonic quantifiers (more on this in the exercise).

   (4.37) a. Nathan is taller than some but not all girls are.
   b. Nathan is taller than between 2 and 4 girls are.
   c. Nathan is taller than an even number of girls are.

To conclude, by moving the negation to the *than*-clause, we can derive the correct truth-conditions for clausal comparatives containing certain types quantifiers such as *every girl* and *no girl*, but not others such as *exactly two girls*.

This problem is in fact still not completely solved, and has been discussed in many different theories of comparatives including the A-not-A analysis we discussed (see the Further Readings section for references).

4.5 Quantifiers in the Matrix Clause

According to the A-not-A analysis of comparatives, -*er* is an existential quantifier. It is natural to wonder whether -*er* exhibits scope interactions with quantifiers.

Let's start with (4.38).

(4.38) Every girl is taller than Nathan is.

If the quantifier *every girl* takes scope over -*er*, this effectively reduces to the conjunction of all the sentences of the form: g_1 is taller than Nathan is, g_2 is taller than

Nathan is, g_3 is taller than Nathan is, etc., because the reading is paraphrased by (4.39).

(4.39)a. For every girl g, g is taller than Nathan is.
b. For every girl g, there is a degree d such that g is d-tall and Nathan is not d-tall.

This is the right reading.

What if *-er* takes scope over *every girl*? The predicted truth-conditions are:

(4.40)There is a degree d such that every girl is d-tall and Nathan is not d-tall.

If the shortest girl is 165 cm, all the degrees smaller than or equal to 165 cm are shared by all the girls. So (4.40) is the case iff the shortest girl is taller than Nathan. This also captures the right reading.

Next, consider (4.41).

(4.41)No girl is taller than Nathan.

If *no girl* takes scope over *-er*, we get the right truth-conditions.

(4.42)a. For no girl g, g is taller than Nathan is.
b. For no girl g, is there a degree d such that g is d-tall and Nathan is not d-tall.
c. For every girl g, there is no degree d such that g is d-tall and Nathan is not d-tall.

What if *-er* takes wide scope? In this case, the predicted truth-conditions are:

(4.43)There is a degree d such that no girl is d-tall and Nathan is not d-tall.

Notice that this is trivially true, because the scale is not upper-bounded. For instance, 2000 km is such a degree (no one is 2000 km tall!). On the assumption that trivial statements are ruled out (cf. the above discussion on downward monotonic quantifiers), we can rule out this reading.

However, this explanation relies on the fact that the truth-conditions are trivially satisfied, which is essentially due to the assumption that the scale of tallness is unbounded. In fact, with an upper bounded scale, the predicted truth-conditions are not trivially true. Take *clean* as an example of an adjective with an upper bounded scale.

(4.44)No (other) table is cleaner than this one is.

With *-er* outscoping *no table*, the truth-conditions should be:

(4.45)There is a degree d such that no (other) table is d-clean and this table is not d-clean.

This can be false, namely, when there is at least one table that is completely clean, because in that case, the completely clean table or tables have all the degrees of cleanness. And this is true whenever none of the tables are completely clean. Thus, this reading is non-trivial and should not be ruled out, but the sentence cannot mean it.

A problem also arises with a non-monotonic quantifier.

(4.46) Exactly two girls are taller than Nathan is.

The intuitively available reading is derived with *-er* taking narrow scope.

(4.47) a. For exactly two girls g, g is taller than Nathan is.
b. For exactly two girls g, there is a degree d such that g is d-tall and Nathan is not d-tall.

If *-er* outscopes the quantifier, the truth-conditions will be:

(4.48) There is a degree d such that exactly two girls are d-tall and Nathan is not d-tall.

This can be true even if there are more than two girls who are taller than Nathan. To see this more concretely, suppose that Nathan is 180 cm tall and there are five girls whose heights are 181 cm, 182 cm, 183 cm, 184 cm and 185 cm. The degrees between 184 cm and 185 cm are the degrees d such that exactly two girls are d-tall. Notice that Nathan is not that tall. Therefore, (4.48) is the case. But there are five girls whose height exceeds Nathan's!

The lesson here is that generally speaking, we don't want *-er* to take scope over a quantifier in the matrix clause. This generalisation was reached by Kennedy (1999) but Heim (2000) discovered certain exceptions to this, namely, she points out *-er* scopally interacts with certain modals and other intensional predicates. As we have not touched on the semantics of modals, we will not discuss this issue here.

4.6 Further Readings

The problem of quantifiers in *than*-clauses was originally pointed out by Larson (1988). It is still largely an unsolved puzzle. The A-not-A analysis of comparatives, which we adopted, was originally proposed by Seuren (1973, 1984), and the problem of quantifiers in *than*-clauses is tackled in this framework by Gajewski (2008), and Van Rooij (2008) (see also Schwarzschild 2008).

There are alternative analyses of the puzzle of quantifiers in comparatives. Schwarzschild & Wilkinson (2002), Heim (2006), Krasikova (2008) and Beck (2011) propose to make use of **intervals** (which are convex sets of degrees) instead of degrees themselves in the semantics of comparatives. Beck (2011) is a nice overview of these approaches. More recently, Beck (2014) and Dotlačil & Nouwen (2014) put forward similar ideas using **plural degrees** instead of intervals.

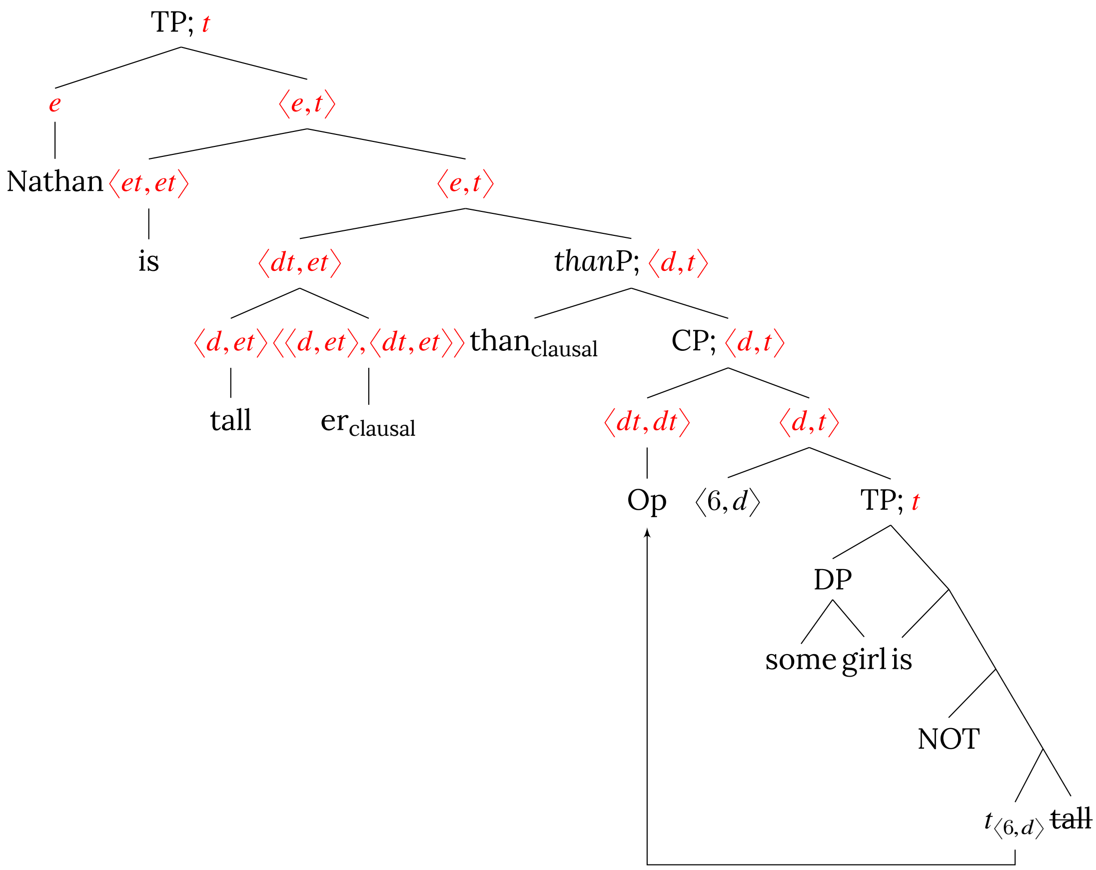
The problem of quantifiers in matrix clauses of comparatives was initially raised by

Kennedy (1999), and has also been discussed by Heim (2000). The generalisation that the scope of *-er* does not interact with quantificational DPs is known as the Kennedy-Heim generalisation. Interestingly, Heim (2000) shows that the scope of *-er* does interact with intensional predicates like *required* and *allowed*, which are also kind of quantifiers. This is also an unsolved puzzle at this moment.

4.7 Exercises

1. Compute the meaning of (4.49), following the negation-in-the-than-clause analysis developed above (with respect to model $\mathcal{M}$ and assignment a).

   (4.49) Nathan is taller than some girl is.



   Are the derived truth-conditions intuitively correct?

2. Recall that downward monotonic quantifiers are unacceptable in *than*-clauses.

   (4.29) a. *Nathan is taller than not every girl is.
   b. *Nathan is taller than few girls are.
   c. *Nathan is taller than fewer than 10 girls are.
   d. *Nathan is taller than neither Daniel nor Andrew is.

   Let's compute the meaning of the *than*-clause of (4.49c). We assume the following lexical entry for *fewer than* 10.

(4.50) $[\![$**fewer than 10**$]\!]^{a,\mathcal{M}}$

$$= \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{there are fewer than 10 individuals } x \in D_e \text{ such that}\\ P(x) = 1 \text{ and } Q(x) = 1\end{array}\right]$$

$$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ |P \cap Q| < 10)$$

What is the denotation of the *than*P?

$[\![$**than Op** $\langle 1, d\rangle$ **fewer than 10 girls are NOT** $t_{\langle 1,d\rangle}$ ~~**tall**~~$]\!]^{a,\mathcal{M}} =$

It's up to you to show the whole computation (although doing so is recommended).

Also explain which set of degrees this function characterises and why the sentence (4.49c) has trivially true truth-conditions.

3. The following sentence is problematic for the negation-in-the-than-clause analysis presented above.

   (4.51) Nathan is taller than an even number of girls are.

   Assume the following denotation for *an even number of* (which is treated as a single lexical item for the sake of simplicity).

   (4.52) $[\![$**an even number of**$]\!]^{a,\mathcal{M}}$

   $$= \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{there are an even number of individuals } x \in D_e\\ \text{such that } P(x) = 1 \text{ and } Q(x) = 1\end{array}\right]$$

   $$(\approx \lambda P \in D_{\langle e,t\rangle}.\lambda Q \in D_{\langle e,t\rangle}.\ |P \cap Q| \text{ is even})$$

   Assuming that the invisible negation NOT takes scope below this quantifier in the *than*-clause, what does the following denote? Again, you can but need not show the computation.

   $[\![$**than Op** $\langle 1, d\rangle$ **an even number of girls are NOT** $t_{\langle 1,d\rangle}$ ~~**tall**~~$]\!]^{a,\mathcal{M}} =$

   What is the set of degrees that this function characterises?

   Also explain why this is a problem with an example situation.

4. (Optional) Do the same thing with NOT taking wide scope.

   $[\![$**than Op** $\langle 1, d\rangle$ **NOT an even number of girls are** $t_{\langle 1,d\rangle}$ ~~**tall**~~$]\!]^{a,\mathcal{M}} =$

Part II

Plurality

Chapter 5

Plural Predication

- By assumption, proper names denote individuals, or equivalently, their semantic type is e.[1]

(5.1) For any assignment a, and for any model $\mathcal{M}$,
a. $[\![\mathbf{John}]\!]^{a,\mathcal{M}} = \text{John}$
b. $[\![\mathbf{Mary}]\!]^{a,\mathcal{M}} = \text{Mary}$

- Starting from this week, we will discuss the meanings of plural phrases.
- Our first question is: What does the conjoined name *John and Mary* denote?
- In this lecture, we propose to enrich our model with new semantic objects called **i(ndividual)-sums** and discuss the semantics of various kinds of predicates.

5.1 A Failed Attempt: Type-Shifting + Generalised Conjunction

- The first thing we should try is to assign a denotation to *John and Mary*, while keeping our semantic theory intact.
- We have three phrases here: *John*, *and*, and *Mary*. We know the meanings of the proper names, i.e. (5.1). But what about *and*?
- We discussed the meaning of *and* in Term 1 (Assignment 8) as generalised conjunction, but generalised conjunction can only apply to two things whose types 'end in t' (e.g. t, $\langle e,t\rangle$, $\langle et,t\rangle$, $\langle et,\langle et,t\rangle\rangle$). Specifically:

(5.2) *Generalised conjunction*
For any assignment a, for any model $\mathcal{M}$, and for any type σ that ends in t,

$$[\![\mathbf{and}]\!]^{a,\mathcal{M}} = \lambda x \in D_\sigma . \lambda y \in D_\sigma . \begin{cases} x = y = 1 & \text{if } \sigma = t \\ \lambda z \in D_{\tau_1} .\ [\![\mathbf{and}]\!]^{a,\mathcal{M}}(x(z))(y(z)) & \text{if } \sigma = \langle \tau_1, \tau_2 \rangle \end{cases}$$

[1]Recall that more precisely, $[\![\mathbf{John}]\!]^{a,\mathcal{M}} = \mathcal{I}(\mathbf{John})$, where $\mathcal{I}(\mathbf{John}) \in \mathcal{D}$, whichever element it is. For expository purposes, we call this individual 'John'.

Since e does not end in t, generalised conjunction cannot apply to individuals!!

- But recall, also from Term 1, that we can type-shift the meanings of type-e objects to type-$\langle et,t\rangle$ functions, as in (5.3).

(5.3) For any assignment a, and for any model $\mathcal{M}$,

a. $[\![\textbf{John}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P(\text{John}) = 1$

b. $[\![\textbf{Mary}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P(\text{Mary}) = 1$

Generalised conjunction can conjoin two type-$\langle et,t\rangle$ functions. For *John and Mary* it will deliver the following function.

(5.4)
$$\left[\!\!\left[\ [\ \textbf{John}_{\langle et,t\rangle}\ [\ \textbf{and}\ \ \textbf{Mary}_{\langle et,t\rangle}\ ]\ ]\ \right]\!\!\right]^{a,\mathcal{M}} = [\![\textbf{and}]\!]^{a,\mathcal{M}}([\![\textbf{Mary}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}})([\![\textbf{John}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}})$$
$$= \lambda P \in D_{\langle e,t\rangle}.\ [\![\textbf{Mary}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}}(P) = [\![\textbf{John}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}}(P) = 1$$
$$= \lambda P \in D_{\langle e,t\rangle}.\ P(\text{Mary}) = P(\text{John}) = 1$$

- This analysis works for examples like (5.5).

(5.5) a. John and Mary smoke.
b. John and Mary are students.
c. John and Mary are British.

For instance, the meaning of (5.5a) will be:

(5.6)
$$[\![\textbf{John}_{\langle et,t\rangle}\ \textbf{and Mary}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}}([\![\textbf{smoke}]\!]^{a,\mathcal{M}})$$
$$= [\lambda P \in D_{\langle e,t\rangle}.\ P(\text{Mary}) = P(\text{John}) = 1]([\lambda x \in D_e.\ x \text{ smokes in } \mathcal{M}])$$
$$= 1 \text{ iff } [\lambda x \in D_e.\ x \text{ smokes in } \mathcal{M}](\text{Mary}) = 1$$
$$\text{and } [\lambda x \in D_e.\ x \text{ smokes in } \mathcal{M}](\text{John}) = 1$$
$$\text{iff Mary smokes in } \mathcal{M} \text{ and John smokes in } \mathcal{M}$$

- Essentially, under this analysis, 'John and Mary VP' ends up meaning 'John VP and Mary VP'.
- So far so good. However, a problem arises with the following types of examples.

(5.7) a. John and Mary met.
b. John and Mary like each other.
c. John and Mary are a couple.

What is predicted for (5.7a), for example, is this:

(5.8)
$$[\![\textbf{John}_{\langle et,t\rangle}\ \textbf{and Mary}_{\langle et,t\rangle}]\!]^{a,\mathcal{M}}([\![\textbf{met}]\!]^{a,\mathcal{M}})$$
$$= [\lambda P \in D_{\langle e,t\rangle}.\ P(\text{Mary}) = P(\text{John}) = 1]([\lambda x \in D_e.\ x \text{ met in } \mathcal{M}])$$
$$= 1 \text{ iff } [\lambda x \in D_e.\ x \text{ met in } \mathcal{M}](\text{Mary}) = 1$$
$$\text{and } [\lambda x \in D_e.\ x \text{ met in } \mathcal{M}](\text{John}) = 1$$
$$\text{iff Mary met in } \mathcal{M} \text{ and John met in } \mathcal{M}$$

The problem here is, first of all, it is not clear what ‘Mary met’ and ‘John met’ mean. Furthermore, even if that made sense, it wouldn’t be straightforward to capture the intended reading that the meeting was between John and Mary, because the predicate is applied to John and Mary separately.

- So (5.4) works for predicates like *smoke*, *are students* and *are British*, but not for predicates like *met*, *like each other* and *are a couple*. We call the former group of predicates **distributive predicates** and the latter group of predicates **collective predicates**.
- What is the difference between these two classes of predicates?

5.2 Three Types of Predicates

- We can define distributive predicates as those predicates that support the **distributivity inference**. (We speak of VPs here, but we can generalise this to other syntactic categories.)

(5.9) We say a VP has a **distributivity inference** if the following holds: ‘John and Mary VP’ entails and is entailed by ‘John VP and Mary VP’.

Here are some examples demonstrating *fell asleep* is a distributive predicate:

(5.10)a. ‘John and Mary fell asleep’ entails ‘John fell asleep and Mary fell asleep’.
b. ‘John and Mary fell asleep’ is entailed by ‘John fell asleep and Mary fell asleep’.

Since mutual entailment amounts to truth-conditional identity, we can say that for distributive predicates, ‘John and Mary VP’ and ‘John VP and Mary VP’ are truth-conditionally synonymous.

- **Collective predicates**, on the other hand, are those predicates that do not validate the distributive inference in either direction. Generally the right-hand side is simply unacceptable, so it doesn’t even make sense to talk about entailment here.

(5.11)a. ‘John and Mary look alike’ does not entail ‘John looks alike and Mary looks alike’.
b. ‘John and Mary look alike’ is not entailed by ‘John looks alike and Mary looks alike’.

- The analysis in the previous section only works for distributive predicates, because it is made to derive the distributivity inference.

(5.12) $$\left[\!\!\left[\; [\, \mathbf{John}_{\langle et,t\rangle} \; [\, \mathbf{and} \;\; \mathbf{Mary}_{\langle et,t\rangle} \,] \,] \; \right]\!\!\right]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ P(\text{Mary}) = P(\text{John}) = 1$$

- Interestingly, there are VPs that are neither distributive nor collective. Here is a concrete example.

  (5.13)John and Mary bought a house.

  This type of predicates validate the entailment in one direction.

  (5.14)a. 'John and Mary bought a house' does not entail 'John bought a house and Mary bought a house'.
  b. 'John and Mary bought a house' is entailed by 'John bought a house and Mary bought a house'.

  It is easy to see that (5.14b) is true. For (5.14a), consider a situation where John and Mary, as a couple, bought a house by splitting the cost. Then it's not true that John bought a house (he did with Mary, but not on his own), and it's not true that Mary bought a house (she did with John, but not on her own). (Of course, there are situations where both sentences are true, namely situations where John and Mary each bought a house, but this is not enough to validate an entailment.)

- Predicates like 'bought a house' which only validate the distributivity inference in one direction are called **mixed predicates**. They are called 'mixed' because they sometimes behave like distributive predicates and sometimes like collective predicates. More on this later.

- Since the analysis in the previous section always derives the distributivity inference, mixed predicates are also problematic.

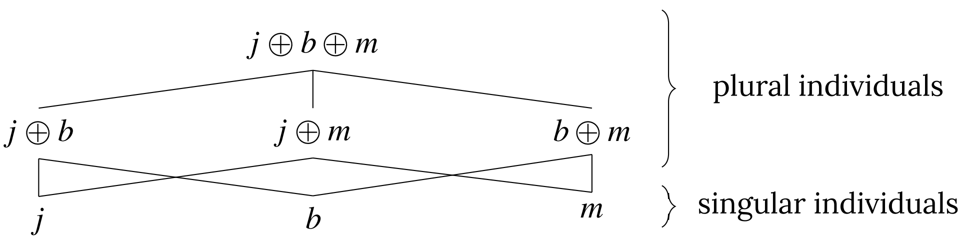
- To sum up, we have the following three types of predicates.
  - **Distributive predicates** (predicates with distributivity inferences)
    - 'John and Mary VP' $\Leftrightarrow$ 'John VP and Mary VP'
  - **Collective predicates** (generally 'John VP and Mary VP' is unacceptable)
    - 'John and Mary VP' $\not\Rightarrow$ 'John VP and Mary VP'
    - 'John and Mary VP' $\not\Leftarrow$ 'John VP and Mary VP'
  - **Mixed predicates**:
    - 'John and Mary VP' $\not\Rightarrow$ 'John VP and Mary VP'
    - 'John and Mary VP' $\Leftarrow$ 'John VP and Mary VP'

5.3 Plural Individuals as 'I(ndividual)-Sums'

- How do we account for collective and mixed predicates, then? We follow the following idea:
  - Plural phrases like *John and Mary* denote **plural individuals**.
  - Collective predicates only take plural individuals as their arguments (We'll come back to mixed predicates later)

- What are plural individuals? They are individuals just like John, Mary, Bill, etc. but unlike these 'simple' individual, they are composed of multiple individuals. For example, the plural individual denoted by *John and Mary* is an individual distinct from John and from Mary, but it has John and Mary as its parts.

- More formally, we model plural individuals as **individual-sums** or **i-sums** for short.
  - Normal individuals like John and Mary are from now on called **singular individuals** (or sometimes **atomic individuals**).
  - The i-sum consisting of John and Mary is represented as 'John$\oplus$Mary'.
- '$\oplus$' is the i-sum forming operator, and it has the following properties. For any individuals (singular or plural) x, y, and z,
  - $x \oplus y = y \oplus x$ (Commutativity)
  - $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ (Associativity)
  - $x \oplus x = x$ (Idempotence)
- Since we have associativity, we often omit parentheses, e.g. John$\oplus$Mary$\oplus$Bill.
- Now we enrich the domain of individuals D_e by 'closing it with $\oplus$'.
  - $\mathcal{D}$, the set of individuals specified by the model $\mathcal{M}$, is a set of singular individuals.
  - We have been assuming so far that for any model $\mathcal{M}$, $D_e = \mathcal{D}$.
  - We change this assumption by putting plural individuals in D_e in addition to the singular individuals in $\mathcal{D}$. That is, for any two members x and y in D_e, we will also have $x \oplus y$ in D_e (this is what it means to 'close' D_e with $\oplus$, i.e. you have all possible plural individuals in D_e based on $\mathcal{D}$).
  - For convenience, we will refer to the set of singular individuals as SG, which is identical to $\mathcal{D}$, and the set of plural individuals as PL. So $D_e = SG \cup PL$.
- Here's a small example with $\mathcal{D} = \{ j, b, m \}$.
  - $D_e = \{ j, b, m, \quad j \oplus b,\ j \oplus m,\ b \oplus m, \quad j \oplus b \oplus m \}$
  - $SG = \{ j, b, m \}$
  - $PL = \{ j \oplus b,\ j \oplus m,\ b \oplus m, \quad j \oplus b \oplus m \}$



- So plural individuals have other individuals as **parts**. We denote the 'part-of' relation by $\sqsubseteq$. As the line in this symbol indicates, we take every individual, singular or plural, to be trivially a part of itself. But only plural individuals have non-trivial parts. For example:

(5.15)a. $j \oplus b \oplus m \sqsubseteq j \oplus b \oplus m$

b. $j \oplus b \sqsubseteq j \oplus b \oplus m$
c. $j \oplus m \sqsubseteq j \oplus b \oplus m$
d. $j \sqsubseteq j \oplus b \oplus m$
e. $j \sqsubseteq j \oplus m$

(5.16)a. $j \not\sqsubseteq b \oplus m$
b. $j \oplus b \not\sqsubseteq j \oplus m$

The part-of relation is represented in the above diagram by the lines.

- So we have:

(5.17)For any assignment a, and for any model $\mathcal{M}$,
a. $[\![\textbf{John}]\!]^{a,\mathcal{M}} = \text{John}$
b. $[\![\textbf{Mary}]\!]^{a,\mathcal{M}} = \text{Mary}$
c. $[\![\textbf{John and Mary}]\!]^{a,\mathcal{M}} = \text{John} \oplus \text{Mary}$

This use of *and* (call it $and_{\langle e,ee \rangle}$) therefore simply denotes $\oplus$ (which is sometimes called 'non-Boolean conjunction').

(5.18)$[\![\textbf{and}_{\langle e,ee \rangle}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ x \oplus y$

(This function is the same thing as '$\oplus$', but we put the λ's here to make the types of the arguments explicit)

- In this setting collective predicates can be understood as simply predicates that are only true of plural individuals.
- Here is an example. Suppose that in $\mathcal{M}_1$, John and Mary look alike (say, because they are siblings), but John and Paul do not. Then, we have:

(5.19)For any assignment a,
a. $[\![\textbf{look alike}]\!]^{a,\mathcal{M}_1}([\![\textbf{John and Mary}]\!]^{a\mathcal{M}_1})$
$= [\![\textbf{look alike}]\!]^{a,\mathcal{M}_1}(\text{John} \oplus \text{Mary}) = 1$
b. $[\![\textbf{look alike}]\!]^{a,\mathcal{M}_1}([\![\textbf{John and Paul}]\!]^{a,\mathcal{M}_1})$
$= [\![\textbf{look alike}]\!]^{a,\mathcal{M}_1}(\text{John} \oplus \text{Paul}) = 0$

Also, since collective predicates are never true of singular individuals, we have the following:

(5.20)For any assignment a, and for any model $\mathcal{M}$,
a. $[\![\textbf{look alike}]\!]^{a,\mathcal{M}}([\![\textbf{John}]\!]^{a,\mathcal{M}}) = [\![\textbf{look alike}]\!]^{a,\mathcal{M}}(\text{John}) = 0$
b. $[\![\textbf{look alike}]\!]^{a,\mathcal{M}}([\![\textbf{Mary}]\!]^{a,\mathcal{M}}) = [\![\textbf{look alike}]\!]^{a,\mathcal{M}}(\text{Mary}) = 0$
c. $[\![\textbf{look alike}]\!]^{a,\mathcal{M}}([\![\textbf{Paul}]\!]^{a,\mathcal{M}}) = [\![\textbf{look alike}]\!]^{a,\mathcal{M}}(\text{Mary}) = 0$

Thus sentences like (5.21) are simply false, according to this analysis.

(5.21)*John looks alike.

- But you might object that sentences like (5.22a) and sentences like (5.22b) seem to have different status. That is, (5.22a) is grammatical and just false (in the model under discussion), while (5.22b) is unacceptable and sounds even ungrammatical.

(5.22) a. John and Paul look alike.
b. *John looks alike.

There are several ways to address this question. Here are two:

- One possibility is that (5.22b) is not only false but also violates the **presupposition** of the predicate *look alike*. Presupposition is a different dimension of meaning from the truth-conditional meaning that we have been discussing. There are different kinds of presupposition, but it is considered that certain predicates put presuppositional constraints on their arguments. For example, *is upset* requires the subject to be a sentient entity. If the subject denotes a non-sentient entity, the sentence is unacceptable, rather than simply false, as in (5.23) (unless you metaphorically personify the subject).

  (5.23) *Binding Condition B is upset.

  So the idea is that a predicate like *look alike* has a presuppositional constraint that the subject is a plural individual. This can differentiate (5.22a) and (5.22b).
- Another possibility is that people judge sentences that are 'analytically false' as unacceptable. According to this idea, the seemingly different status comes from the fact that (5.22a) could be true in a different situation, while (5.22b) can never be true. Cf. the discussion from last week on trivially true sentences such as **Nathan is taller than no girl is*.

• What about mixed predicates? Mixed predicates are those predicates that are compatible with both singular and plural individuals.

(5.24) Suppose that in the situation described by the model $\mathcal{M}_1$, John bought a house on his own, and Mary bought a house on his own, but John and Mary did not buy a house together. Then, we have: (for any assignment a)
a. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_1}(\text{John}) = 1$
b. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_1}(\text{Mary}) = 1$
c. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_1}(\text{John} \oplus \text{Mary}) = 0$

(5.25) Suppose that in the situation described by the model $\mathcal{M}_2$, John and Mary bought a house together, but they did not buy one on their own. Then, we have: (for any assignment a)
a. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_2}(\text{John}) = 0$
b. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_2}(\text{Mary}) = 0$
c. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_2}(\text{John} \oplus \text{Mary}) = 1$

• Now, you might wonder how we get the one-way entailment from (5.26a) to (5.26b).

(5.26) a. John bought a house and Mary bought a house.
b. John and Mary bought a house.

In order to see this problem, assume for the moment that (5.26a) is true iff (5.27a) and (5.27b) are the case, and that (5.26b) is the case iff (5.27c) is the case.

(5.27) a. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_1}(\text{John}) = 1$

b. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_1}(\text{Mary}) = 1$
c. $[\![\textbf{bought a house}]\!]^{a,\mathcal{M}_2}(\text{John} \oplus \text{Mary}) = 1$

But the truth of (5.27a) and (5.27b) does not guarantee the truth of (5.27c)! (and it shouldn't, because we have situations like $\mathcal{M}_2$ above).

- A related issue arises with distributive predicates. Given their meanings, we would want to say that they can only be true of singular individuals. For example, suppose that in the situation described by model $\mathcal{M}_3$, John fell asleep and Mary fell asleep. Then we have:

(5.28) a. $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(\text{John}) = 1$
b. $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(\text{Mary}) = 1$
c. $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(\text{John} \oplus \text{Mary}) = 0$

But then why is the following sentence true in $\mathcal{M}_3$?

(5.29) John and Mary fell asleep.

We will answer this question in the next section.

5.4 Distributivity

- The sentence in (5.30) with a mixed predicate *bought a house* is true in two different kinds of situations.

(5.30) John and Mary bought a house.

That is, (5.30) is true if John bought a house and Mary bought a house, or if John and Mary bought a house together. As explained in the previous section, it is not immediately clear why (5.30) is true in the former kind of situation.

- A key to the solution to this issue comes from a slightly different version of the sentence, (5.31).

(5.31) John and Mary both bought a house.

This sentence is only true if John bought a house and Mary bought a house, and is not true if John and Mary bought a house together but they did not separately. Thus, in a way, the word *both* disambiguates the meaning.

- What is the meaning of *both*? The standard answer is that it is a **distributivity operator**, whose function is the follwoing: in the above sentence, it applies the predicate *bought a house* to each atomic part of John $\oplus$ Mary, i.e. John and Mary, separately. The lexical entry of *both* looks as follows.

(5.32) For any assignment a and for any model $\mathcal{M}$,
$[\![\textbf{both}]\!]^{a,\mathcal{M}} = \lambda P_{\langle e,t\rangle}.\lambda x \in PL.$ for all $y \in SG$ such that $y \sqsubseteq x$, $P(y) = 1$

The variable y here ranges over singular individuals (it is a member of SG) that comprise the plural individual x, and the predicate P is applied to each such y.

- This accounts for the truth-conditions of (5.31):

(5.33)
$$\left[\!\left[[_{\mathbf{S}}\ [_{\mathbf{DP}}\ \textbf{John}\ [\textbf{and}\ \textbf{Mary}]]\ [\textbf{both}\ [_{\mathbf{VP}}\ \textbf{bought}\ [_{\mathbf{DP}}\ \textbf{a}\ \textbf{house}]]]] \right]\!\right]^{a,\mathcal{M}}$$

$$= \left[\!\left[[\textbf{both}\ [_{\mathbf{VP}}\ \textbf{bought}\ [_{\mathbf{DP}}\ \textbf{a}\ \textbf{house}]]] \right]\!\right]^{a,\mathcal{M}} \left(\left[\!\left[[_{\mathbf{DP}}\ \textbf{John}\ [\textbf{and}\ \textbf{Mary}]] \right]\!\right]^{a,\mathcal{M}} \right)$$

$$= \left[\!\left[[\textbf{both}\ [_{\mathbf{VP}}\ \textbf{bought}\ [_{\mathbf{DP}}\ \textbf{a}\ \textbf{house}]]] \right]\!\right]^{a,\mathcal{M}} (\text{John} \oplus \text{Mary})$$

$$= \left[\begin{array}{l} \lambda P_{\langle e,t\rangle}.\lambda x \in PL.\ \text{for all } y \in SG \\ \text{such that } y \sqsubseteq x, P(y) = 1 \end{array} \right] \left(\left[\!\left[[_{\mathbf{VP}}\ \textbf{bought}\ [_{\mathbf{DP}}\ \textbf{a}\ \textbf{house}]] \right]\!\right]^{a,\mathcal{M}} \right) (\text{John} \oplus \text{Mary})$$

$$= \text{for all } y \in SG \text{ such that } y \sqsubseteq \text{John} \oplus \text{Mary},$$
$$\left[\!\left[[_{\mathbf{VP}}\ \textbf{bought}\ [_{\mathbf{DP}}\ \textbf{a}\ \textbf{house}]] \right]\!\right]^{a,\mathcal{M}} (y) = 1$$

$= 1$ iff John bought a house in $\mathcal{M}$ and Mary bought a house in $\mathcal{M}$

- You might have noticed that part of the meaning of *both* should somehow say that the subject plural individual consists of exactly two singular individuals. In fact, with bigger plural individuals, *both* in infelicitous, and either *each* or *all* needs to be used.

(5.34)a. #John, Paul and Mary *both* bought a house.
b. John, Paul and Mary *each* bought a house.
c. John, Paul and Mayr *all* bought a house.

A promising analysis of this restriction on *both* is that it *presupposes* that the subject plural individual has two singular individuals as its components. Since we have not discussed how to formally represent presuppositions, we will leave this aspect of meaning unaccounted for here.

- Now, what about (5.30), which does not contain *both*? We pursue the hypothesis that it is ambiguous between an LF that contains a **covert distributivity operator** Δ and an LF that does not contain one. The former derives the same meaning as (5.33), while the latter is only true if John and Mary bought a house together.

  (5.30)John and Mary bought a house.

  In other words, we assume that a sentence like (5.30) is ambiguous between a **distributive reading** and **collective reading**.

  (5.35)a. John and Mary Δ bought a house. $\Rightarrow$ Distributive reading
  b. John and Mary bought a house. $\Rightarrow$ Collective reading

- Now we account for the entailment from (5.36a) to (5.36b).

  (5.36)a. John bought a house and Mary bought a house.
  b. John and Mary bought a house.

  That is, if (5.36b) is parsed as (5.35a), the entailment goes through. In fact, the entailment goes through in the other direction too, i.e. (5.35a) also entails (5.36b). That (5.36b) also has a parse (5.35b) without Δ, however, makes it look as if the entailment is only one way.

  To put it differently, Δ turns a mixed predicate to a distributive predicate.

- Turning now to distributive predicates, we assume that they are only true of singular individuals. Consequently, when the subject is plural, there is always Δ to make the sentence acceptable.

  (5.37)a. John and Mary Δ fell asleep.
  b. *John and Mary fell asleep.

- Finally, Δ simply cannot apply to a collective predicate, because by assumption a collective predicate cannot be true for singular individuals.

  (5.38)a. John and Mary look a like.
  b. *John and Mary Δ look alike.

- Consequently, each LF is disambiguated.

5.5 Further Readings

Bennett (1974) and Hausser (1974) are the earliest analyses of plurality in the framework of Montague Grammar. These authors analyse plural individuals as sets, rather than i-sums. But such theories require a lot of redundancy in the lexicon for predicates that can apply to both singular and plural arguments. For this reason, Scha (1981) proposes to treat singular and plural individuals on a par, i.e. they are all sets. This is easy to do, the domain of singular individuals and the domain of singleton sets are obviously isomorphic (i.e. they are formally 'identical').

While these authors treat plural individuals as sets (other authors that do so include Landman 1989a,b, 2000, Schwarzschild 1996, Winter 2001a), Link (1983) advocates a mereological approach where the 'part-of' relation is taken to be the primitive and atomic entities are not necessarily required (see Champollion & Krifka 2014 for a linguistically-oriented overview of mereology; Varzi 2015 is also an accessible survey article on this topic). Link makes use of this ontology for the semantics of mass nouns. We will discuss mass nouns in Week 8.

The crucial difference between the set approach and the i-sum approach is that in the set approach, we can talk about sets of sets of individuals, sets of sets of sets of individuals, sets of sets of sets of sets of individuals, etc., while in the i-sum approach, the structure of plural individuals is 'flat', so-to-speak. There is a lot of discussion on whether such extra structure is necessary to account for the meanings of plural nouns phrases in natural language. If you are interested, read Landman (1989a,b) and Schwarzschild (1996), among others.

However, if we are only interested in sets of individuals, there is no formal difference between the set approach and the i-sum approach, because the domain of sets of individuals and the domain of i-sums are isomorphic.

Winter (2001a:Ch.5) (a shorter version appeared as Winter 2002) discusses a different classification of predicates than the three-way classification we discussed above. His idea is motivated by plural vs. singular quantificational phrases, which we will discuss next week.

The distributivity operator Δ was originally put forward by Link (1987) and Roberts (1987). Scha (1981) proposed to build in the distributivity to the lexical entry of predicates, but there are cases involving distributivity at a non-lexical level. See also Landman (2000) and Winter (2001a) for discussion on this. Schwarzschild (1996) discusses cases involving 'intermediate distributivity', which can be understood as distributivity over non-singular parts.

5.6 Exercises

a) Give one example of distributive VPs (e.g. *smoke*), one example of collective VPs (e.g. *look alike*), and one example of mixed VPs (e.g. *wrote a paper*) that are not mentioned above. Motivate your answer with examples demonstrating their behaviour with respect to the distributivity inference.

b) The three-way classification of predicates applies to predicates of other syntactic categories, too. Give one collective AP, one collective PP and one collective NP. Show that they are collective by giving examples demonstrating that the distributivity inference does not hold in either direction.

c) There are transitive predicates (verbs, nouns, adjectives, or presuppositions) that require a plural noun phrase as its object (or complement), e.g. *group X and*

Y together, *marry X and Y* (in the transitive/causative sense of *marry*). These predicates can be said to be collective with respect to the object. Give two other examples of such transitive predicates.

Chapter 6

Plural Nouns

- When we discussed Generalised Quantifier Theory in Term 1, we did not distinguish singular and plural NPs.

  (6.1) For any assignment a and for any model $\mathcal{M}$,
  a. $[\![\textbf{boy}]\!]^{a,\mathcal{M}} = [\![\textbf{boys}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a boy in $\mathcal{M}$
  b. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = [\![\textbf{books}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a book in $\mathcal{M}$

  Or equivalently, in set talk, they are assumed to be both sets of individuals.

  (6.2) For any assignment a and for any model $\mathcal{M}$,
  a. $[\![\textbf{boy}]\!]^{a,\mathcal{M}} = [\![\textbf{boys}]\!]^{a,\mathcal{M}} = \{\, x \in D_e \mid x \text{ is a boy in } \mathcal{M} \,\}$
  b. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = [\![\textbf{books}]\!]^{a,\mathcal{M}} = \{\, x \in D_e \mid x \text{ is a book in } \mathcal{M} \,\}$

  But this assumption is too simplistic, perhaps unsurprisingly.

- For example, the following contrasts illustrate that singular and plural noun phrases have different meanings.

  (6.3) a. John is a boy.
  b. *John and Bill are a boy.

  (6.4) a. *John is boys.
  b. John and Bill are boys.

  One might think that this is a syntactic phenomenon, i.e. the two DPs in predicational sentences need to agree in number. However, number agreement is not always required, as shown by (6.5).

  (6.5) a. John and Bill are a couple.
  b. These assignments are a nightmare.

- Another reason to believe that singular and plural nouns have different semantics comes from the fact that plural noun phrases mean 'plural', i.e. more than one (at least in certain cases; we'll discuss exceptions later).

  For instance, (6.6b) sounds false if John only read one book, unlike (6.6a).

  (6.6) a. John read a book.

b. John read books.

So the plural marking seems to have some meaning.

- The simplest hypothesis is that plural nouns like *books* have plural individuals and only plural individuals in their denotation, and singular nouns like *book* only has singular individuals in their denotation, as in (6.7).

(6.7) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\ x$ is a book in $\mathcal{M}$
b. $[\![\textbf{books}]\!]^{a,\mathcal{M}} = \lambda x \in PL.$ each singular part y of x is a book in $\mathcal{M}$

Or in set talk:

(6.8) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \{ b_1, b_2, b_3 \}$
b. $[\![\textbf{books}]\!]^{a,\mathcal{M}} = \left\{ \begin{array}{c} b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\}$

However, it turns out that the semantics of plural nouns is not that straightforward.

6.1 Plural is Unmarked

6.1.1 Singular Nouns

- It seems that we do not have to revise our analysis for singular nouns. That is, they denote functions of type $\langle e,t \rangle$ that are true of singular individuals.
- This semantics gives a (partial) explanation as to why singular nouns cannot be true of plural individuals.

(6.9) *John and Bill are a boy.

That is, John$\oplus$Bill is never in the extension of a singular noun *boy*.

- **Excursus**: You might wonder why the distributivity operator Δ cannot be used to make the sentence true. In fact, for sentences like (6.10), we postulate Δ to account for the distributive reading, which entails that two beers were ordered.

(6.10)John and Bill Δ(ordered a beer).

The meaning of Δ is:

(6.11)For any assignment a and for any model $\mathcal{M}$,
$[\![\Delta]\!]^{a,\mathcal{M}} = \lambda P_{\langle e,t \rangle}.\lambda x \in PL.$ for all $y \in SG$ such that $y \sqsubseteq x$, $P(y) = 1$

In words, Δ applies the VP denotation to each singular part of the subject (but recall the complication that Patrick mentioned last time regarding the landing site of the object quantifier).

If this operator can appear in the predicational copula construction, we do indeed predict (6.9) to be fine. That is, (6.12) is true iff John is a boy and Bill is a boy.

(6.12)*John and Bill Δ(are a boy).

We could stipulate a constraint that prohibits Δ in predicational sentences, but why such a constraint exists needs to be explained.

6.1.2 Plural Nouns

- There are two possible analyses for plural nouns.
  1. Plural means more than one (> 1)

     (6.13) For any assignment a and for any model $\mathcal{M}$,
     $$[\![\textbf{books}]\!]^{a,\mathcal{M}} = \left[\lambda x \in D_e.\ \begin{array}{l} x \text{ is a plural individual} \\ \text{each of whose singular part is a book in } \mathcal{M} \end{array}\right]$$

     In set talk:

     (6.14) For any assignment a and for any model $\mathcal{M}$,
     $$[\![\textbf{books}]\!]^{a,\mathcal{M}} = \{\, x \in PL \mid \text{each singular part of } x \text{ is a book in } \mathcal{M} \,\}$$

  2. Plural means one or more (> 0)

     (6.15) For any assignment a and for any model $\mathcal{M}$,
     $$[\![\textbf{books}]\!]^{a,\mathcal{M}} = \left[\lambda x \in D_e.\ \begin{array}{l} x \text{ is a book or a plural individual} \\ \text{each of whose singular part is a book in } \mathcal{M} \end{array}\right]$$

     In set talk:

     (6.16) For any assignment a and for any model $\mathcal{M}$,
     $$[\![\textbf{books}]\!]^{a,\mathcal{M}} = \{\, x \in SG \mid x \text{ is a book in } \mathcal{M} \,\} \cup \left\{\, x \in PL \;\middle|\; \begin{array}{l} \text{each singular part} \\ \text{of } x \text{ is a book in } \mathcal{M} \end{array} \right\}$$

- At first sight, the first option seems to be better. However, there are arguments for the second analysis. That is, there are some cases where the plural means one or more (> 0), rather than more than one (> 1).

  - Plural indefinites in questions

    (6.17) Do you have children?
    a. Yes, I have one.
    b. #No, I (only) have one.

    This question is neutral with respect to the number. Compare this to (6.18):

    (6.18) Do you have two or more children?
    a. #Yes, I have one.
    b. No, I (only) have one.

    These two questions clearly have different meanings.

  - Plural indefinites in negative sentences

    (6.19) John doesn't have children.

    This sentence entails that John does not have a child. Again, compare this to (6.20), which does not entail it.

(6.20) John doesn't have two or more children.

– Plural indefinites in conditionals

(6.21) If you have coins in your pocket, put them in a tray.

This sentence is number neutral in the sense that it does not exclude situations where you only have one coin in your pocket. Compare this to (6.22), which does exclude such situations.

(6.22) If you have two or more coins in your pocket, put them in a tray.

– Plural definites in ignorance situations

Consider the following scenario (this example is taken from Sauerland, Anderssen & Yatsushiro 2005):

(6.23) You are inviting an old friend who you have not seen in years. you heard that he has a family now, but you have no idea how many children he has.

In this scenario, it is more natural to use a plural:

(6.24) a. You are welcome to bring your children.
b. #You are welcome to bring your child.

Similarly, *two or more* would be strange in this context.

(6.25) #You are welcome to bring your two or more children.

Rather, (6.24a) is closer to *child or children*.

(6.26) You are welcome to bring your child or children.

- The above observations lead us to assume that plural includes singular individuals. In set talk:

(6.27) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \{\, b_1, b_2, b_3 \,\}$

b. $$[\![\textbf{books}]\!]^{a,\mathcal{M}} = \left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \\ b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\}$$

So the plural does not mean more than one (> 1), but rather, it is number neutral!

- What's interesting is that the plural does mean more than one in other cases like the following examples.

(6.28) a. John has children.
b. I like Paul's books on semantics.

(6.28a) entails that John has more than one child, and the possessive DP in (6.28b) refers to more than one book on semantics. Let us call these inferences that there is more than one individual **plurality inferences**.

- The question is: when do we get the plurality inference and when do we not? And why?

- Here one way to account for the distribution of the plurality inference. Let us postulate the following rule (see Sauerland et al. 2005 for a more precise formulation of this; see also Pearson, Khan & Snedeker 2010)

  (6.29) **Unmarked Plural Rule (UPR)**
  If you mean 'exactly one' you cannot use the plural.

- Let's go through an example to see how the plurality inference arises. Consider (6.30).

  (6.30) John has children.

  According to our number neutral semantics for plural nouns, (6.30) means: John has one or more children. The UPR demands that if you want to mean 'John has exactly one child', you cannot use this sentence. Conversely, you can use (6.30) if you do not mean 'John has exactly one child'. Thus, together with the meaning of the sentence (John has at least one child), it follows that John has two or more children.

  Incidentally, the UPR says nothing about situations where you do *not* mean 'exactly one'. So in such contexts, you can use (6.30) or (6.31).

  (6.31) John has a child.

  (You might think (6.31) sounds like John only has one child, but this does not generalise too all situations. The 'exactly one' inference might be a scalar implicature.)

- Let's now go through cases without plurality inferences.

  - <u>Plural indefinites in questions</u>

    (6.32) Does John have children?

    The UPR demands that if you want to mean 'exactly one', i.e. 'Does John have exactly one child'?, you cannot use (6.32). In all other contexts, you can use (6.32), including when you want to ask 'Does John have one or more children?'.

    Again, if you do not mean 'exactly one', the UPR has nothing to say. In particular, it does not prevent you from using (6.33) to mean the same thing as (6.32).

    (6.33) Does John have a child?

  - The same reasoning applies to plural indefinites in negative sentences and conditionals. Let us take conditionals (the negation example is left for an exercise).

    (6.34) If you have coins in your pocket, put them in a tray.

    The UPR says, if you want to mean 'If you have exactly one coin in your pocket, put it in a tray', you cannot use (6.34). So you can use (6.34) to mean 'If you have one or more coins in your pocket, put them in a tray'.

    Again, there are no restrictions on the singular counterpart, so nothing prevents (6.35) from meaning the same thing.

(6.35) If you have a coin in your pocket, put it in a tray.

– Plural definites in ignorance contexts

(6.36) You are welcome to bring your children.

If you know that your old friend has exactly one child, the UPR says you cannot use (6.36), because you would mean 'You are welcome to bring your exactly one/sole child'. But in all other contexts, including contexts where you do not know how many children your friend has, you can use (6.36) to mean 'You are welcome to bring your child or children'.
Again, the UPR says nothing about (6.37).

(6.37) You are welcome to bring your child.

But (6.37) is only fine in contexts where your friend has exactly one child, due to the meaning of the singular noun and the definiteness (we'll come back to definiteness below).

- To sum up, in some cases the plural is number neutral but in other cases it gives rise to a plurality inference ('more than one'). We assume a number neutral meaning of the plural and derive the plurality inference via the Unmarked Plural Rule (UPR).

  In set talk, the denotation of the nouns look like (6.38).

(6.38) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \{\, b_1, b_2, b_3 \,\}$

b. $[\![\textbf{books}]\!]^{a,\mathcal{M}} = \left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \\ b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\}$

In function talk:

(6.39) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\ x$ is a book in $\mathcal{M}$

b. $[\![\textbf{books}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ for each $y \in SG$, if $y \sqsubseteq x$, y is a book in $\mathcal{M}$

(6.40) **Unmarked Plural Rule (UPR)** If you mean 'exactly one' you cannot use the plural noun.

- We can assign the following meaning to the plural morpheme. It takes the denotation of a singular noun and 'closes it with $\oplus$':

(6.41) $[\![\textbf{-s}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda x \in D_e.$ for each $y \in SG$, if $y \sqsubseteq x$, $P(x) = 1$

  Nothing prohibits the second application of (6.41), but it won't change the meaning.

- For the rest of today's class, we will discuss the meanings of various DPs, assuming the above semantics of singular and plural NPs.

- Suppose that there are three books, b_1, b_2 and b_3 (call this situation $\mathcal{M}_1$). Then, the phrase *the book* is infelicitous, while *the books* denotes the plural individual consisting of these three books.

(6.42)a. $[\![\textbf{books}]\!]^{a,\mathcal{M}_1} = \left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \\ b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\}$

b. $[\![\textbf{the books}]\!]^{a,\mathcal{M}_1} = b_1 \oplus b_2 \oplus b_3$

- Suppose now that there is only one book b_1 (call this situation $\mathcal{M}_2$). Then, the phrase **the book** denotes this unique book, while **the books** is infelicitous, because of the UPR.

(6.43)a. $[\![\textbf{book}]\!]^{a,\mathcal{M}_2} = [\![\textbf{books}]\!]^{a,\mathcal{M}_2} = \{ b_1 \}$

b. $[\![\textbf{the book}]\!]^{a,\mathcal{M}_2} = b_1$

- Generally, ‘the NP(s)’ denotes the **unique maximal individual** satisfying the NP, if any.
  - In $\mathcal{M}_1$, $[\![\textbf{books}]\!]^{a,\mathcal{M}_1}$ has a unique maximal individual, $b_1 \oplus b_2 \oplus b_3$. It’s maximal in the sense that everything else in $[\![\textbf{books}]\!]^{a,\mathcal{M}}$ is a part of it.
  - On the other hand, $[\![\textbf{book}]\!]^{a,\mathcal{M}_1}$ does not have a maximal individual in $\mathcal{M}_1$, because there are three independent books.

    (6.44) $[\![\textbf{book}]\!]^{a,\mathcal{M}_1} = \{ b_1, b_2, b_3 \}$

    Consequently *the book* has nothing to denote (we call such a case a presupposition failure; see discussion in Heim & Kratzer 1998:§4.4 for more on this).
  - In $\mathcal{M}_2$, $[\![\textbf{book}]\!]^{a,\mathcal{M}_2}$ does have a unique maximal individual, namely b_1, so it denotes it.
  - As noted above, *the books* is infelicitous in $\mathcal{M}_2$ because you would mean ‘the exactly one book’, and violate the UPR.
- The unique maximal individual with respect to a noun denotation P is denoted by $\sigma(P)$ (or the supremum of P). We can simply analyse the meaning of *the* as this operator.

(6.45) For any model $\mathcal{M}$, and for any assignment a,

$$[\![\textbf{the}]\!]^{a,\mathcal{M}} = [\lambda P \in D\langle e,t\rangle.\ \sigma(P)] = \sigma$$

Crucially, the σ-operator is not defined for all predicates. It requires there to be a unique maximal element (a supremum) in the predicate.

(6.46) σ is defined for $P \in D_{\langle e,t\rangle}$ only if there is a unique maximal element in P
i.e. there is x such that $P(x) = 1$ and for all y such that $P(y) = 1$, $y \sqsubseteq x$.

(6.47) Whenever defined, $\sigma(P)$ is the unique maximal element.

(In the literature, the definedness condition like (6.46) is treated as a presupposition. Again, see Heim & Kratzer 1998:§4.4 for more on this)

6.3 Generalised Quantifiers with Plurality

- Some quantificational determiners combine with a singular noun:

  (6.48) a. every book
  b. *every books

  Others select for a plural noun:

  (6.49) a. *most book
  b. most books

  Others are neutral:

  (6.50) a. some book
  b. some books

- The analysis of quantifiers we discussed in Term 1 did not distinguish singular and plural noun phrases. Let us modify Generalise Quantifier Theory to incorporate the semantics of plural noun phrases developed above.

- <u>Indefinites</u>

  According to the generalised quantifier analysis of indefinites, a singular indefinite like *a book* is an existential quantifier:

  (6.51) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{a book}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}$. there is a book x in $\mathcal{M}$ such that $P(x) = 1$

  The meaning of the determiner is:

  (6.52) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{a}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}$. there is $x \in D_e$ such that $Q(x) = P(x) = 1$
  (or in set talk: $[\![\textbf{a}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ Q \cap P \neq \varnothing$)

  We gave the same meaning to *some*:

  (6.53) For any model $\mathcal{M}$ and for any assignment a,
  a. $[\![\textbf{some book}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}$. there is a book x in $\mathcal{M}$ such that $P(x) = 1$
  b. $[\![\textbf{some}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}$. there is $x \in D_e$ such that $Q(x) = P(x) = 1$
  (or in set talk: $[\![\textbf{some}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ Q \cap P \neq \varnothing$)

  We can use the same meaning for *some* to account for plural indefinites:

  (6.54) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{some books}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}$. there is $x \in D_e$ such that $[\![\textbf{books}]\!]^{a,\mathcal{M}}(x) = P(x) = 1$

$$= \lambda P \in D_{\langle e,t\rangle}. \left[\begin{array}{l}\text{there is } x \in D_e \text{ such that } x \text{ is a book in } \mathcal{M} \\ \text{or each singular part of } x \text{ is a book in } \mathcal{M} \\ \text{and } P(x) = 1\end{array}\right]$$

This analysis accounts for the following sentences (recall that the UPR prohibits the 'exactly one' meaning).

(6.55)a. Some children Δ cried. (Distributive)
b. Some children gathered in the park. (Collective)

Notice in particular, since $[\![\textbf{children}]\!]^{a,\mathcal{M}}$ includes plural individuals, collective predication is made possible, (6.55b). That collective predication is incompatible with a singular indefinite is also as expected.

(6.56)a. *A child gathered in the park.
b. *Some child gathered.

Since $[\![\textbf{child}]\!]^{a,\mathcal{M}}$ does not include a plural individual in its extension, the sentence cannot be true (as we discussed last time, this can be though of as a purely semantic anomaly or syntactic anomaly, or both).

We can give the same meaning to bare plural indefinites:

(6.57)a. Children Δ cried.
b. Children gathered in the park.

That is, we can assume a null determiner ∃ in these cases, which has the same meaning as *some* (and is realised overtly in other languages like French and Spanish).

(6.58) $[\![\exists]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.$ there is $x \in D_e$ such that $Q(x) = P(x) = 1$
$= [\![\textbf{some}]\!]^{a,\mathcal{M}}$
(or in set talk: $[\![\exists]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ Q \cap P \neq \varnothing$)

Side note: However, there are several differences between *some* NPs and bare plurals. Firstly, bare plurals always take narrow scope, while *some* NPs is a so-called Positive Polarity Item and cannot take scope below a clause-mate negation. So the following two sentences are not synonymous.

(6.59)a. John didn't read some books.
b. John didn't read books.

Secondly, bare plurals have kind readings, while *some* NPs only have sub-kind readings.

(6.60)a. Some dogs are mammals.
b. Dogs are mammals.

Thirdly, only *some* can appear in partitive DPs.

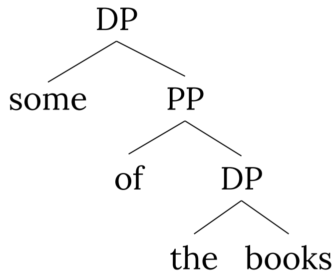
(6.61)a. Some of the books are interesting.
b. *Of the books are interesting.

Some argue that bare plurals do not actually involve determiners (Carlson 1977, Chierchia 1998b). See Carlson (1977), Chierchia (1998b), Chung & Ladusaw (2003), Diesing (1992) and Van Geenhoven (1998) for more on bare plurals and other bare NPs, and cross-linguistic facts.

- Partitive Indefinites

  Some of the books is also an existential quantifier, but it contains a definite DP. Let's assume the following LF for this DP.

  (6.62)

  Recall from the previous section that *the books* denotes the unique maximal individual:

  (6.63) $[\![\textbf{the books}]\!]^{a,\mathcal{M}} = \sigma([\![\textbf{books}]\!]^{a,\mathcal{M}})$

  Since *some* requires a predicate of type $\langle e,t\rangle$ as its argument, let's assume that *of* turns the (plural) individual in (6.63) into predicates in the following manner:

  (6.64) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{of}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ \lambda y \in D_e.\ y \sqsubseteq x$

  Assuming that y in (6.64) is not restricted to singular individuals, we have:

  (6.65) a. $[\![\textbf{of the books}]\!]^{a,\mathcal{M}} = [\![\textbf{books}]\!]^{a,\mathcal{M}}$
  b. $[\![\textbf{some of the books}]\!]^{a,\mathcal{M}} = [\![\textbf{some books}]\!]^{a,\mathcal{M}}$

  Side note: As Patrick mentioned in class, the semantics of *each of the books* might require some assumptions, as *each* can only combine with a singular NP in non-partitive constructions. One possibility is that there is a hidden singular NP *each ~~book~~ of the books*. Another possibility is that *each* quantifies over singular individuals.

  (6.66) For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{each}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.$ for each $x \in SG$ such that $Q(x) = 1$, $P(x) = 1$

  So even if the noun is plural as in *each of the books*, it effectively quantifies over singular books. Notice that **each books* is ruled out by the UPR.

- No is another number neural determiner.

  (6.67) a. no book
  b. no books

  In Term 1, we gave the following analysis, which is the negation of *some*:

(6.68) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{no}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\text{ there is no } x \in D_e \text{ such that } Q(x) = P(x) = 1$$

(or in set talk: $[\![\textbf{no}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ Q \cap P = \emptyset$)

This will work as is with plural noun phrases.

(6.69) a. No students Δ were late for class.
b. No students gathered.

(Notice that this is another case where the plural should not mean 'more than one')

- Unlike *some* and *no*, *all* gives rise to a problem. In Term 1, we assigned the same meanings to *all* and *each*, i.e. they express the subset relation.

(6.70) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{each}]\!]^{a,\mathcal{M}} = [\![\textbf{all}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{for each } x \in D_e \\ \text{such that } Q(x) = 1,\ P(x) = 1\end{array}\right]$$

(or in set talk: $[\![\textbf{each}]\!]^{a,\mathcal{M}} = [\![\textbf{all}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.Q \subseteq P$)

One notable difference between *each* and *all* is number marking.

(6.71) a. each book
b. *each books

(6.72) a. *all book
b. all books

Their difference goes beyond their morphosyntactic properties. Only *all* is compatible with collective predication.

(6.73) a. *Each student gathered.
b. All (of) the students gathered.

(As discussed in class, *every* is compatible with collective predicates, e.g. *every student gathered.*)

The above subset-meaning works for *every* (assuming that it selects for singular nouns). But for *all*, it predicts too strong a meaning. For instance, (6.73b) is predicted to be true iff $[\![\textbf{of the students}]\!]^{a,\mathcal{M}}$ is a subset of $[\![\textbf{gathered}]\!]^{a,\mathcal{M}}$. Suppose that there are four students, $s_1, s_2, \ldots, s_4$ and there is only one gathering by $s_1 \oplus s_2 \oplus s_3 \oplus s_4$. Then, in set talk, we have:

(6.74) a.
$$[\![\textbf{students}]\!]^{a,\mathcal{M}} = \left\{\begin{array}{c} s_1,\quad s_2,\quad s_3,\quad s_4, \\ s_1 \oplus s_2,\quad s_1 \oplus s_3,\quad s_1 \oplus s_4,\quad s_2 \oplus s_3,\quad s_2 \oplus s_4,\quad s_3 \oplus s_4, \\ s_1 \oplus s_2 \oplus s_3,\quad s_1 \oplus s_2 \oplus s_4,\quad s_1 \oplus s_3 \oplus s_4,\quad s_2 \oplus s_3 \oplus s_4, \\ s_1 \oplus s_2 \oplus s_3 \oplus s_4 \end{array}\right\}$$

b.
$$[\![\textbf{gathered}]\!]^{a,\mathcal{M}} = \{\, s_1 \oplus s_2 \oplus s_3 \oplus s_4 \,\}$$

Intuitively, the sentence (6.73b) is true, but (6.74a) is not a subset of (6.74b)! In fact, it is predicted that (6.73b) can never be true, because a collective predicate is never true of singular individuals (recall the discussion from last time).

One possibility is that *all* is not a quantifier after all, as Brisson (2003) proposes. The idea is that *all (of) the students* denotes the maximal individual, just like *the students*. Then, the sentence is correctly predicted to be true in the scenario depicted in (6.74).

But of course there is a difference between *all of the students* and *the students*. One such difference is the (in)tolerance of exceptions. It is widely observed that definite descriptions can be used somewhat loosely. For instance, (6.75) sounds true, even when, say, 5 of the 20,000 students are not satisfied.

(6.75)The students are satisfied.

But with *all* the sentence does sound false.

(6.76)All the students are satisfied.

Brisson (2003) claims that the elimination of this looseness is the meaning of *all*. We will not try to formalise this here, but ultimately, the idea relates to the theory of *pragmatic slacks* mentioned in the first half of the term. See Lasersohn (1999), among others, for useful discussion.

However, observe the following contrast:

(6.77)a. The students are numerous.
b. *All the students are numerous.

If we analyse *all the* NP as the same thing as *the* NP, we cannot account for this contrast. Notice that (6.77) implies that there are two types of collective predicates, those that are compatible with quantifiers and those that are not. For more on this issue and its theoretical consequences, see Dowty (1987) and Winter (2001a, 2002)

6.4 Numerals

In Term 1, we analysed numerals as quantificational determiners:

(6.78)For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{there are three individuals } x\\ \text{such that Q(x)=P(x)=1}\end{array}\right]$

(In set talk: $[\![\textbf{three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ |P \cap Q| = 3$)

b. $[\![\textbf{at least three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{there are at least three individuals } x\\ \text{such that Q(x)=P(x)=1}\end{array}\right]$

(In set talk: $[\![\textbf{at least three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ |P \cap Q| \geqslant 3$)

c. $[\![\textbf{at most three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{there are at most three individuals } x\\ \text{such that Q(x)=P(x)=1}\end{array}\right]$

(In set talk: $[\![\textbf{at most three}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ |P \cap Q| \leqslant 3$)

But numerals are morphosyntactically not determiners, since they can co-occur with *the*, e.g. *the three books I read*, etc.

So it makes sense to analyse numerals as modifiers of type $\langle e,t \rangle$, which combines with $[\![\textbf{books}]\!]^{a,\mathcal{M}}$ via Predicate Modification. For instance, we assume the following LF for *three books*. $\exists$ is the null existential determiner, synonymous with *some*, mentioned above.

(6.79) [$_{\text{DP}}$ $\exists$ [$_{\text{NP}}$ three books]]

Recall that *books* is true of any individual whose singular parts are books (including singular books). We assume that *three* sieves out those individuals that are not comprised of three books:

(6.80) For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{three books}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has three singular parts and $[\![\textbf{books}]\!]^{a,\mathcal{M}}(x) = 1$

b. $[\![\textbf{three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has three singular parts

For instance, suppose that there are four books in $\mathcal{M}_4$. Then,

(6.81) a. $$[\![\textbf{books}]\!]^{a,\mathcal{M}_4} = \left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \quad b_4, \\ b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_1 \oplus b_4, \quad b_2 \oplus b_3, \quad b_2 \oplus b_4, \quad b_3 \oplus b_4, \\ b_1 \oplus b_2 \oplus b_3, \quad b_1 \oplus b_2 \oplus b_4, \quad b_1 \oplus b_3 \oplus b_4, \quad b_2 \oplus b_3 \oplus b_4, \\ b_1 \oplus b_2 \oplus b_3 \oplus b_4 \end{array} \right\}$$

b. $$[\![\textbf{three books}]\!]^{a,\mathcal{M}_4} = \{\, b_1 \oplus b_2 \oplus b_3, \quad b_1 \oplus b_2 \oplus b_4, \quad b_1 \oplus b_3 \oplus b_4, \quad b_2 \oplus b_3 \oplus b_4 \,\}$$

Then,

(6.82) $[\![\exists\ \textbf{three books}]\!]^{a,\mathcal{M}_4} = \lambda P \in D_{\langle e,t \rangle}.$ there is a member of (6.81b) for which P is true.

Notice that the definite *the three books* will be infelicitous in $\mathcal{M}_4$, because *three books* has no unique maximal element. In fact, *the three books* is only felicitous in contexts where there are exactly three books. This is a good result.

It should also be noticed that what we derive for sentences like (6.83) is an 'at-least' reading.

(6.83) $\exists$ three students Δ were late for class.

According to our semantics, (6.83) is true iff there is a plural individual consisting of three students and each of these students was late for class. This is going to be true in a situation where there are four or more students who were late, because in such a situation you can just take three of them to make the sentence true!

This is not a problem because it's possible that the 'exact'-reading is a scalar implicature. That is, when somebody asserts (6.83), you compare it to the versions of sentences with different numerals:

(6.84)a. $\exists$ four students Δ were late for class.
b. $\exists$ five students Δ were late for class.
c. $\exists$ six students Δ were late for class.
...

Since these sentences are stronger, i.e. they asymmetrically entail (6.83), you conclude that they are not true. Therefore, exactly three students were late for class.

Notice that you do not want to derive scalar implicatures for collective sentences. In order to see this, consider (6.85), where *combined* is collective with respect to its object.

(6.85)John combined $\exists$ three PDFs (into a single PDF).

This sentence does not implicate that he did not make another PDF with four or more PDFs, which you would infer by negating the alternatives in (6.86).

(6.86)a. John combined $\exists$ four PDFs.
b. John combined $\exists$ five PDFs.
...

Importantly, with collective predicates like *combine*, these sentences do not stand in an entailment relation. That is, (6.86a) does *not* entail (6.85). If you believe that scalar implicatures are only generated based on alternatives that asymmetrically entail what is uttered, the incorrect inference is blocked.

Notice furthermore, that this means that we do not have a 'at least' reading for (6.85), because in a situation where John combined four PDFs, (6.85) is simply not true. It is only true in a situation where John combined a plurality that has three parts, each of which is a PDF file.

One remaining problem is modified numerals, however. The following semantics predicts the wrong readings.

(6.87)For any model $\mathcal{M}$ and for any assignment a,
a. $[\![\textbf{at least three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has at least three singular parts
b. $[\![\textbf{exactly three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has exactly three singular parts
c. $[\![\textbf{at most three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has at most three singular parts

(6.87a) actually is not a bad analysis, but there will be an open question regarding the scalar implicatures of distributive sentences. In order to see this, consider (6.88).

(6.88)$\exists$ at least three students Δ were late for class.

This is true iff there is a plural individual consisting of three or more singular parts that are all students, and each of them was late for class. Notice that this has exactly the same truth-conditions that we assigned to (6.89).

(6.89)$\exists$ three students Δ were late for class.

But there is a notable difference, i.e. (6.88) does not have an 'exact'-reading, unlike (6.89). Or to put it differently, (6.88) has no scalar implicatures. Why this is the case is left unsolved here (but see Büring 2009 and Fox & Hackl 2006 for some ideas).

For *exactly three* and *at most three*, the above analysis predicts truth-conditions that are too weak (this problem is sometimes called **Van Bentham's trap**). Let us illustrate the problem with *at most three* (the problem for *exactly three* is left for an exercise). The following sentence is true iff there is a plural individual consisting of at most three students such that each of them was late for class.

(6.90) $\exists$ at most three students Δ were late for class.

But notice that if four (or more) students were late for class, you can always find a plural individual consisting of three or fewer students who were late. Then it is predicted that the sentence is true in such a situation!

The above semantics also becomes problematic with collective predicates.

(6.91) $\exists$ at most three students gathered.

This sentence is predicted to be true in the following scenario: there are two gatherings, one by $s_1 \oplus s_2$ and another by $s_3 \oplus s_4 \oplus s_5 \oplus s_6$. In this context, there is a plural individuals consisting of three or fewer students , namely $s_1 \oplus s_2$, that gathered. But intuitively the sentence is false.

However, this might not be a problem. A recent work by Marty, Chemla & Spector (2015) found that in certain experimental settings, the seemingly problematic reading is indeed detected. See Spector (2014) for a pragmatic explanation of this, and related discussion.

6.5 Further Readings

For the unmarkedness of plural noun phrases, see Sauerland (2003), Sauerland et al. (2005), Pearson et al. (2010). Farkas & De Swart (2010) pursue a different analysis where the plural is not unmarked. There are also some works on the meanings of plural indefinites in particular: Spector (2007), Zweig (2009), and Ivlieva (2013). See Heim (2008) and Sauerland (2008) for related discussion. Among these papers, you should find Sauerland et al. (2005) and Sauerland (2008) particularly accessible.

For the semantics of quantifiers, see Scha (1981), Van der Does (1993), and Winter (2001a, 2002). These tend to be a little complicated. In addition, I find Van den Berg (1996) very useful in understanding the meanings of plural quantifiers, but it is highly technical (in part because it deals with dynamic semantics).

For bare plurals, Carlson (1977), Chierchia (1998b), Chung & Ladusaw (2003), Diesing (1992) and Van Geenhoven (1998) are major works, as mentioned above. There are also some overview articles: Delfitto (2005) and Dayal (2011).

There is a lot of recent studies on numerals, especially modified numerals, starting from Krifka (1999). See for instance, Hackl (2000), Takahashi (2006), Nouwen (2010), Geurts & Nouwen (2007), and the works cited in these papers. There is also a very useful overview article, Spector (2014). Also see Marty et al. (2015) for an experimental work on the 'weak' reading of certain modified numerals mentioned above.

6.6 Exercises

1. The UPR and negative sentences

   Assuming the analysis of plural nouns phrases developed in §6.1, explain why the plurality inference is blocked in a negative sentence like (6.92).

   (6.92) John does not have children.

2. Van Bentham's Trap with *exactly three*

   Assume the following modifier meaning for *exactly three*:

   (6.93) $[\![\textbf{exactly three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has exactly three singular parts

   Explain why it works for the collective sentence in (6.94a) but not for the distributive sentence in (6.94b):

   (6.94) a. $\exists$ exactly three students gathered.
   b. $\exists$ exactly three students Δ are blond.

   In showing the problem of (6.94b), give an example situation where the intuitions and the predictions do not match up.

3. Proportional Quantifiers

   In Term 1, we treated *most* as a quantificational determiner with the following meaning (maybe the proportion should be higher than $\frac{1}{2}$ and also vague, but we simplify this aspect of the meaning here):

   (6.95) For any model $\mathcal{M}$ and for any assignment a,
   $$[\![\textbf{most}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\left[\begin{array}{l}\text{for most individuals } x\\ \text{such that } Q(x) = 1,\ P(x) = 1\end{array}\right]$$
   (In set talk: $[\![\textbf{most}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ \dfrac{|Q \cap P|}{|Q|} > \dfrac{1}{2}$)

   However, it turns out that this meaning predicts the wrong truth-conditions for sentences with collective predicates like (6.96).

   (6.96) Most students surrounded Chandler House.

   Here, *surrounded Chandler House* is a collective predicate (as you cannot surround a building alone).

   Assuming the semantics of plural nouns developed above, explain why the above analysis of *most* is wrong using the sentence (6.96) (and other relevant examples, if you need any). In particular, discuss what the predicted truth-conditions are and describe an example situation where the prediction and the intuition diverge.

4. Optional exercise (This is challenging)

   Propose a semantics of *most* that works both for distributive and collective predicates.

Chapter 7

Mass Nouns and Group Nouns

7.1 Mass Nouns

We have so far only considered **count nouns**, which are nouns that have a singular and plural form. English and many other languages also have **mass nouns**, which have different morphosyntactic properties from count nouns. Interestingly, the morphosyntactic mass-count distinction seems to correlate with their semantic properties. Let us review the major problems posed by mass nouns.

Firstly, what are mass nouns? The most reliable way of distinguishing mass nouns from count nouns is by looking at their morphosyntactic differences. Since the following diagnostics refer to the morphosyntactic properties of certain expressions in English, they are largely language specific. However, you can often find similar properties, at least in other European languages.

1. Mass nouns, but not count nouns, can be used without a determiner or a plural morphology:

   (7.1) a. I didn't receive **mail**.
   b. *I didn't receive **letter**.

   (NB: Some count nouns have identical singular and plural forms, some with two plural forms, *sheep*, *deer*, *fish*, *shrimp*, etc., but they are still count nouns, as they exhibit all the other features of count nouns)

2. Count nouns, but not mass nouns, have plural forms.

   (7.2) a. *I received **mails**.
   b. I received **letters**.

3. Some determiners/modifiers have two forms, one for count nouns, one for mass nouns.

   (7.3)

	Mass	*Count*
	much	many
	little	few

Not all expressions of this category have two forms. For instance, *some*, *any*, *the*, *no*, *a lot*, *more*, and *less* are insensitive to the mass-count distinction and compatible with both kinds of nouns. On the other hand *a*, *each* and *every* are only compatible with (singular) count nouns.

4. Cardinal expressions can directly combine with count nouns, but not with mass nouns. Mass nouns require so-called 'classifier expressions', e.g. *piece of*.

(7.4) a. *One mail
b. One letter
c. One piece of mail

One thing to notice at this point is that most nouns have both mass and count 'modes'. For example, *beer* is canonically a mass noun, but in a bar/restaurant, it can be used as a count noun. In such contexts, counting is based on the portion per serving.

(7.5) a. John is drinking **beer**.
b. We've just ordered six **beers**.

(7.6) I ordered a pizza, not a slice of pizza! (Gillon 1999:57)

This is not limited to bar/restaurant contexts:

(7.7) a. Kim produces **sculpture**.
b. Kim is producing **a sculpture**. (Pelletier 2012:14)

Conversely, *apple* is typically a count noun, but it admits a mass use as in (7.8b).

(7.8) a. This salad contains **a lot of apple**.
b. John ate **two apples**.

Some more examples of canonically count nouns used as mass nouns:[1]

(7.9) a. Leslie has more **car** than garage.
b. He's got **woman** on his mind. (Pelletier 2012:14)

(7.10)a. Bill got a lot of **house** for $100,000.
b. How much **floor** did you lay today? (Gillon 1999:58)

Notice that there seems to be some semantic change between the two modes. This intuition is expressed by the Mapping Hypothesis in (7.11) (which is often attributed to Quine 1960; see also Landman 1989a,b, Link 1983 for similar ideas).

(7.11) **Mapping Hypothesis:**
Mass nouns describe stuff/substance. Count nouns describe concrete,

[1]Gillon (1999) says there are several sub-regularities. Nouns denoting animals (e.g. *chicken*, *duck*, *lamb*), plants (e.g. *potato*, *turnip*, *rutabaga*) can be used as mass nouns to denote aggregates of their parts that are suitable for human consumption. Nouns denoting trees (e.g. *oak*, *maple*, *birch*) can be used as mass nouns to denote aggregates of these parts useful for human use. Also, those nouns that denote products can be used to denote their parts that contribute to the enlargement or enhancement of the products.

discrete objects.

While this might look like a plausible principle behind the mass-count distinction (and life would be so much easier if it was!), but this hypothesis turns out to be too naive. Above all, the same object/stuff can be described by either types of nouns.

(7.12)

Mass	*Count*
We need **much more chocolate**.	We need **many more chocolates**.
We need **high-quality paper**.	We need **high-quality papers**.
We don't need **much rope**.	We don't need **many ropes**.
There isn't **much discussion**.	There aren't **many discussions**.
There is **reason** for this.	There is **a reason** for this.
There is **a lot of difference**.	There are **a lot of differences**.
There is **war** here.	There is **a war** here.

Other such nouns include: *detail*, *complexity*, *error*, *effort*, *shortage*, *exposure*, *change*, *variation*, etc. (see Gillon 1999 for some more).

The conclusion that emerges from this is that the mass-count distinction cannot be solely rooted in the properties of the objects/stuff described by the noun. Rather, it is a property of the expressions and language use. However, the mass-count distinction is not only a morphosyntactic issue, since there seems to be a semantic correlate of the morphosyntactic difference. Above all, it is undeniable that the intuition expressed by the Mapping Hypothesis captures the overall tendency. Furthermore, one can sense a semantic difference in the minimal pairs in (7.12).

An alternative way of thinking about the mass-count distinction is that the distinction reflects whether the speaker/conversational participants are interested in the **individuation** of the objects described, i.e. the way to determine what counts as one instance of that noun. If yes, use count nouns, if not, mass nouns. Let's state it (a bit loosely) as (7.13) (see Bunt 1985 for a similar idea):

(7.13) **Individuation Hypothesis**:
If a particular way of individuating the referents is salient/relevant, a count nouns is used. If not, a mass noun is used.

The idea is that count nouns somehow 'presuppose' a particular way of individuation. The underlying intuition is this:

- When I say *a dog*, I can safely presuppose shared knowledge about what constitutes one dog. For instance, when you count dogs, you don't count their heads, tails, paws, eyes, etc. separately. But in certain contexts, a salient way of counting becomes either irrelevant or unavailable. A classical example of such a situation is this: if you blow up a dog, there'll be dog all over the place!!
- On the other hand, in a typical situation, *juice* is used as a mass noun because there is clearly no salient way of individuating different portions of juice. However, if you are in a restaurant, for instance, there is a clear way, i.e. one juice corresponds to one serving. So you use it as a count noun.

This sounds like a plausible analysis, but there is a complication that comes from so-called **fake mass nouns** (also known as: 'count mass nouns', 'object-mass nouns', 'collective mass nouns', 'collectives', etc.; note that this is not a strictly defined term). That is, there are many pairs of predominantly mass and predominantly count nouns that are used to describe similar objects (McCawley 1975, Chierchia 1998a, Rothstein 2010, Pelletier 2012).

(7.14)

Mostly Mass	*Mostly Count*
mail	letter
change	coin
laughter	laugh
spaghetti	noodle
garlic	onion
toast	sandwich
luggage	suitcase
rice, corn	bean, pea, lentil
footware	shoe, sandle
carpeting	carpet
foliage	leaf
wildlife	animal
software	app(lication)
sushi	fishcake
baklava	brownie
fruit	vegetable
flu	cold
success	failure
advice	suggestion
knowledge	belief
information (English)	information (French)
hair (English)	cheveu (French), capello (Italian)
furniture (English), meubilair (Dutch)	meuble (French), mobile (Italian), meubel (Dutch)
parentela (Italian)	relative (English)
chalk (English)	giyr (Hebrew)

Many of these mass nouns seem to describe concrete objects, but morphosyntactically they are *bona fide* mass nouns. Take, for instance, *furniture*. It is pretty clear what counts as a piece of furniture, which in fact can be made explicit with the help of *a piece of*, and in this regard, it's not so different from *chair* or *desk*. Then, under the Individuation Hypothesis, *furniture* should be used as a count noun!

Barner & Snedeker (2005) conducted experiments whose results confirm these intuitions. In particular, the results of their Experiment 1 show that people behave differently for mass nouns with and without natural ways of individuation, e.g. *silverware* (a fake mass noun) vs. *toothpaste* (a pure mass noun). In their Experiment 1, participants were asked to answer questions by selecting one of two pictures. The questions looked like (7.15):

(7.15)a. Who has more silverware? (Fake/Object mass noun)

b. Who has more shoes? (Count noun)
c. Who has more toothpaste? (Mass noun)

One picture contained one huge object and the other picture a group of three small objects of the same type, as shown below (taken from Barner & Snedeker 2005):

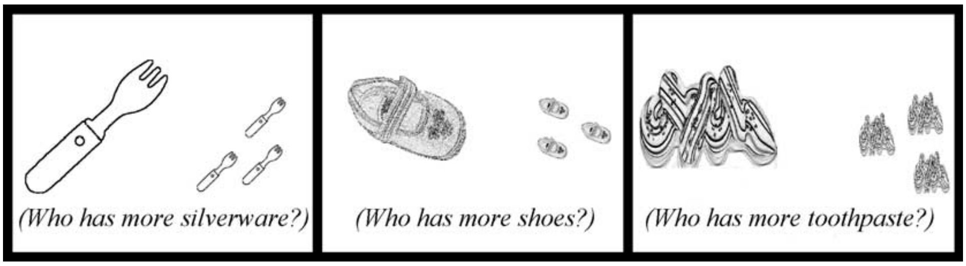


Fig. 1. Images of selected stimuli from Experiment 1 (object-mass: *silverware*; count: shoes; substance-mass: *toothpaste*).

The results indicate that people relied on numbers, rather than quantities, for fake mass nouns (*silverware*) and count nouns (*shoe*), but not for (pure) mass nouns (*toothpaste*):

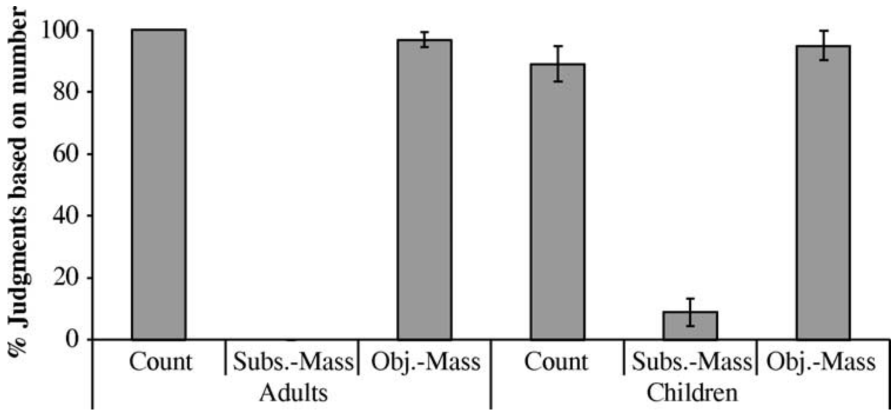


Fig. 2. Adults' and children's quantity judgments, as a percentage of judgments based on number of individuals.

All in all, it seems to me that there is a one-way generalisation (see Chierchia 2010 for a similar remark): A count noun always presupposes a salient way of individuating the objects, while a mass noun might or might not. In this sense, count nouns, or more precisely, count uses of nouns are more 'marked'.

Assuming that this markedness of count nouns is morphosyntactically encoded, fake mass nouns can be analysed as follows (see Borer 2005, Bale & Barner 2009, Rothstein 2010 for similar ideas).

- Every noun is a mass noun when it enters syntax.
- There is a functional head F^0 that turns a mass noun into a count noun, which semantically requires there to be a salient way of individuation (via presupposition). If a noun appears with F^0, it behaves as a count noun. If not, it behaves as a mass noun.
- Certain nouns, i.e. fake mass nouns, are morphosyntactically marked as incompatible with F^0.

- But whenever possible, the use of F^0 is preferred.

However, this is still problematic. For some count nouns, the way of individuation is not very clear:

(7.16) reason cloud puddle ripple mountain valley instant detail wave

Figure 7.1: How many clouds are there in this picture?

We leave this issue open.

7.1.1 The Issue of the Denotations of Mass Nouns

Now, let us turn to the denotations of mass nouns (or mass uses of nouns). As we have not identified the common denominators for mass and count nouns, the discussion here will be inevitably inconclusive. But we will review major views on the denotations of mass nouns.

First, to remind you of our theory of count nouns (or count uses of nouns) from last week, their denotations characterise sets of individuals.

(7.17) For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{book}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\ x$ is a book in $\mathcal{M}$

b. $[\![\textbf{books}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part y of x is a book in $\mathcal{M}$

If there are three books in the model $\mathcal{M}_1$, b_1, b_2 and b_3, we have:

(7.18) a. $[\![\textbf{book}]\!]^{a,\mathcal{M}_1}$ characterises $\{ b_1, b_2, b_3 \}$

b. $[\![\textbf{books}]\!]^{a,\mathcal{M}_1}$ characterises $\left\{ \begin{array}{ccc} b_1, & b_2, & b_3, \\ b_1 \oplus b_2, & b_1 \oplus b_3, & b_2 \oplus b_3, \\ & b_1 \oplus b_2 \oplus b_3 & \end{array} \right\}$

Importantly, these sets have 'atoms' = those individuals that do not have proper parts, i.e. singular individuals. Numerals and 'counting expressions' like *a lot*, *several*, *most* etc. count the number of these atoms.

(7.19) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{two}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has two singular parts

(7.20) $[\![\textbf{two books}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3 \,\}$

What does mass nouns like *water* and *furniture* denote? The first clue comes from the property called **cumulative reference** (Quine 1960).

(7.21) A predicate P refers cumulatively iff for any x, y, whenever $P(x) = 1$ and $P(y) = 1$, $P(x \oplus y) = 1$.

As we can see, plural count nouns and mass nouns refer cumulatively, while singular count nouns do not.

- Suppose that we have a pile of books, x, and another pile of books, y. Both x and y are 'books' ($[\![\textbf{books}]\!]^{a,\mathcal{M}}(x) = [\![\textbf{books}]\!]^{a,\mathcal{M}}(y) = 1$). If we put them together and create another pile of books, $x \oplus y$, it also counts as books ($[\![\textbf{books}]\!]^{a,\mathcal{M}}(x \oplus y) = 1$).
- Suppose we have a portion of milk x in a glass and another portion of milk y in another glass. x is milk and y is milk ($[\![\textbf{milk}]\!]^{a,\mathcal{M}}(x) = [\![\textbf{milk}]\!]^{a,\mathcal{M}}(y) = 1$). If we put them together in one jug, we have $x \oplus y$, which is also milk ($[\![\textbf{milk}]\!]^{a,\mathcal{M}}(x \oplus y) = 1$).
- Singular nouns do not refer cumulatively. Suppose we have a book x and another book y ($[\![\textbf{book}]\!]^{a,\mathcal{M}}(x) = [\![\textbf{book}]\!]^{a,\mathcal{M}}(y) = 1$). If we put them together, we have books, not a book! ($[\![\textbf{book}]\!]^{a,\mathcal{M}}(x \oplus y) = 0$).
- Note that being able to sometimes describe $x \oplus y$ by the noun is not enough. It must be always possible. Take two lines, x and y ($[\![\textbf{line}]\!]^{a,\mathcal{M}}(x) = [\![\textbf{line}]\!]^{a,\mathcal{M}}(y) = 1$). You can certainly combine them to produce another line $x \oplus z$. But this requires a particular geometrical arrangment.

So cumulative reference distinguishes mass nouns and plural count nouns from singular count nouns. Notice that our analysis does assign a cumulative denotation to plural count nouns, and a non-cumulative denotation to singular count nouns. Since mass nouns refer cumulatively, their denotations should be closed by $\oplus$, just like the denotations of plural count nouns are.

This view is further supported by the observation that mass nouns, just like plural count nouns and unlike singular count nouns, give rise to a distributive-collective ambiguity (Gillon 1992, 1999, Nicolas 2005, Schwarzschild 2009, Lasersohn 2011):

(7.22) These boxes are expensive.
  a. Distributive: Each of these boxes is expensive.
  b. Collective: The group consisting of these boxes as a whole is expensive.

(7.23) [pointing at 6 bottles of wine]
This wine is expensive.
  a. Distributive: Each bottle of wine is expensive.
  b. Collective: The group consisting of the 6 bottles of wine as a whole is expensive.

Then, what distinguishes mass nouns from plural count nouns? Cheng (1973) proposes that it is another property called the **divisive reference** (see also Bunt 1985):

(7.24) A predicate P refers divisively iff for any x, if $P(x) = 1$, then for any part y of x, $P(y) = 1$.

This says: If P is true of something, it is true of any of its parts. The thought is, if there's, say, something that can be described by the noun *time*, any part of it also counts as *time*. Although this might work for nouns like *time*, this is clearly problematic as a characterisation of mass nouns in general.

- Firstly, for fake mass nouns, (7.24) simply does not hold. Consider *furture*. It's clearly not the case that every part of a piece of furniture is also furniture. E.g., a leg of a desk is not furniture.
- Secondly, even for a more canonical mass noun like *coffee*, it does not hold: if you keep breaking a portion of coffee into its components, at some point, you will have pure water, which is not coffee!
- And even for more pure mass nouns like *water*, if one keeps dividing water, at some point, there will be hydrogen atoms and oxygen atoms (and you can go on to break them too). These things themselves are not water![2]

This is called the **minimal parts problem** (Quine 1960)[3] Therefore, divisive reference is not a general property of mass nouns.[4]

In the literature, there are two major views on the denotations of mass nouns, but they have pros and cons.

- Link (1983) assumes that there is a domain of individuals D_e, and a separate domain of substance D_s, where divisive reference may hold. Count nouns denote subsets of D_e, while mass nouns denote subsets of D_s. So they are inherently about two different types of things.

  However, Link (1983) completely ignores fake mass nouns. It is highly counterintuitive to say, as Chierchia (1998a,b, 2010) argues, that *furniture* is about different types of things from what *chairs* and *desks* are about.

  Also, D_s is meant to allow divisive reference, but it is not a crucial property, as we have just discussed.
- Chierchia (1998a,b, 2010) claims that mass nouns have the same kind of denotations as plural count nouns (see also Gillon 1992). According to Chierchia, the

[2] For this, one might say that the ontology of entities in semantic models does not have to reflect the reality. Rather, it is a function of how we conceive of the world (in the spirit of 'natural language metaphysics' in the sense of Bach 1986). And maybe we regard water that way. However, as Pelletier (2012) points out, this would entail that we use language in a way that does not reflect our beliefs, for almost all of us in fact believe that water consists of hydrogen and oxygen atoms.

[3] In the words of (Quine 1960:99): "there are parts of water, sugar, and furniture too small to count as water, sugar, furniture. Moreover, what is too small to count as furniture is not too small to cunt as water or sugar; so the limitation needed cannot be worked into any general adaptation of 'is' or 'is a part of' but must be left rather as the separate reference-dividing business of the several mass terms".

[4] There are also count nouns that refer divisively to some extent, e.g. *sequence*, *twig*, *fence*, etc. See Zucchi & White (2001) and Rothstein (2010) for discussion on these nouns.

main difference between mass and plural count nouns is that for mass nouns, the criterion for individuation is 'vague'.

However, I think it's counterintuitive to assume that the criteria for individuation for fake mass nouns like *furniture* are vague, while those for nouns like *cloud* are not.

As remarked above, we need to leave this issue unsolved for the moment.

7.1.2 'Plural Mass Nouns'

To add another layer of complexity, there is one class of nouns in English that have not been paid enough attention in the theoretical literature (Ojeda 2005 contains interesting discussion on them; see also McCawley 1975, Gillon 1992, Schwarzschild 2009). They are sometimes called **plural mass nouns**. In short, they are plural nouns that behave like mass nouns, and often lack singular forms. Here are some examples:

(7.25) clothes dregs guts bowels brains dues annals
earnings goods spirits shavings belongings valuables

These nouns show certain features of plural count nouns, e.g.

(7.26) My **clothes are/*is** in this locker. (McCawley 1975:320)

(7.27) a. The club requires **these/*this dues** to be paid immediately.
b. **Dues are/*is** to be paid upon joining.
c. The person who collects dues knows how much **they are/*it is**.
(Gillon 1992:612)

However, in other respects, they behave like mass nouns, e.g.

(7.28) a. *I've just bought **several/five clothes**.
b. ***Many clothes** are too expensive for me to buy. (McCawley 1975:320)

(7.29) a. How **much/*many brains** does Bill have?
b. How **little/*few brains** does Bill have? (Gillon 1992:613)

7.2 'Group Nouns'

There are nouns that denote groups of entities, sometimes called **groups nouns**:

(7.30) team family committee faculty staff class

Unsurprisingly, they can behave like normal nouns in the sense that they characterise sets of groups as entities, e.g. *commettee* characterises a set of committees. Our semantics works for examples like (7.31) straightforwardly.

(7.31)a. One committee was founded two years ago.
b. The team has a lot of supporters.

What is interesting about these nouns is that they sometimes behave like plural individuals consisting of the describe group, even when they are singular.

(7.32)a. The committee is smiling.
≈ The members of the committee are smiling.
b. The Dutch team is very tall.
≈ The members of the Dutch team are very tall.

These nouns also license collective predication.

(7.33)a. The family gathered in the living room.
b. The team can't stand each other.

As you might know, in British and Canadian English, plural agreement is possible (Barker 1992, Pearson 2011, De Vries 2012)

(7.34)a. The committee **hope** that you will accept the job.
b. The basketball team have surpassed **themselves** with their recent performance. (Pearson 2011:161)

Pearson (2011) also points out that these nouns can appear in cardinal partitive quantifiers:

(7.35)a. Three of the committee came to the meeting.
b. Several of the family objected to Bill marrying Mary.
c. Many of the present cabinet will have to resign. (Pearson 2011:162)

Thus the puzzle is that group nouns sometimes behave like regular nouns but sometimes like plural nouns denoting the members of the described groups.

It is also interesting that in British and Canadian English, the plural agreement is optional. but there is an interpretive difference, as observed by Barker (1992). In order to see this, consider:

(7.36)a. The committee is old.
b. The committee are old.

The singular agreement (7.36a) is ambiguous here. Either the committee is an old committee, or the members of the committee are old people. On the other hand, (7.36b) is unambiguous and only has the latter reading. In other words, with plural agreement, only the distributive reading is available.

To reinforce this generalisation, consider (7.37) with a collective predicate. Here, only singular agreement is possible, even in British and Canadian English.

(7.37)a. The team was formed in 1991.
b. *The team were formed in 1991.

7.2.1 'Collection Nouns'

Interestingly, Pearson (2011) points out that nouns like (7.38), which she calls 'collection nouns', behave differently from group nouns like *committee* (see also Barker 1992).

(7.38) bunch pile group list heap pile

Firstly, these nouns are incompatible with collective predicates:

(7.39) a. *The bunch of flowers looks nice together.
b. *The heap of papers is equally interesting. (Pearson 2011:163)

Secondly, even in British and Canadian English, plural agreement is impossible.

(7.40) a. *The bunch of flowers are tall.
b. *The pile of dishes are touching each other.
c. *The group of statues resemble themselves. (*Ibid.*)

Thirdly, these nouns cannot appear in cadinal partitives (unless plural):

(7.41) a. *Three of the bunch of flowers had died.
b. *Several of the deck of cards had gone missing.
c. *Many of the pile of dishes needed to be washed. (*Ibid.*)

So simply put, collection nouns do not have the special properties that group nouns have, i.e. they lack the plural behaviour. In other words, collection nouns are well-behaved normal count nouns.

Generally, nouns that describe groups of inanimate individuals (e.g. *bunch of flowers*) do not give rise to the 'member reading', while those that describe groups of animate individuals (e.g. *committee*) do.

7.2.2 Towards the Semantics of Group Nouns

What are the meanings of group nouns? What cannot be maintained is that *the committee* has the same semantics as *the members of the committee*. As Schwarzschild (1996) and Pearson (2011) point out, this analysis fails to account for the following contrast.

(7.42) a. The committee was formed.
b. *The members of the committee were formed.

In other words, as we already saw above, *the committee* can describe a group as a single entity, which *the members of the committee* cannot.

There are three possible analyses:

- Group nouns and collection nouns are like normal nouns and characterise sets of groups as entities.

(7.43) For any model $\mathcal{M}$, and for any assignment a,
$$[\![\textbf{committee}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x \text{ is a committee in } \mathcal{M}$$

But there is a special process that terns animate groups into the plural individual consisting of the members.

(7.44) For any model $\mathcal{M}$, and for any assignment a,
$$[\![\textbf{committee}]\!]^{a,\mathcal{M}} \rightsquigarrow \left[\lambda x \in D_e \begin{array}{l}\text{for some committee } c \text{ in } \mathcal{M},\\ \text{each singular part of } x \text{ is a member of } c\end{array}\right]$$

We can assume that this rule has animacy restrictions, so does not apply to collection nouns.

- Another possibility is the converse of the above analysis. *Committee* characterises a set of committee members, as in (7.44), and get turned into their groupings, based on who belongs to which committee, i.e. (7.43). The animacy restriction is stated as follows: for group nouns (with animate members), the process is optional, but for collection nouns (with inanimate members), the process is obligatory.

This is a topic that is still actively investigated and we do not have a definitive conclusion yet.

7.3 Further Readings

There is tons of papers and books on mass nouns and how they are different from count nouns, both in linguistics and philosophy. Lasersohn (2011) is a nice overview article. Some of the major works on the semantics of mass nouns are cited in the main text (Link 1983, Bunt 1985, Gillon 1992, Chierchia 1998a,b, Nicolas 2005, Bale & Barner 2009, Chierchia 2010, Rothstein 2010), but this list is by no means exhaustive.

Roger Schwarzschild has a very interesting take on this issue. Based on the behaviour of predicates like *large*, *small*, *round* and *long*—which he calls *stubbornly distributive predicates*—Schwarzschild (2009) claims that there are two types of mass nouns, which he calls *multi-participant* and *mixed-participant nouns*. He further claims that count nouns are *single-participant nouns*. He cashes out this idea in a theory where nouns are predicates of 'events', rather than predicates of individuals. He develops this view further in Schwarzschild (2014) (which will be available as a paper soon).

The discussion on mass vs. count nouns has spawned a lot of work on cross-linguistic variation, especially since the seminal works by Krifka (1989) and Chierchia (1998a,b) where they discuss differences between European languages and so-called 'classifier languages' like Chinese and Japanese. Doetjes (2012) is a very useful overview of cross-linguistic issues surrounding mass-count distinction.

Lisa Cheng and Rint Sybesma have a series of papers on this issue on languages of China (Cheng & Sybesma 1999, 2005, 2012, Cheng, Doetjes & Sybesma 2008, Cheng, Doetjes, Sybesma & Zamparelli 2012). They are particularly interested in the function of so-called ‘classifiers’ in the number system in Chinese languages. Chierchia (1998a,b) and Borer (2005) claim that in classifier languages like Mandarin and Cantonese, all nouns are in some sense mass, they are all ‘pluralised’, unlike in English, and numerals require classifiers to make the noun ‘count’ (see also Krifka 1989, Lucy 1992, Cheng & Sybesma 1999). However, this has been questioned by several authors, based on experimental evidence that the noun meanings do not differ cross-linguistically in any essential ways (Barner, Inagaki & Li 2009, Cheng et al. 2008, Cheung, Li & Barner 2012, Li, Dunham & Carey 2009). In particular, Cheung et al. (2012) reports that Mandarin speakers behave the same as English speakers in counting tasks similar to Barner & Snedeker’s (2005) (see Barner et al. 2009, Inagaki & Barner 2009 for similar observations on Japanese).

The cross-linguistic work on the semantics of nouns has gone beyond classifier languages. Matthieu (2012) discusses gender shift and its effects on the number and mass-count in Ojibwe, Breton, etc. (e.g. in Breton, the masculine noun *geot* is a mass noun meaning ‘grass’, but its feminine form *geot-enn* is a count noun meaning ‘a blade of grass’). Lima (2014) investigate a native language of Brazil that has both singular-plural distinction and classifiers. Khanjian (2008) looks at (Western) Armenian, and Grimm (2012) at Welsh. See also papers in Massam (2012).

Nouns like *group* and *committee* are extensively discussed by Landman (1989a,b), Barker (1992), Schwarzschild (1996), Pearson (2011) and De Vries (2012). While Barker (1992) and Schwarzschild (1996) analyse them as denoting singular entities, Landman (1989a,b) assigns a special ontological status to it that is different from both singular and plural individuals of a normal kind (Link 1983 seems to countenance this view). Pearson (2011) explores a possibility that these nouns have singular intensions and plural extensions. In De Vries (2013) (as well as in her dissertation to be defended on 13 March 2015), Hanna de Vries argues that group nouns denote plural individuals (which are sets for her), rather than singular individuals.

7.4 Exercises

In this week’s exercises, you will discuss problems of two analyses of mass nouns. The discussion is essentially open-ended.

1. The first analysis is a ‘straw man’. Suppose that the denotations of mass nouns characterise set of singular individuals, as in (7.45).

   (7.45)For any model $\mathcal{M}$ and for any assignment a,

   a. $[\![\textbf{water}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\ x$ is water in $\mathcal{M}$

   b. $[\![\textbf{water}]\!]^{a,\mathcal{M}}$ characterises $\{ w_1, w_2, w_3, \ldots \}$ where each w_1 is water.

   Your task in this exercise to construct arguments that would convince somebody that this analysis is wrong. For example, try to show that this analysis does not account for some of the observations we made in the main text. You can also make use of the material from the previous two weeks to construct counter-arguments. It’s all up to you.

2. Similarly, what are potential problems of analysing mass nouns as denoting sets of singular and plural individuals, pretty much in the same manner that we analyse plural nouns.

   (7.46) For any model $\mathcal{M}$ and for any assignment a,

   a. $[\![\textbf{water}]\!]^{a,\mathcal{M}} = \lambda x \in D_e$. each singular part of x is water in $\mathcal{M}$

   b. $[\![\textbf{water}]\!]^{a,\mathcal{M}}$ characterises $\{w_1, w_2, w_3, \ldots, w_1 \oplus w_2, w_1 \oplus w_3, \ldots\}$ where each w_1 is water.

   Does this account for some of the problems you raised above? What kind of problems still remain? Again, try to construct cogent arguments by referring to data mentioned in the main text or your own observations.

Chapter 8

Distributivity and Cumulativity

8.1 Review

- Plural individuals and singular individuals (Lecture 6):
  - The set of singular individuals SG.
  - The set of plural individuals PL.
  - The domain D_e of individuals is $SG \cup PL$.

(8.1) a. $SG = \{ a, b, c, d \}$

b. $$PL = \left\{ \begin{array}{c} (a \oplus b), \quad (a \oplus c), \quad (a \oplus d), \quad (b \oplus c), \quad (b \oplus d), \quad (c \oplus d), \\ (a \oplus b \oplus c), \quad (a \oplus b \oplus d), \quad (a \oplus c \oplus d), \quad (b \oplus c \oplus d), \\ (a \oplus b \oplus c \oplus d) \end{array} \right\}$$

c. $$D_e = \left\{ \begin{array}{c} a, \quad b, \quad c, \quad d, \\ (a \oplus b), \quad (a \oplus c), \quad (a \oplus d), \quad (b \oplus c), \quad (b \oplus d), \quad (c \oplus d), \\ (a \oplus b \oplus c), \quad (a \oplus b \oplus d), \quad (a \oplus c \oplus d), \quad (b \oplus c \oplus d), \\ (a \oplus b \oplus c \oplus d) \end{array} \right\}$$

- There are three types of predicates (Lecture 6).
  - **Distributive predicates** are those that describe properties of singular individuals, e.g. *fall asleep*, *is blond*.
  - **Collective predicates** are those that describe properties of plural individuals, e.g. *gather*, *form a circle*, *are a couple*.
  - **Mixed predicates** are those that describe properties of either kind of individuals, e.g. *carry the piano*, *write poems*.
- There are several types of DPs that denote plural individuals, e.g. conjoined proper names (Lecture 6):

(8.2) a. $[\![\textbf{John}]\!]^{a,\mathcal{M}} = j \qquad [\![\textbf{Mary}]\!]^{a,\mathcal{M}} = m$

b. $[\![\textbf{John and}_{\langle e,ee\rangle}\ \textbf{Mary}]\!]^{a,\mathcal{M}} = j \oplus m$

The connective simply denotes $\oplus$ ('non-Boolean conjunction'):

(8.3) $[\![\textbf{and}_{\langle e,ee\rangle}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\lambda y \in D_e.\ x \oplus y$

Plural NPs are true of plural individuals, unlike singular NPs (Lecture 7):

(8.4) Suppose that there are two students (s_1 and s_2) and three books (b_1, b_2 and b_3) in $\mathcal{M}_1$.

a. $[\![\textbf{student}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, s_1, \quad s_2 \,\}$

b. $[\![\textbf{students}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, s_1, \quad s_2, \quad s_1 \oplus s_2 \,\}$

c. $[\![\textbf{book}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, b_1, \quad b_2, \quad b_3 \,\}$

d. $[\![\textbf{books}]\!]^{a,\mathcal{M}_1}$ characterises $\left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \\ b_1 \oplus b_2, \quad b_2 \oplus b_3, \quad b_1 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\}$

Given this, we can assign the following meaning to the plural morpheme:

(8.5) For any model $\mathcal{M}$, and for any assignment a,
$[\![\textbf{-s}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ \lambda x \in D_e.$ for each singular part y of x, $P(y) = 1$

Probably mass nouns are also true of 'plural individuals', i.e. individuals that have non-trivial parts with respect to $\oplus$, but it's not clear what exactly their denotations are (Lecture 8).

Numerals denote functions of type $\langle e,t \rangle$ and modify a noun via Predicate Modification (Lecture 7):

(8.6) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{two}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has two distinct singular parts

(8.7) a. $[\![\textbf{two students}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, s_1 \oplus s_2 \,\}$

b. $[\![\textbf{two books}]\!]^{a,\mathcal{M}_1}$ characterises $\{\, b_1 \oplus b_2, \quad b_1 \oplus b_3, \quad b_2 \oplus b_3 \,\}$

(But recall Van Bentham's Trap for certain modified numerals)

The definite article *the* denotes the supremum-operator (denoted by σ), which picks out the unique maximal element of a given set, if it exists (Lecture 7).

(8.8) a. $[\![\textbf{the students}]\!]^{a,\mathcal{M}_1} = \sigma(\{\, s_1, \quad s_2, \quad s_1 \oplus s_2 \,\}) = s_1 \oplus_2$

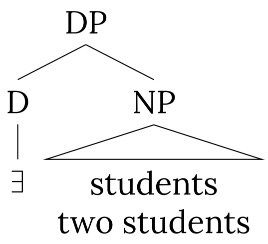
b. $[\![\textbf{the two students}]\!]^{a,\mathcal{M}_1} = \sigma(\{\, s_1 \oplus s_2 \,\}) = s_1 \oplus_2$

c. $[\![\textbf{the books}]\!]^{a,\mathcal{M}_1} = \sigma \left(\left\{ \begin{array}{c} b_1, \quad b_2, \quad b_3, \\ b_1 \oplus b_2, \quad b_2 \oplus b_3, \quad b_1 \oplus b_3, \\ b_1 \oplus b_2 \oplus b_3 \end{array} \right\} \right) = b_1 \oplus b_2 \oplus b_3$

d. σ is not defined for $[\![\textbf{two books}]\!]^{a,\mathcal{M}_1}$, because there is no unique maximal element (similarly for $[\![\textbf{book}]\!]^{a,\mathcal{M}_1}$ and $[\![\textbf{student}]\!]^{a,\mathcal{M}_1}$).

We also postulated an invisible existential determiner $\exists$ for determiner-less plural DPs (Lecture 7):

(8.9)



∃ means the same thing as *some* (although there are several differences, see Lecture 7).

(8.10) For any model $\mathcal{M}$ and for any assignment a,
$[\![\exists]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.$ there is $x \in D_e$ such that $P(x) = Q(x) = 1$
(In set talk, $[\![\exists]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\ |P \cap Q| \neq \varnothing$)

- Collective predicates select for plural arguments.[1]

(8.11) a. *John is a couple.
b. John and Mary are a couple.
c. The two students are a couple.

Distributive predicates semantically require singular individuals as their arguments, but on the surface, they are compatible with plural DPs.

(8.12) a. John sneezed.
b. John and Mary sneezed.
c. The two students sneezed.

In order to make sense of this , we postulated the phonologically null distributivity operator Δ.

(8.13) For any model $\mathcal{M}$ and for any assignment a,
$[\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda x \in D_e.$ for each singular part y of x, $P(y) = 1$

(8.14) For any model $\mathcal{M}$ and for any assignment a,
$[\![\Delta\ \textbf{sneezed}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part of x sneezed in $\mathcal{M}$

We account for the two readings of mixed predicates with plural subjects with the help of Δ.

(8.15) a. John carried the box.
b. John and Mary carried the box. (Collective reading)

[1] As noted last time and in the updated version of the lecture notes for Week 7, there are two types of collective predicates distinguished by the compatibility with quantified DPs.

(i) a. The UCL students are politically homogeneous.
b. *All the UCL students are politically homogeneous.

We don't account for this difference here (see Dowty 1987, Winter 2001a, 2002), but notice that it suggests that *the* NPs and *all the* NPs cannot be synonymous).

c. John and Mary Δ carried the box. (Distributive reading)

Today, we talk more about distributive readings and how they come about. In the second half, we will look at another type of reading that plural nouns phrase give rise to, **cumulative readings**.

8.2 More on Distributive Readings

8.2.1 Distributivity = Plurality

Notice that the plural morpheme -s and Δ have the same meaning!

(8.16)For any model $\mathcal{M}$, and for any assignment a,
$[\![\textbf{-s}]\!]^{a,\mathcal{M}} = [\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\ \lambda x \in D_e.$ for each singular part y of x, $P(y) = 1$

This leads to the following hypothesis (cf. Link 1983, Landman 1989a,b, 2000): **Distributivity = plurality**. The idea is that just like you can pluralise nouns, you can pluralise verb and other non-nominal predicates.

In order to simplify the discussion, let us concentrate on verbs and adjectives that denote one-place predicates (i.e. functions of type $\langle e,t\rangle$) for the moment. We'll come back to transitive predicates (functions of type $\langle e,et\rangle$) later.

- Consider *is blond*. This predicate is only true of singular individuals.

  (8.17)For any model $\mathcal{M}$, and for any assignment a,
  $[\![\textbf{is blond}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\ x$ is blond in $\mathcal{M}$

- If you pluralise this predicate:

  (8.18)For any model $\mathcal{M}$, and for any assignment a,
  $[\![\textbf{are blond}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part of x is blond in $\mathcal{M}$

  This pluralised predicate can take a plural individual as its argument.

From this example, one might be tempted to think that the plural agreement indicates pluralisation of the predicate, just like -s does for a noun. However, the correspondence between plural agreement in morphosemantics, and semantic plurality (=distributivity) does not obtain in the general case:

- Group nouns allow singular agreement with distributive readings (in all dialects of English).

  (8.19)The Dutch team is tall.

- Mass nouns also give rise to distributive reading without triggering plural agreement.

  (8.20)All my furniture is wooden.

- A mixed predicate with plural agreement can receive a non-distributive reading. E.g., (8.21) is still ambiguous between a collective and distributive reading, although the agreement is plural.

(8.21) John and Mary have carried the piano.

So number agreement is a matter of morphosyntax, and does not necessarily correlate with semantic plurality (unlike the plural morpheme -*s* on nominals).

8.2.2 Lexical vs. Phrasal Distributivity

Recall that according to our analysis, 'plural' does not mean plural (quite confusingly), but is number neutral. So plural predicates may be true of singular individuals as well.

Keeping this in mind, let us consider the following hypothesis (cf. Scha 1981, Krifka 1992, Landman 1996, 2000, Kratzer 2013).

(8.22) **Lexical Plurality Hypothesis** (ver. 1): Non-nominal predicates (verbs, adjectives, prepositions) are always 'plural' (and are exempt from the Unmarked Plural Rule).

According to this hypothesis, there's no need to say that *is blond* has separate singular and plural versions, because it is plural to begin with.

(8.23) For any model $\mathcal{M}$, and for any assignment a,
$[\![\textbf{is/are blond}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part of x is blond in $\mathcal{M}$

So to analyse the following examples, we don't need to posit Δ!

(8.24) a. John is blond.
b. The boys are blond.

(8.25) a. John sneezed.
b. The boys sneezed.

This is quite nice, as you don't need to postulate a covert operator like Δ to account for these sentences.

The next question, then, is, can we do away with the distributivity operator Δ altogether? Unfortunately, the answer is no. The reason is because we need to be able to pluralise a non-lexical, phrasal predicate, and in order to do so, you need an operator. To see this argument, consider (8.26).

(8.26) a. John sneezed or yawned.
b. The students sneezed or yawned.

Unlike (8.26a), (8.26b) is ambiguous between two readings.

- Each student sneezed or each student student yawned (they all did the same thing).

- Each student did one of two things: sneezing or yawning (different students did different things).

If we do not have Δ and only have lexical pluralisation, we can only account for the first reading. Here are the details.

- Firstly, we assume that there is a version of disjunction *or* that combines two functions of type $\langle e,t\rangle$ and produces another function of the same type.

  (8.27) $[\![\mathbf{or}_{\langle et,\langle et,et\rangle\rangle}]\!]^{a,\mathcal{M}} = \lambda Q \in D_{\langle e,t\rangle}.\lambda P \in D_{\langle e,t\rangle}.\lambda x \in D_e.\ Q(x) = 1 \text{ or } P(x) = 1$

  NB: this is a particular version of 'generalised disjunction' (a.k.a. Boolean disjunction) we considered in Term 1, where $\sigma = \langle e,t\rangle$.

  (8.28) *Generalised disjunction*
  For any model $\mathcal{M}$, for any assignment a, and for any type σ that ends in t,

  $$[\![\mathbf{or}]\!]^{a,\mathcal{M}} = \lambda x \in D_\sigma.\lambda y \in D_\sigma.\begin{cases} x = 1 \text{ or } y = 1 & \text{if } \sigma = t \\ \lambda z \in D_{\tau_1}.\ [\![\mathbf{or}]\!]^{a,\mathcal{M}}(x(z))(y(z)) & \text{if } \sigma = \langle \tau_1, \tau_2\rangle \end{cases}$$

- By assumption, the verbs are pluralised (Lexical Plurality Hypothesis):

  (8.29) For any model $\mathcal{M}$, and for any assignment a,
  a. $[\![\mathbf{sneezed}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part of x sneezed in $\mathcal{M}$
  b. $[\![\mathbf{yawned}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.$ each singular part of x yawned in $\mathcal{M}$

- If these are combined via *or*, we obtain:

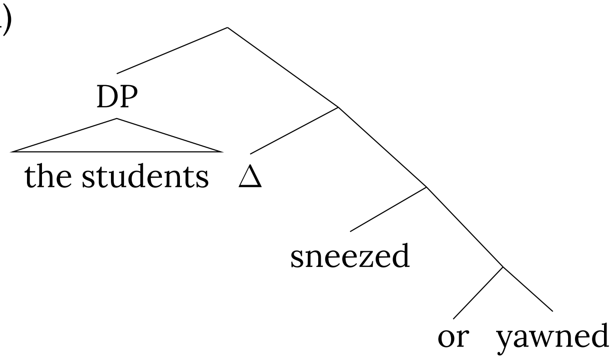
  (8.30) For any model $\mathcal{M}$, and for any assignment a,

  $$\begin{aligned} & [\![\mathbf{sneezed\ or\ yawned}]\!]^{a,\mathcal{M}} \\ & = [\![\mathbf{or}_{\langle et,\langle et,et\rangle\rangle}]\!]^{a,\mathcal{M}}([\![\mathbf{yawned}]\!]^{a,\mathcal{M}})([\![\mathbf{sneezed}]\!]^{a,\mathcal{M}}) \\ & = \lambda x \in D_e.\ [\![\mathbf{yawned}]\!]^{a,\mathcal{M}}(x) = 1 \text{ or } [\![\mathbf{sneezed}]\!]^{a,\mathcal{M}}(x) = 1 \\ & = \lambda x \in D_e.\ \left[\begin{array}{l} \text{each singular part of x yawned in } \mathcal{M} \text{ or} \\ \text{each singular part of x sneezed in } \mathcal{M} \end{array}\right] \end{aligned}$$

  Notice that the universal quantifier contributed by the pluralisation ('for each singular part ...') is taking scope below the disjunction. If (8.30) is combined with a plural subject, the resulting truth-conditions are: each singular part of the plural subject yawned or each singular part of the plural subject sneezed. This is the first reading.

In order to derive the second reading, we need to be able to pluralise the disjoined predicate, as depicted in the following diagram in terms of Δ:

(8.31)



If Δ is applied to (8.30), we get:

(8.32) $[\![\Delta\ \textbf{sneezed or yawned}]\!]^{a,\mathcal{M}}$

$$= [\![\Delta]\!]^{a,\mathcal{M}}([\![\textbf{sneezed or yawned}]\!]^{a,\mathcal{M}})$$
$$= \lambda z \in D_e.\ \text{for each singular part } y \text{ of } z,\ [\![\textbf{sneezed or yawned}]\!]^{a,\mathcal{M}}(y) = 1$$
$$= \lambda z \in D_e. \begin{bmatrix} \text{for each singular part } y \text{ of } z, \\ \quad \text{each singular part of } y \text{ yawned in } \mathcal{M} \text{ or} \\ \quad \text{each singular part of } y \text{ sneezed in } \mathcal{M} \end{bmatrix}$$
$$= \lambda z \in D_e.\ \text{each singular part } y \text{ of } z \text{ yawned in } \mathcal{M} \text{ or sneezed in } \mathcal{M}$$

(NB: each singular individual x has a trivial singular part, namely x itself)

To conclude, distributivity can be thought of as the same thing as plurality. But **lexical distributivity** is not enough and we need a way to derive distributive readings above a lexical level, or **phrasal distributivity**. The distributivity operator Δ that 'pluralises' its argument achieves exactly this.

The Lexical Plurality Hypothesis is in principle compatible with the distributivity operator Δ, but assuming both creates a certain degree of redundancy. Since we cannot dispense with Δ anyway, let us drop the Lexical Plurality Hypothesis. This, however, does not mean that the hypothesis is wrong, and you might be able to construct arguments for a position with both the Lexical Plurality Hypothesis and Δ.

8.2.3 Non-Atomic Distributivity

Our assumption so far is that pluralisation amounts to universal quantification over singular parts.

(8.33) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda x \in D_e.\ \text{for each singular part } y \text{ of } x,\ P(y) = 1$$

There is reason to regard this as a special case of more general notion of plurality/distributivity.

Consider the following example due to Gillon (1987:212). Suppose there are three men in the model $\mathcal{M}_2$, (Richard) Rodgers, (Oscar) Hammerstein (II), and (Lorenz) Hart. Rodgers and Hammerstein wrote musicals together. Similarly, Rodgers and

Hart wrote musicals together. And these collaborative works are the only musicals we have in the model. Then, we have (forget about the plural object for the moment):

(8.34)For any assignment a,

a. $[\![\textbf{Rodgers, Hammerstein, and Hart}]\!]^{a,\mathcal{M}_2}$
$= \text{Rodgers} \oplus \text{Hammerstein} \oplus \text{Hart}$

b. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Rodgers} \oplus \text{Hammerstein}) = 1$

c. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Rodgers} \oplus \text{Hart}) = 1$

d. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Rodgers}) = 0$

e. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Hammerstein}) = 0$

f. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Hart}) = 0$

g. $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}(\text{Rodgers} \oplus \text{Hammerstein} \oplus \text{Hart}) = 0$

Now consider the sentence in (8.35).

(8.35)Rodgers, Hammerstein, and Hart wrote musicals.

The sentence is intuitively judged true in $\mathcal{M}_2$, but this is not accounted for by our distributive operator, which universally quantifies over the singular parts of the subject.

(8.36)$[\![\textbf{Rodgers, Hammerstein, and Hart} \, \Delta \, \textbf{wrote musicals}]\!]^{a,\mathcal{M}_2} = 1$ iff Rodgers wrote musicals (alone), Hammerstein wrote musicals (alone), and Hart wrote musicals (alone) in $\mathcal{M}_2$.

Thus, the sentence is wrongly predicted to be false! In order to account for the relevant reading of the sentence, we need a way to distribute over non-singular parts.

The same thing can be illustrated with collective predicates.

(8.37)The semanticists and the syntacticians outnumber the phonologists.

There is a reading of (8.37), which is equivalent to: the semanticists outnumber the phonologists and the syntacticians outnumber the phonologists (There is also a completely collective reading; the semanticists and syntacticians together outnumber the phonologists).

Again, Δ does not account for this, and in fact it would make the sentence unacceptable, because *outnumber the phonologists*, being a collective predicate, cannot take a singular subject.

In order to account for such 'non-atomic distributive readings', we will generalise the notion of distributivity as follows.

- The distribution over singular parts is a special case, where the plural subject is decomposed into its singular parts.
- A more general notion of distributivity is that the plural subject is decomposed

into any kind of part.

We define an operator $\star$ (in our metalanguage) that denotes this. Here's one way to formalise it:

(8.38)For any predicate P of type $\langle e,t\rangle$,

$$\star(P) = \lambda x \in D_e . \left[\begin{array}{l} P(x) = 1 \text{ or} \\ x = y \oplus z \text{ and } \star(P)(y) = \star(P)(z) = 1 \end{array} \right]$$

By way of illustration, consider the musical example above (8.35). Here, the relevant predicate P of type $\langle e,t\rangle$ is $[\![\textbf{wrote musicals}]\!]^{a,\mathcal{M}_2}$, which characterises

$$\{\,\text{Rodgers} \oplus \text{Hammerstein}, \quad \text{Rodgers} \oplus \text{Hart}\,\}$$

$\star(P)$ is true of any individual x that is either (i) in this set, or (ii) can be decomposed into two parts y and z that $\star(P)$ is true of. In our example, the subject $x =$ Rodgers$\oplus$ Hammerstein $\oplus$ Hart, so (i) is clearly not the case. But it can be decomposed into $y =$ Rodgers$\oplus$Hammerstein and $z =$ Rodgers$\oplus$Harts, because $x = y \oplus z$ (the overlap doesn't matter!). And for these y and z, P is true, so $\star(P)(y) = 1$ and $\star(P)(z) = 1$.

This subsumes distribution down to atoms as a special case. Here is an example. Suppose in the model $\mathcal{M}_3$ that there are three boys b_1, b_2, and b_3, and each of them fell asleep. Given that this is a distributive predicate, it is not true of plural individuals (recall we dropped the Lexical Plurality Hypothesis).

(8.39)For any assignment a,

a. $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_1) = [\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_2) = [\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_3) = 1$

b. $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_1 \oplus b_2) = [\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_1 \oplus b_3)$
$= [\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_2 \oplus b_3) = [\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_3}(b_1 \oplus b_2 \oplus b_3) = 0$

Now suppose the subject is *the boys*, whose denotation in $\mathcal{M}_3$ is $b_1 \oplus b_2 \oplus b_3$. In order to account for (8.40),

(8.40)The boys fell asleep.

we pluralise the predicate via $\star$.

- By definition, $\star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})(b_1 \oplus b_2 \oplus b_3) = 1$ iff either (i) $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2}(b_1 \oplus b_2 \oplus b_3) = 1$ or (ii) $b_1 \oplus b_2 \oplus b_3 = y \oplus z$ and $\star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})(y) = \star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})(z) = 1$.]
- As before, (i) is simply not the case. So the sentence is true if $b_1 \oplus b_2 \oplus b_3$ can be decomposed into two parts y and z for which $\star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})$ is true.
- Let's choose $y = b_1$ and $z = b_2 \oplus b_3$ (there are other ways too, but we just need one way to make the sentence true).
- $\star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})(b_1) = 1$, because $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2}(b_1) = 1$.
- What about $\star([\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2})(b_2 \oplus b_3)$? Here, we need to decompose again into b_2 and b_3, because the predicate is not true of plural individuals by assumption. Suppose $y = b_2$ and $z = b_3$. Since the original predicate $[\![\textbf{fell asleep}]\!]^{a,\mathcal{M}_2}$ without $\star$ is true, the sentence is true, as desired.

So, whether the subject gets decomposed into singular atoms or not depends on the denotations of the original predicate P without $\star$. If P is only true of singular individuals, as in the case of distributive predicates, the plural subject gets decomposed into its singular parts. But if P is true of plural parts of the subject, this does not have to hold, as in the musical example (8.35).

Another (perhaps more perspicuous) way of looking at the $\star$-operator is that it closes the set characterised by P with $\oplus$.

(8.41)a. Suppose P characterises $\{a, b, c\}$
b. $\star(P)$ characterises $\{a, b, c, \quad a \oplus b, \ b \oplus c, \ a \oplus c, \quad a \oplus b \oplus c\}$

(8.42)a. Suppose Q characterises $\{a \oplus c, \quad b \oplus c\}$
b. $\star(Q)$ characterises $\{a \oplus c, \ b \oplus c, \ a \oplus b \oplus c\}$

From now on, we assume that both the plural marker -s and the distributivity operator Δ denote $\star$:

(8.43)For any model $\mathcal{M}$, and for any assignment a,
$[\![\textbf{-s}]\!]^{a,\mathcal{M}} = [\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ \lambda x \in D_e.\ \star(P)(x) = 1$

8.2.4 Distributive Readings of Object DPs

We can find distributive readings of object DPs. The synonymy of the two sentences in (8.44) demonstrates that *kissed* is distributive with respect to the object.

(8.44)a. John kissed Mary and Sue.
b. John kissed Mary and John kissed Sue.

In fact, *kissed* is also distributive with respect to the subject.

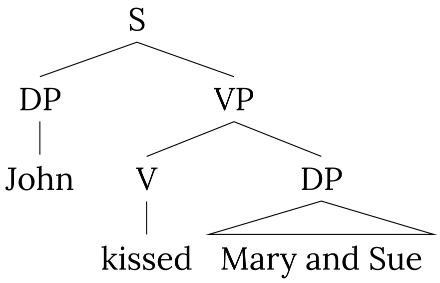
(8.45)a. John and Bill kissed Mary.
b. John kissed Mary and Bill kissed Mary.

So, given the idea that distributive predicates describe properties of singular individuals, *kissed* is only true of singular arguments (again, we don't assume the Lexical Plurality Hypothesis):

(8.46)For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{kissed}]\!]^{a,\mathcal{M}} = \lambda x \in SG. \lambda y \in SG.\ y$ kissed x in $\mathcal{M}$

But then, we have a problem with combining *likes* with the object *Mary and Sue*, because the latter denotes a plural individual. In other words, (8.47) will be trivially false.

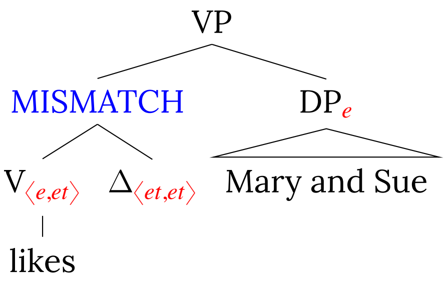
(8.47)



In order to solve this, we want to use Δ, but we cannot insert it between *likes* and *Mary and Sue*, because of the type-mismatch. Notice that the distributivity operator Δ is of type $\langle et, et \rangle$.

(8.48) For any model $\mathcal{M}$, and for any assignment a,
$$[\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ \lambda x \in D_e.\ \star(P)(x) = 1$$

(8.49)



We can do one of two things to solve this problem.

1. Postulate a distributivity operator for the object.

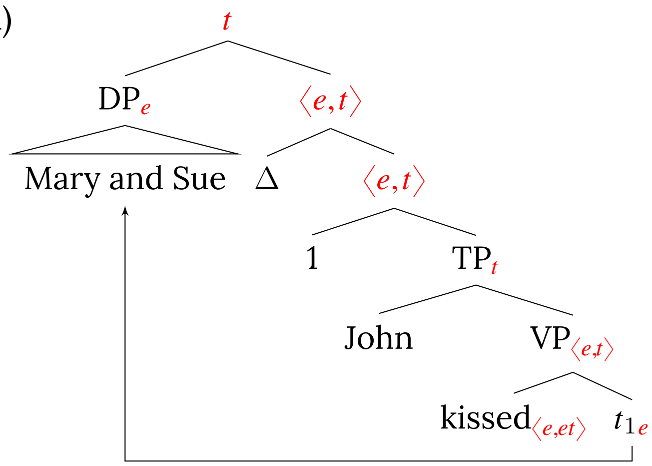
   (8.50) For any model $\mathcal{M}$, and for any assignment a,

   a. $[\![\Delta_{\textbf{subj}}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ \lambda x \in D_e.\ \star(P)(x) = 1$

   b. $[\![\Delta_{\textbf{obj}}]\!]^{a,\mathcal{M}} = \lambda R \in D_{\langle e,et \rangle}.\ \lambda x \in D_e. \lambda y \in D_e.\ \star(\lambda z \in D_e. R(z)(x) = 1)(y)$

2. Q(uantifier)R(aise) the object, and insert Δ between the landing site and the index node.

   (8.51)



In an exercise for this week, you will compute the meaning of (8.51) and verify that it does give rise to the correct meaning.

Note that the nature of the problem is very similar to the issue of quantifiers in object position. The solutions are either 'type-shift' Δ into Δ_{obj}, or move the object to resolve the type mismatch.

Due to the time constraints, we will not compare these two approaches in detail (a good Long Essay topic!), but here's one crucial piece of evidence showing that something like the movement is necessary anyway. This has to do with **phrasal distributivity**. Consider (8.52).

(8.52) John or Bill kissed the girls.

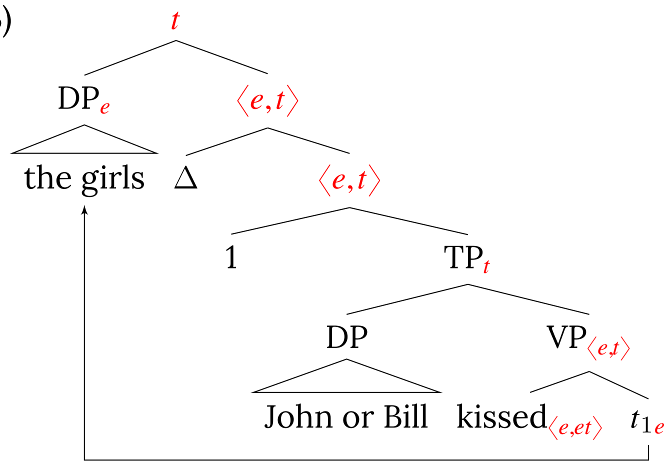
This has two readings.

- John kissed each girl or Bill kissed each girl.
- Each girl was kissed by John or Bill.

The *in situ* derivation with Δ_{obj} only gives rise to the first reading, where the girls were kissed by the same boy. Without going into the details, the reason is because the disjunction takes scope over $\star$. So it would mean, roughly, "John or Bill has the following property: he kissed each of the girls."

In order to derive the second reading, we need the following LF.

(8.53)



Again, for reasons of time, I will not go into the details of the computation, but (8.53) will mean roughly the following: "Each of the girls has the following property: John or Bill kissed her". In other words, $\star$ takes scope over the disjunction.

So, covert movement (QR) of the plural object is necessary anyway to account for phrasal distributivity. And with it, we can actually derive both of the readings, because if you QR the subject as well above the object in (8.53), the disjunction will take scope over $\star$. So for the sake of simplicity, let's assume that distributive always objects undergo QR.

However, just to stress, this is a tentative hypothesis, as we do not have arguments against also postulating Δ_{obj}. What we do know is that the object *can* undergo QR and give rise to a distributive reading via $\Delta_{\text{(subj)}}$.

8.2.5 Distributivity and QR

As depicted in the tree diagram in (8.53), in order to get a distributive reading of *the girls* over *John or Bill*, the former needs to undergo QR above *John or Bill*. This makes a prediction, namely, this covert movement is subject to the general constraints on QR. This prediction is borne out. Let us go through some examples.

Firstly, as we discussed in Term 1, finite clauses are islands for QR (not for overt wh-movement). For instance, (8.54) does not have a reading where *every girl* takes scope over *somebody*.

(8.54) Somebody said that Nathan read every paper.

a. *for every paper, somebody (possibly different for different paper) said that Nathan read it.

b. ^{ok}Somebody said something to the following effect: Nathan read every paper.

Now consider the following example:

(8.55) Andrew or Daniel said that Nathan read the five papers.

This does not have a distributive reading paraphrased as (8.56):

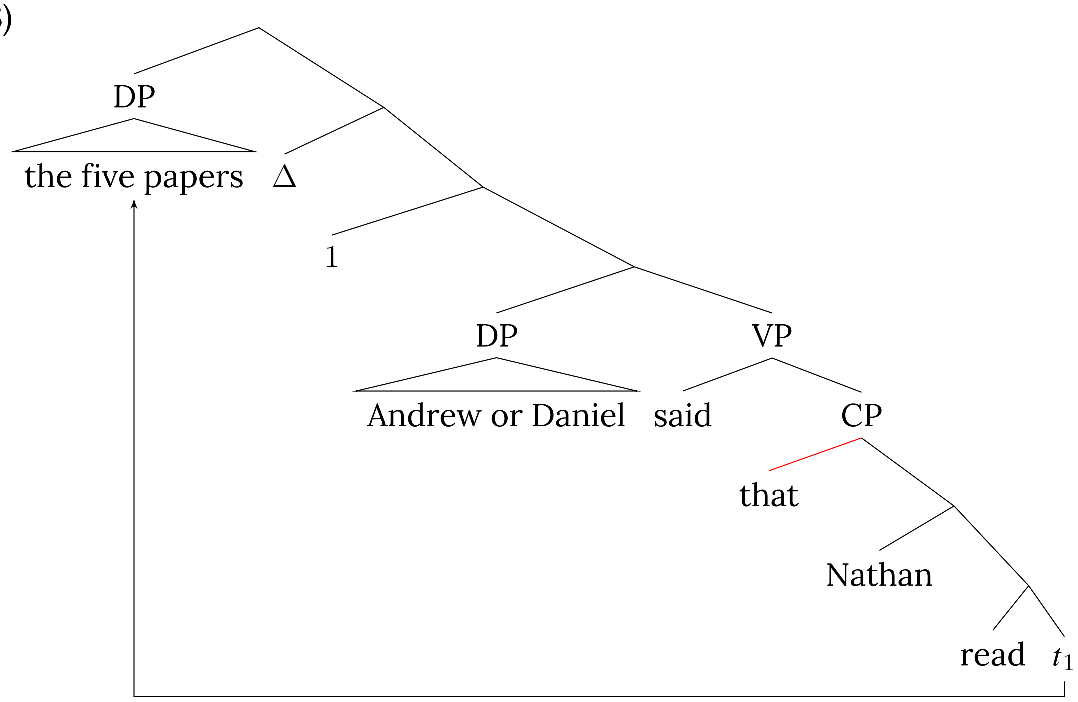
(8.56) For each of the five papers, Andrew or Daniel said that Nathan read it.

Rather, the only reading we have is:

(8.57) Andrew or Daniel said the following: Nathan read the five papers.

We have an explanation for this: in order to derive the reading in (8.56), we need to move *the five papers* above *Andrew or Daniel*, but this movement violates the locality constraint on QR.

(8.58)



Another constraint on QR is the Complex NP Island:

(8.59) Somebody invited the author who wrote every book.

a. *For every book, somebody (possibly different for different books) invited the man who wrote it.

b. [ok]Somebody invited the man who authored all the books.

We observe a parallel constraint on distributivity:

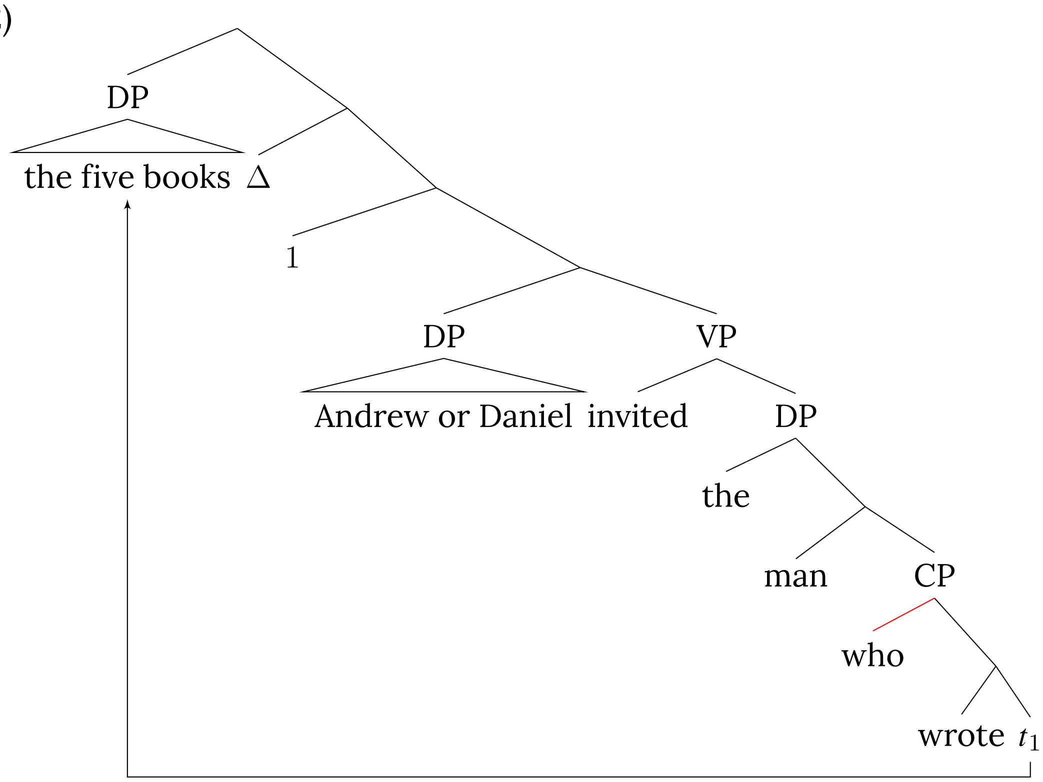
(8.60) Andrew or Daniel invited the author who wrote the five books.

We do not have the distributive reading paraphrased by (8.61).

(8.61) For each of the five books, Andrew or Daniel invited the author who wrote it.

Again, in order to derive this reading, QR has to cross an island.

(8.62)



8.3 Cumulative Reading

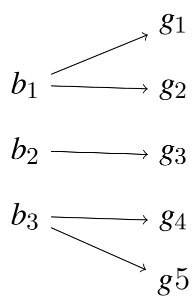
We have discussed two readings that plural DPs give rise to, namely distributive and collective readings. There is actually one more reading called **cumulative readings**. A cumulative reading arises when there are more than one plural noun phrase in the sentence.

(8.63)The three boys kissed the five girls.

This sentence has (at least) the following two readings:

- <u>Double-distributive reading:</u>
  Each of the three boys kissed each of the five girls.
- <u>Cumulative reading:</u>
  Each of the boys kissed at least one of the girls, and each of the girls was kissed by at least one of the boys.

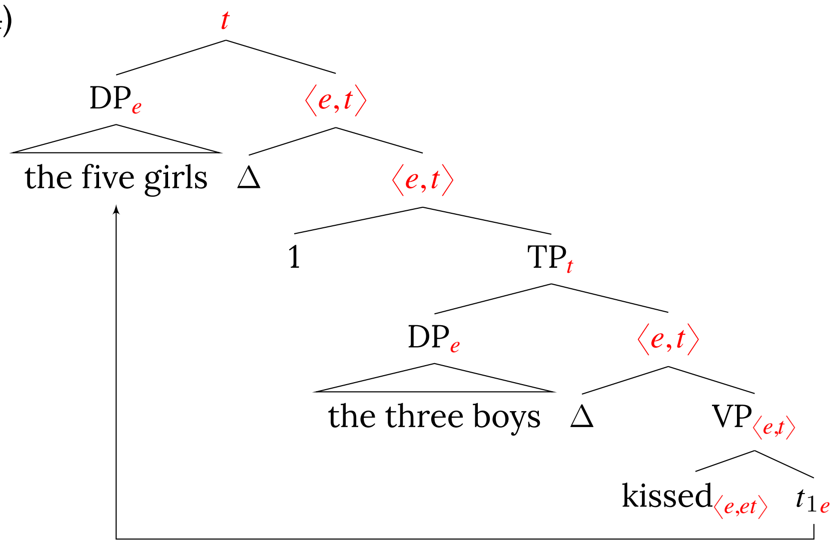
Only the cumulative reading is true in the following situation:



How can we derive these readings? Notice first of all that without Δ, the sentence will be trivially false, because *kissed* is distributive with respect to both arguments.

If we use two instances of Δ, we will have something like (8.64):

(8.64)



Roughly, this means: "for each of the five girls, the following is the case: for each of the boys, they stand in a kissing relation." This is the double-distributive reading.

In order to derive the other reading, we need another operator. But before going there, let's look at some more examples:

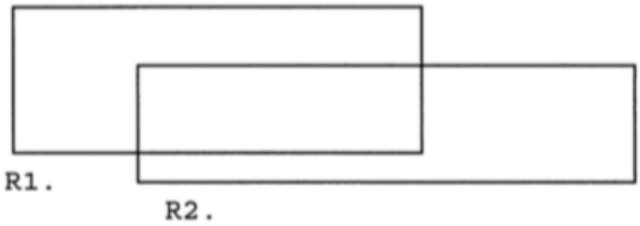
(8.65) John and Bill kissed five girls.

a. Double-distributive reading with wide scope obj:
   There is a group of five girls. Each of John and Bill kissed each of them. (5 girls)

b. Double-distributive reading with narrow scope obj
   There is a group of five girls each of whom John kissed. And there is a group of five girls each of whom Bill kissed (up to 10 girls)

c. Cummulative reading:
   There is a group of five girls. Each of John and Bill kissed at least one of them, and each of them was kissed by John or Bill (or both).

(8.66)The sides of R1 run parallel to the sides of R2.



(Schwarzschild 1996:82)

8.3.1 The Cumulativity Operator Γ and $\star\star$

In order to derive the cumulative reading, we define a new operator (in the mata-language), $\star\star$, that applies to a relation (i.e. a function of type $\langle e, et\rangle$).

(8.67)For any relation R of type $\langle e, et\rangle$,

$$\star\star(R) = \lambda x \in D_e.\lambda y \in D_e.\left[\begin{array}{l} R(x)(y) = 1 \text{ or} \\ \quad x = x_1 \oplus x_2 \text{ and } y = y_1 \oplus y_2 \text{ and} \\ \quad \star\star(R)(x_1)(y_1) = \star\star(R)(x_2)(y_2) = 1 \end{array}\right]$$

This operator applies to a two-place predicate R and distributes over the two arguments at the same time, so to speak. Let's us see how it works with a concrete example.

(8.63)The three boys kissed the five girls.

- Suppose that in $\mathcal{M}_4$, there are three boys, b_1, b_2 and b_3 and five girls g_1, g_2, g_3, g_4 and g_5. Then,

  (8.68)For any assignment a,
  a. $[\![\textbf{the three boys}]\!]^{a,\mathcal{M}_4} = b_1 \oplus b_2 \oplus b_3$
  b. $[\![\textbf{the five girls}]\!]^{a,\mathcal{M}_4} = g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5$

- In $\mathcal{M}_4$, the following kissing-relations whole:
  Let us assume the following situation:

$$\begin{tikzcd} & g_1 \\ b_1 \arrow[ur] \arrow[r] & g_2 \\ b_2 \arrow[r] & g_3 \\ b_3 \arrow[r] \arrow[dr] & g_4 \\ & g5 \end{tikzcd}$$

- The predicate *kissed* is distributive with respect to both arguments:

  (8.69)For any model $\mathcal{M}$ and for any assignment a,
  $[\![\textbf{kissed}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\lambda y \in SG.\ y$ kissed x in $\mathcal{M}$

- Now we will apply the $\star\star$-operator, and compute the meaning of the sentence:

(8.70)$[\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})](g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5)(b_1 \oplus b_2 \oplus b_3) = 1$ iff
(i) $[\![\textbf{kissed}]\!]^{a,\mathcal{M}_4}(g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5)(b_1 \oplus b_2 \oplus b_3) = 1$ or
(ii) $g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5 = x_1 \oplus x_2$ and $b_1 \oplus b_2 \oplus b_3 = y_1 \oplus y_2$ and $\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(x_1)(y_1) = 1$ and $\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(x_2)(y_2) = 1$

Here, (i) is clearly not the case, because *kissed* is only true of singular individuals. So in order for the sentence to be true, (ii) must be the case.

- Let us decompose the two arguments as follows:
  - $x_1 = g_1$
  - $x_2 = g_2 \oplus g_3 \oplus g_4 \oplus g_5$
  - $y_1 = b_1$
  - $y_2 = b_1 \oplus b_2 \oplus b_3$

  There are other ways as well, but if there's one way of satisfying (ii), the sentence will be true. Notice that $y_1 \oplus y_2 = b_1 \oplus b_2 \oplus b_3$, because the overlap doesn't count.

- First, consider $\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(x_1)(y_1)$. This is equivalent to $\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(g_1)(b_1)$, and this is 1 (True), because $[\![\textbf{kissed}]\!]^{a,\mathcal{M}_4}(g_1)(b_1) = 1$.

- The other thing to check is: $\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(x_2)(y_2)$, which in this case is:

$$\star\star([\![\textbf{kissed}]\!]^{a,\mathcal{M}_4})(g_2 \oplus g_3 \oplus g_4 \oplus g_5)(b_1 \oplus b_2 \oplus b_3)$$

  This requires further decomposition.

- In the end, you have to decompose the plural arguments to its singular parts and make sure that the kissing-relation holds. In effect, this amounts to: Each of the three boys kissed at least one of the five girls, and each of the five girls was kissed by at least one of the three boys.

Notice that the decomposition is down to the singular parts in this case, because *kissed* is distributive. If we take a mixed predicate (or a collective predicate), we might not have to go that far. Consider:

(8.71)The three men wrote 200 musicals.

write is distributive with respect to the object but mixed with respect to the subject.

(8.72)a. John and Bill wrote this paper.
b. #John wrote this paper. Bill wrote this paper.

(8.73)a. John wrote this paper and that paper.
b. John wrote this paper and John wrote that paper.

The cumulative reading of (8.71) requires decomposing the object into the singular parts, but not the subject. In other words, we have non-atomic cumulative readings.

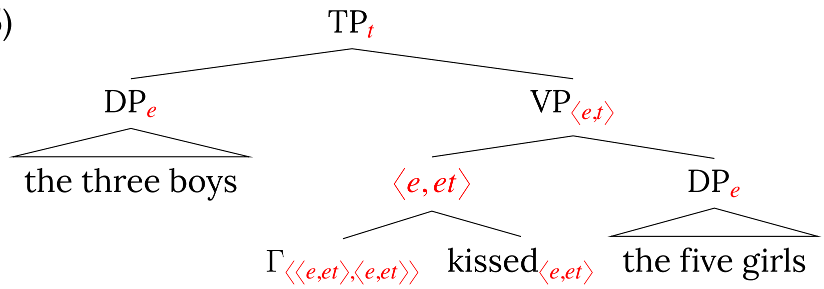
8.3.2 Lexical vs. Phrasal Cumulativity

Now, how do we implement the $\star\star$-operator? We can assume that there is an invisible operator Γ that denotes it.

(8.74) For any model $\mathcal{M}$ and for any assignment a,
$$[\![\Gamma]\!]^{a,\mathcal{M}} = \lambda R \in D_{\langle e,et\rangle}.\lambda x \in D_e.\lambda y \in D_e.\ [\star\star(R)](x)(y) = 1$$

The example (8.63) can be analysed with the following LF:

(8.75)



An alternative way of analysing this data is by assuming the Lexical Plurality Hypothesis, mentioned above. The idea is that plural for one-place predicates means $\star$, while plural for two-place predicates means $\star\star$:

(8.76) **Lexical Plurality Hypothesis** (ver. 2):
Non-nominal predicates (verbs, adjectives, prepositions) are always 'plural' (and are exempt from the Unmarked Plural Rule).
a. The plural for one-place predicate P of type $\langle e,t\rangle$ = $\star(P)$.
b. The plural for two-place predicate R of type $\langle e,et\rangle$ = $\star\star(R)$.

If we assume (8.76), we don't need to postulate Γ to account for the example (8.63).

However, for essentially the same reason why we cannot do away with Δ, we cannot do away with Γ. That is, we need to be able to compute the cumulative reading at a phrasal level (**phrasal cumulativity**), in addition to a lexical level (**lexical cumulativity**).

Consider the following example:

(8.77) The three boys kissed or hit the five girls.

This example has several readings, but let's consider the following reading: For each of the boys, there is a girl who he kissed or hit, and for each of the five girls, there is a boy who kissed or hit her. This reading requires applying Γ at a phrasal, non-lexical level. Here are the details:

- We assume the following version of *or* (which is a special case of generalised disjunction given in (8.28)):

(8.78) $[\![\mathbf{or}_{\langle\langle e,et\rangle,\langle\langle e,et\rangle,\langle e,et\rangle\rangle\rangle}]\!]^{a,\mathcal{M}} = \lambda R \in D_{\langle e,et\rangle}.\lambda S \in D_{\langle e,et\rangle}.\lambda x \in D_e.\lambda y \in D_e.\ R(x)(y) = 1 \text{ or } S(x)(y) = 1$

- The Lexical Plurality Hypothesis says that the verbs *kissed* and *hit* are already pluralised, i.e. it comes with $\star\star$:

(8.79) For any model $\mathcal{M}$ and for any assignment a,

a. $[\![\textbf{kissed}]\!]^{a,\mathcal{M}} = \star\star\,(\lambda x \in SG.\lambda y \in SG.\ y \text{ kissed } x \text{ in } \mathcal{M})$

b. $[\![\textbf{hit}]\!]^{a,\mathcal{M}} = \star\star\,(\lambda x \in SG.\lambda y \in SG.\ y \text{ hit } x \text{ in } \mathcal{M})$

- Then, the disjunction of these predicates is:

(8.80)
$$\left[\!\!\left[\ [\textbf{kissed}\ [\textbf{or}\ \ \textbf{hit}]]\ \right]\!\!\right]^{a,\mathcal{M}} = [\![\textbf{or}]\!]^{a,\mathcal{M}}([\![\textbf{hit}]\!]^{a,\mathcal{M}})([\![\textbf{kissed}]\!]^{a,\mathcal{M}})$$
$$= \lambda x \in D_e.\ \lambda y \in D_e.\ [\![\textbf{hit}]\!]^{a,\mathcal{M}}(x)(y) = 1 \text{ or } [\![\textbf{kissed}]\!]^{a,\mathcal{M}}(x)(y) = 1$$
$$= \lambda x \in D_e.\ \lambda y \in D_e.\ \begin{bmatrix} \star\star\,(\lambda z \in SG.\lambda w \in SG.\ w \text{ kissed } z \text{ in } \mathcal{M})(x)(y) = 1 \text{ or} \\ \star\star\,(\lambda z \in SG.\lambda w \in SG.\ w \text{ hit } z \text{ in } \mathcal{M})(x)(y) = 1 \end{bmatrix}$$

- Notice that the disjunction takes scope over the $\star\star$-operators. If we combine this with the plural arguments, say $b_1 \oplus b_2 \oplus b_3$ and $g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5$, the sentence is true iff

$\star\star\,(\lambda z \in SG.\lambda w \in SG.\ w \text{ kissed } z \text{ in } \mathcal{M})(g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5)(b_1 \oplus b_2 \oplus b_3) = 1$
or
$\star\star\,(\lambda z \in SG.\lambda w \in SG.\ w \text{ hit } z \text{ in } \mathcal{M})(g_1 \oplus g_2 \oplus g_3 \oplus g_4 \oplus g_5)(b_1 \oplus b_2 \oplus b_3) = 1$

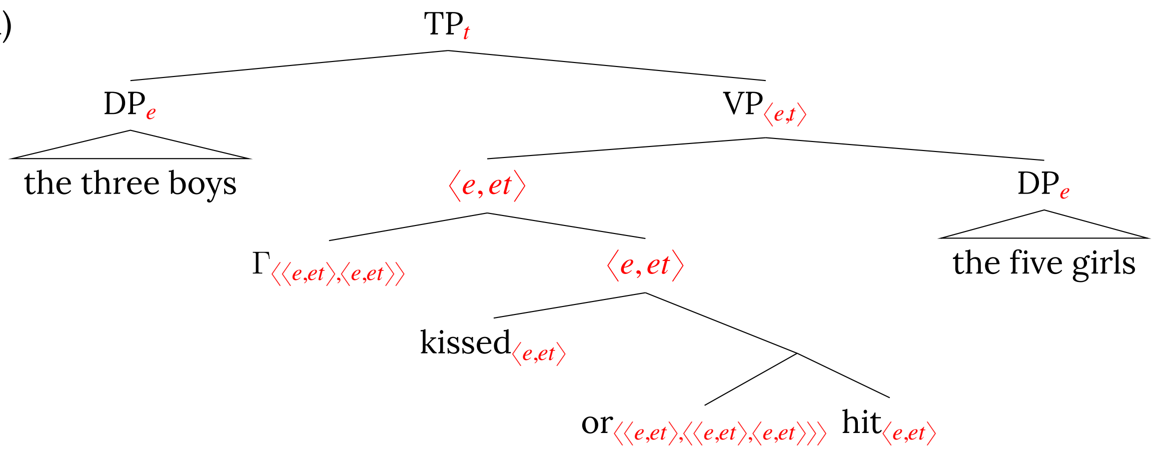
In words, the sentence is true iff either of the following is the case:

– Each of the boys kissed at least one of the girls and each of the girls was kissed by at least one of the boys; or

– each of the boys hit at least one of the girls and each of the girls was hit by at least one of the boys. This a possible reading, but not the reading we are after.

In order to derive the reading we are interested in, we need to apply the Γ-operator above the disjunction, as depicted in the following tree diagram.

(8.81)



This will generate the reading we are after (the computation is left for an exercise).

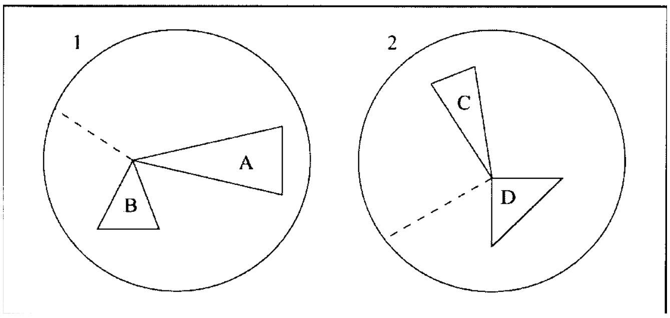
The upshot is that we need Γ for certain cases and cannot reduce all cumulative readings to a lexical issue. Again, we have no empirical reason to doubt the Lexical

Plurality Hypothesis, but since we need the cumulative operator Γ anyway, let us abandon the Lexical Plurality Hypothesis, for the sake of simplicity. As a result, every time you have a cumulative reading, you use Γ.

8.3.3 Cumulativity and QR

Certain cumulative readings involving phrasal cumulativity require QR. Consider the following sentence, which is judged true against the picture.

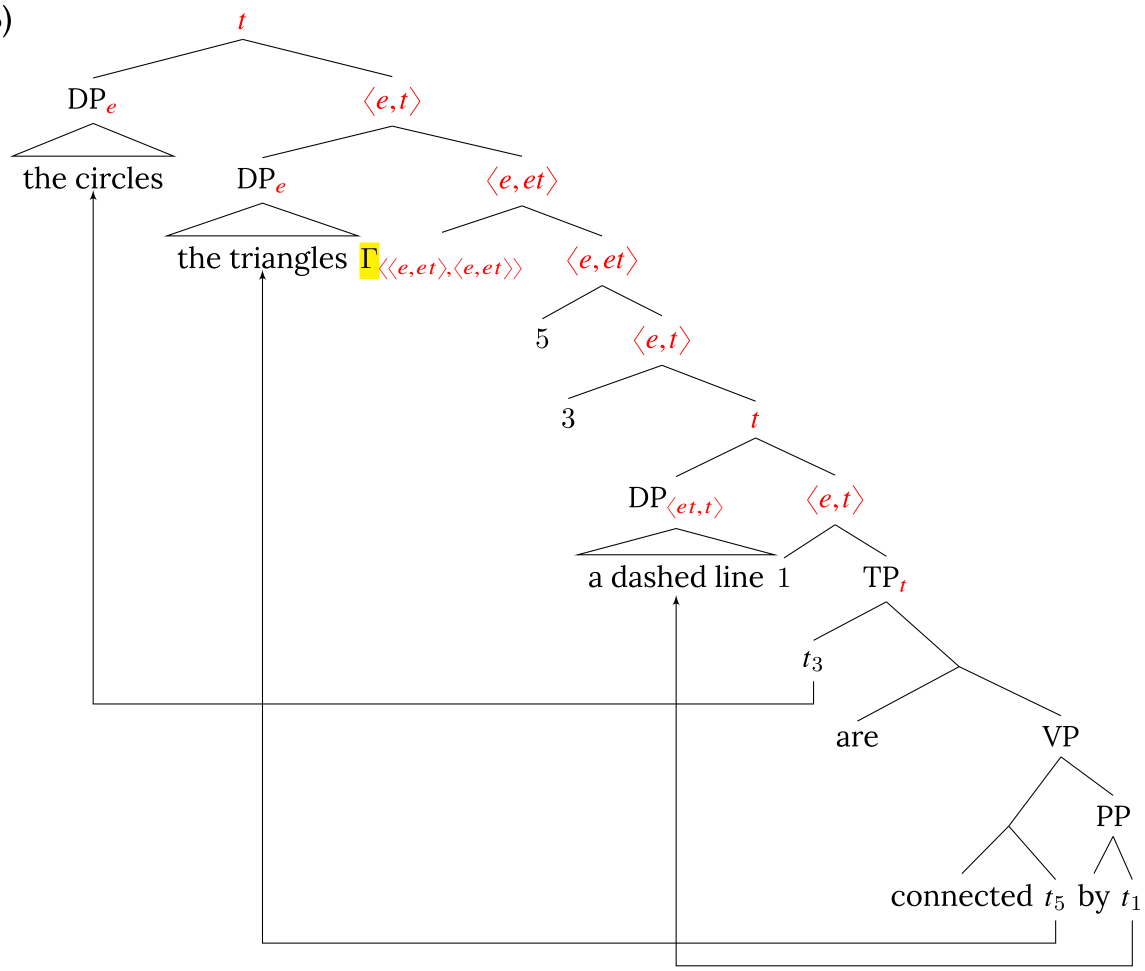
(8.82)The circles are connected to the triangles by a dashed line.



(Winter 2000:41)

Notice in particular that there are two dashed lines. This is because the cumulativity operator Γ (highlighted) takes scope over the indefinite *a dashed line*, as depicted in the following tree diagram.

(8.83)



Notice that the object *the triangles* QR. This is necessary, because we want it to be one of the arguments of the predicate that Γ creates.

OK, maybe you don't see the results but if you compute the meaning of (8.83), you will get the desired reading that is true in the above picture. The point here is that when Γ appears high in the structure, the arguments that stand in the cumulative relation need to move.

This makes a prediction: If the object is in an island, the cumulative reading does not arise. This prediction is borne out, as observed by Beck & Sauerland (2000).

Firstly, let us consider the tensed clause island. As we saw above, a quantifier cannot take scope outside of the tensed clause that contains it. For instance, the following example contrast in terms of the availability of the inverse scope reading.

(8.84)a. A lawyer has pronounced every proposal to be against the law.
b. A lawyer has pronounced that every proposal is against the law.

Unlike (8.84) with a tensed clause, (8.84a) has an inverse scope reading: For every proposal, there is a lawyer who has pronounced it to be against the law.

Now consider the following examples with two plural definites.

(8.85)a. The two lawyers have pronounced the two proposals to be against the law.

b. #The two lawyers have pronounced that the two proposals are against the law. (Beck & Sauerland 2000:365)

As indicated by #, (8.85b) does not have a cumulative reading that (8.85a) has: Each lawyer has pronounced one of the two proposals to be against the law.

Similarly, the following examples demonstrate that the covert movement is sensitive to the Complex NP island constraint.

(8.86) Andrew and John talked to a man who came from the two countries.

This cannot mean: Andrew talk to a man who came from one of the two countries and John talked to a man who came from the other country.

8.4 Further Readings

The recent overview article by Scha & Winter (to appear) is an accessible and very detailed review of the issues of distributivity and cumulativity. In addition, another overview article by Nouwen (to appear) is also recommended. For book-length discussion of these issues, see I recommend Schwarzschild (1996). Landman (2000) and Winter (2001a) are also worth reading. These books summarise the important works preceding them, including Scha (1981), Gillon (1987), Landman (1989a,b), Gillon (1992) and Lasersohn (1995).

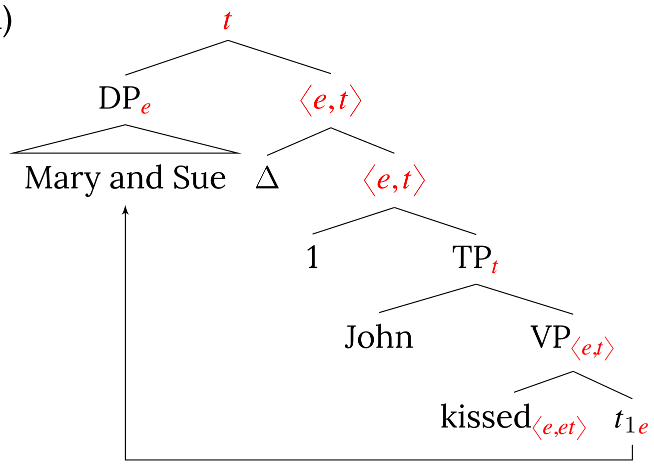
The representative works on the distributivity operator are Link (1983), Dowty (1987), Roberts (1987), Lasersohn (1995), Winter (2001a) and a series of works by Lucas Champollion. Non-atomic distributivity is discussed by Gillon (1992) and Schwarzschild (1996), among many others.

For cumulative readings in particular, the debate between Winter (2000) and Beck & Sauerland (2000) is recommendable. There is also a series of works on cumulative readings in event semantics that are very interesting, including Landman (1996, 2000) and Kratzer (2013).

8.5 Exercises

1. Compute the meaning of the following LF (with respect to $\mathcal{M}$ and a), and make sure that what you get is a distributive reading.

(8.51)



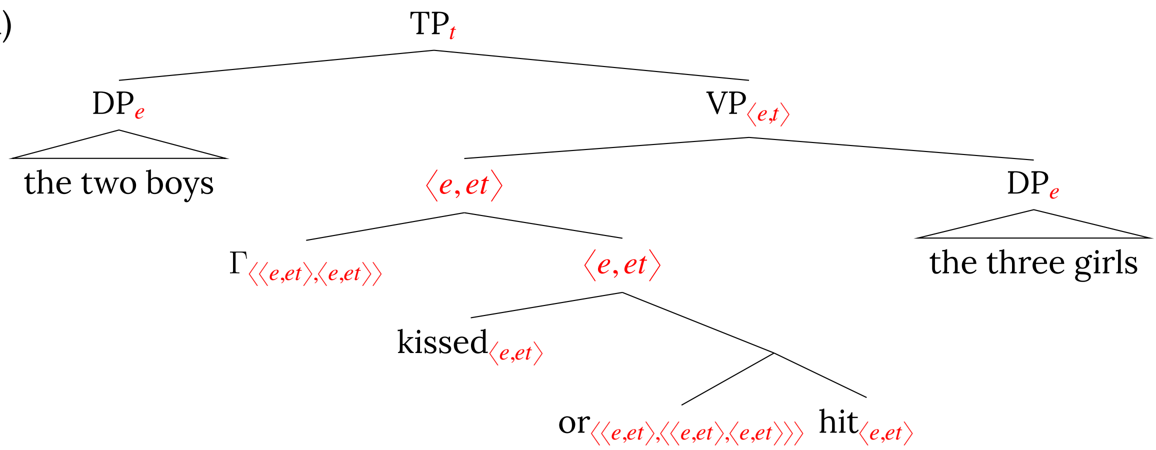
The lexical entries you need to assume are:

- $[\![\textbf{Mary}]\!]^{a,\mathcal{M}} = \text{Mary}$ $\quad [\![\textbf{Sue}]\!]^{a,\mathcal{M}} = \text{Sue}$ $\quad [\![\textbf{John}]\!]^{a,\mathcal{M}} = \text{John}$
- $[\![\Delta]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ \lambda x \in D_e.\ \star(P)(x) = 1$
- $[\![\textbf{kissed}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\lambda y \in SG.\ y \text{ kissed } x \text{ in } \mathcal{M}$

(Recall that we do *not* assume the Lexical Plurality Hypothesis)

2. Compute the meaning of the following LF (with respect to $\mathcal{M}$ and a), and make sure that what you get is a cumulative reading.

(8.81)



The lexical entries are:

- $[\![\Gamma]\!]^{a,\mathcal{M}} = \lambda R \in D_{\langle e,et \rangle}.\lambda x \in D_e.\lambda y \in D_e.\ [\star\star(R)](x)(y) = 1$
- $[\![\textbf{kissed}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\lambda y \in SG.\ y \text{ kissed } x \text{ in } \mathcal{M}$
- $[\![\textbf{hit}]\!]^{a,\mathcal{M}} = \lambda x \in SG.\lambda y \in SG.\ y \text{ kissed } x \text{ in } \mathcal{M}$

Also, assume that there are three boys, b_1 and b_2 and three girls g_1, g_2, and g_3 in the model $\mathcal{M}$. You do not need to compute the meanings of the two DPs and just use the following.

- $[\![\textbf{the two boys}]\!]^{a,\mathcal{M}} = b_1 \oplus b_2$

- $[\![\textbf{the three girls}]\!]^{a,\mathcal{M}} = g_1 \oplus g_2 \oplus g_3$

Chapter 9

Pseudo-Partitives

9.1 Pseudo-Partitives

English has two types of 'partitive' constructions (Selkirk 1977):

- (True) **partitive**:

(9.1) a. three of the students
b. many of the linguists
c. most of the books
d. both of the idiots
e. a few pieces of the bread

- **Pseudo-partitive**:

(9.2) a. two inches of cable
b. three bottles of water
c. many boxes of books
d. six pounds of cherries
e. five grains of rice

Pseudo-partitives are 'pseudo' because they are not really about one thing being sub-part of another, unlike true partitives. Rather, pseudo-partitives are about 'measurement'. Related to this point:

- True partitives involve a definite DP, as the 'whole', and a quantificational DP as the 'part' (though there are exceptions, e.g. *one of two things*; see Jackendoff 1977, Ladusaw 1982, Barker 1998 on this).
- Pseudo-partitives are formed with a bare mass or plural count noun, rather than a DP.

Pseudo-partitives involve two nouns: Q N1 of N2

- The first noun (N1) is one of the following:
  - Measure nouns: *kilo*, *inch*, *pound*, etc.
  - Container nouns: *bottle*, *glass*, *box*, etc.

– ‘Atomiser’ nouns: *grain*, *piece*, etc.

(You might also want to include *lot*, *bit*, *tad*, etc. as a fourth kind)

- The second noun (N2) describes the substance or objects in question, and is either a bare mass noun or a plural noun.

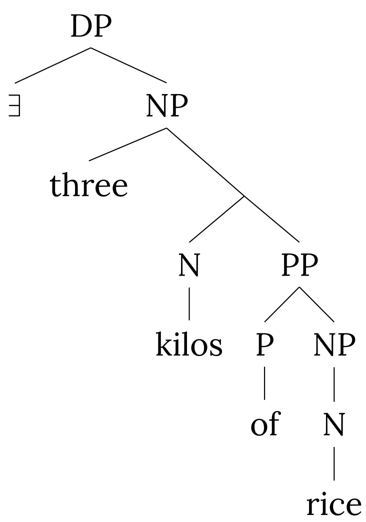
Pseudo-partitives often involve a quantificational expression Q, e.g. a numeral, *many*, *several*, etc.

9.1.1 Measure Nouns and Measure Functions

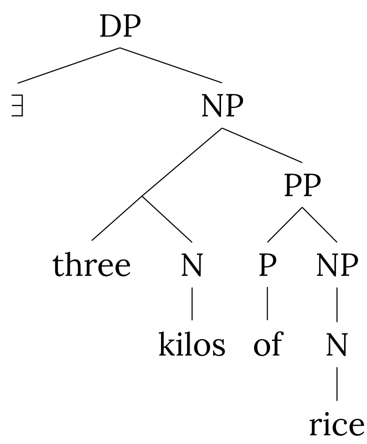
Let us start or analysis with measure nouns. We’ll come back to container nouns and ‘atomisers’ later.

Firstly, what is the syntax of pseudo-partitives? We have at least the following two possibilities. (Recall from Lecture 7 that ∃ is an invisible existential quantifier)

(9.3)



(9.4)



The choice here does not matter much for our discussion to follow, so following Selkirk (1977), Scontras (2014), among others, let’s assume (9.3) (for the other structure, see e.g. Rothstein 2009, Partee & Borschev 2012).

What is the semantics of the pseudo-partitive construction? We assume that the noun *rice* is true of any portion of rice (but remember the complications we discussed in Lecture 8). That is, if c is a portion of rice in $\mathcal{M}$, then $[\![\textbf{rice}]\!]^{a,\mathcal{M}}(c) = 1$ (for any assignment a).

We also assume that the semantic function of the partitive construction—or more precisely, of *three kilos of*—is to intersectively modify the noun *rice*. Thus, the denotation of *three kilos of rice* is a function of type $\langle e,t \rangle$ that is true of any portion of rice that weighs 3 kg.

(9.5) For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{three kilos of rice}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a portion of rice that weighs 3 kg in $\mathcal{M}$

To put it differently, the semantic function of *three kilos of* is the same as that of the relative clause in (9.6).

(9.6) rice that weighs 3 kg

For the sake of exposition, let us re-write (9.5) as (9.7).

(9.7) $[\![\textbf{three kilos of rice}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ is a portion of rice and $\mu_{\text{kg}}(x) = 3$ in $\mathcal{M}$

μ_{kg} is called a **measure function**. It takes an individual and returns its weight—which is a degree!—in terms of kilograms. Together with the existential determiner $\exists$, this will allow us to account for the meanings of sentences like (9.8).

(9.8) a. $\exists$ three kilos of rice is in the kitchen.
b. I bought $\exists$ three kilos of rice.

Now we know what kind of meaning the DP should get. What about the meanings of the individual items that make up the pseudo-partitive? Here is one of of decomposing the meaning (see Scontras 2014 for more careful discussion). Notice that *of* is semantically vacuous (i.e. denotes an identity function).

(9.9) For any model $\mathcal{M}$ and for any assignment a,
a. $[\![\textbf{three}]\!]^{a,\mathcal{M}} =$ the degree 3
b. $[\![\textbf{of}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\ P$
c. $[\![\textbf{kilo}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda d \in D_d.\lambda x \in D_e.\ P(x) = 1$ and $\mu_{\text{kg}}(x) = d$ in $\mathcal{M}$

The key is the meaning of the measure noun *kilo*. It has a degree argument d and also includes the measure function μ_{kg}.

We can assign similiar denotations to other measure nouns. The main difference is in the measure functions.

(9.10)For any model $\mathcal{M}$ and for any assignment a,
a. $[\![\textbf{metre}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda d \in D_d.\lambda x \in D_e.\ P(x) = 1$ and $\mu_{\text{m}}(x) = d$ in $\mathcal{M}$
b. $[\![\textbf{degrees}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda d \in D_d.\lambda x \in D_e.\ P(x) = 1$ and $\mu_{^\circ\text{C}}(x) = d$ in $\mathcal{M}$

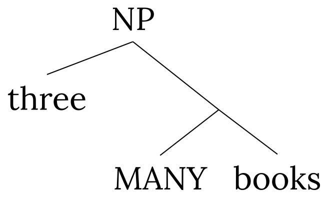
Notice that we are assuming that numerals like *three* denote degrees, rather than an intersective modifier of type $\langle e,t \rangle$ as in Lecture 8, i.e. (9.11).

(9.11)For any model $\mathcal{M}$ and for any assignment a,
$[\![\textbf{three}]\!]^{a,\mathcal{M}} = \lambda x \in D_e.\ x$ has three distinct singular parts in $\mathcal{M}$

According to (9.11), $[\![\textbf{three books}]\!]^{a,\mathcal{M}}$ characterises a set of plural individuals that have three singular parts.

Let us re-analyse numerals as degrees so that we can analyse *three books* with degree semantics. First, we assume that *three* denotes a degree. Since a degree and a noun *books* cannot directly compose, we also assume a hidden morpheme, which introduces a measure function $\mu_{\#}$, which measures the size of a given individual in terms of the number of singular parts (cf. Hackl 2000). Concretely, we postulate a covert version of *many*, MANY, as in (9.12):

(9.12)



MANY has the same semantic type as a measure noun, and introduces the measure function $\mu_{\#}$.

(9.13) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{MANY}]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t \rangle}.\lambda d \in D_d.\lambda x \in D_e.\ P(x) = 1 \text{ and } \mu_{\#}(x) = d \text{ in } \mathcal{M}$$

So the only difference from the pseudo-partitive construction is that MANY is phonologically null and the semantically vacuous item *of* is missing.

The predicted meanings are the same as before:

(9.14) a. $[\![\textbf{MANY books}]\!]^{a,\mathcal{M}} = [\![\textbf{MANY}]\!]^{a,\mathcal{M}}([\![\textbf{books}]\!]^{a,\mathcal{M}})$
$= [\lambda P \in D_{\langle e,t \rangle}.\lambda d \in D_d.\lambda x \in D_e.\ P(x) = 1 \text{ and } \mu_{\#}(x) = d \text{ in } \mathcal{M}]([\![\textbf{books}]\!]^{a,\mathcal{M}})$
$= \lambda d \in D_d.\lambda x \in D_e.\ [\![\textbf{books}]\!]^{a,\mathcal{M}}(x) = 1 \text{ and } \mu_{\#}(x) = d \text{ in } \mathcal{M}$

b. $[\![\textbf{three MANY books}]\!]^{a,\mathcal{M}} = [\![\textbf{MANY books}]\!]^{a,\mathcal{M}}([\![\textbf{three}]\!]^{a,\mathcal{M}})$
$= [\![\textbf{MANY books}]\!]^{a,\mathcal{M}}(3)$
$= \lambda x \in D_e.\ [\![\textbf{books}]\!]^{a,\mathcal{M}}(x) = 1 \text{ and } \mu_{\#}(x) = 3 \text{ in } \mathcal{M}$
$= \lambda x \in D_e.\ \left[\begin{array}{l}\text{each singular part of } x \text{ is a book in } \mathcal{M} \text{ and} \\ \mu_{\#}(x) = 3 \text{ in } \mathcal{M}\end{array}\right]$

So what we did is to 'decompose' the meaning of *two* into the degree part *two* and the measurement part MANY. This allows us to uniformly analyse various constructions involving numerals, including pseudo-partitives.

Notice that since degrees are more abstract than individuals *per se*, this alternative analysis opens up a possibility of analysing sentences like (9.15).

(9.15) There are one and a half/1.5 apples.

According to our analysis, this sentence is true iff there is a (plural) individual x that is in the denotation of *apples* and $\mu_{\#}(x) = 1.5$.

However, notice that there is a problem. In order for the sentence to be true, there must be one apple and a half-apple, but the latter is not an apple! Since the denotation of *apples* only includes whole apples and their combinations, whatever x that measures 1.5 with respect to $\mu_{\#}$ is not in the denotation of *apples*!! Thus, we need to say a bit more to analyse sentences like (9.15), in particular about the treatment of **sub-atomic** individuals in the denotation. This is left open here.

9.1.2 Monotonicity Constraint

Going back to pseudo-partitives, Schwarzschild (2006) points out that this construction has a constraint on what kind of measure functions can be used. The constraint is illustrated by (9.16) vs. (9.17).

(9.16) a. 30 kg of water
b. 180 cm of rope
c. five litres of oil
d. 10 kilos of books
e. two hours of driving

(9.17) a. *40 degrees of water
b. *180 cm of men
c. *80 km/h of driving

According to our semantics noun phrases like (9.17) should make sense, e.g. (9.17a) should denote a function of type $\langle e,t\rangle$ that is true of any portion of water that is 40°C.

Schwarzschild (2006) proposes the following constraint (see Krifka 1998 proposes a stronger version in terms of 'extensiveness'):

(9.18) Pseudo-partitives are only compatible with measure functions that are **monotonic** on the part-whole relation in the domain of the noun.

The notion of monotonicity is defined as (9.19).

(9.19) A measure function μ_D is monotonic on the part-whole relation in the domain of the noun N iff for any x, y that are Ns such that $y \sqsubset x$, x measures more than y along the dimension, i.e. $\mu_D(y) \prec_D \mu_D(x)$.

The intuition is this: The measure function must track the part-whole relation.

- *30 kg of water* is acceptable, because μ_{kg} is monotonic on the part-whole relation. If you have a portion of water, x, and take a sub-portion of it, y, then the weight of the original portion in kilograms, $\mu_{\text{kg}}(x)$, is necessarily more than the weight of the sub-portion in kilograms, $\mu_{\text{kg}}(y)$.
- 40 *degrees of water* is unacceptable, because temperature is not monotonic on the part-whole relation. If you have a portion of water, w_1 and take a sub-portion of it, w_2, then the temperature of w_2 is the same as that of w_1!

Another example from Schwarzschild (2006:74): *karat* has two meanings, one has to do with the purity, and one has to do with the weight. Only weight is monotonic. Consequently, (9.20) only has a weight reading.

(9.20) 18 karat of gold

Schwarzschild (2006) further discusses a different modification construction, which he calls the **attributive construction**, where the dimension is required to be non-monotonic on the part-whole relation in the domain of the noun. For example, (9.21) now only has a purity reading.

(9.21) 18 karat gold

In this construction, dimensions like temperature are acceptable.

(9.22) 40 degree water

Also consider the contrast between (9.23a) and (9.23b):

(9.23) a.	2 inches of cable	(length)	
b.	2 inch cable	(diameter)	(Schwarzschild 2006:74)

Both of these examples are about length, but while the length of the cable is monotonic, the diameter of the cable is not monotonic.

So in the attributive construction, a dimension that is not monotonic is felicitous. The converse is also true: a non-monotonic dimension gives rises to infelicity. However, there is a caveat: as Schwarzschild notices, singular count nouns, unlike mass nouns and plural count nouns, allow any kind of modification in the attributive construction. This contrast is shown with monotonic measure functions in (9.24).

(9.24) Singular	Mass	
2 hour job	*2 hour work	(duration)
2 hour trip	*2 hour traveling	(duration)
2 millilitre drop	*2 millilitre blood	(volume)
2 pound bean	*2 pound coffee	(weight)
2 page poem	*2 page poetry	(page count)
		(Schwarzschild 2006:77)

In order to make sense of this observation, Schwarzschild defines the notion of non-monotonicity as follows:

(9.25) A measure function μ_D is non-monotonic on the part-whole relation in the domain of the noun N iff for any x, y that are Ns such that $y \sqsubseteq x$, then x measure the same as y along the dimension, i.e. $\mu_D(y) =_D \mu_D(x)$.

Given that singular count nouns are only true of atomic individuals—which do not have parts other than themselves—any dimension becomes non-monotonic. Thus, the constraint on the attributive construction can be stated as (9.26).

(9.26) Attributives are only compatible with measure functions that are **non-monotonic** on the part-whole relation in the domain of the noun.

Schwarzschild (2006) claims that the (non-)monotonicity requirement is syntactically encoded in a particular way. The rough idea is that if the modifier appears high in the structure and far from the noun, it needs to be monotonic, and if it appears low in the structure and close to the noun, it needs to be non-monotonic. He also observes that similar constraints hold in a number of other languages.

One interesting observation in this connection is that adjectives also show similar constraints. In the attributive position, you only have a distributive reading, which is non-monotonic, while in the predicative position (which is outside the DP), it also gives rise to a collective reading (see Schwarzschild 2006:87):

(9.27)a. the heavy bottles
b. The bottles are heavy.

9.1.3 Container Nouns

When N1 of the pseudo-partitive is a container noun, there are two possible interpretations (Rothstein 2009, Scontras 2014).[1]

- **Container reading**: What is referred to is a container or containers with N2 substance in it.
- **Measure reading**: What is referred to is a portion of N2 substance. The container itself might or might not exist.

This ambiguity is illustrated by the following examples:

(9.28)a. I knocked off two glasses of wine. (container reading)
b. I put two glasses of wine in the soup. (measure reading)

(9.28a) is about two glasses, because you can't knock off wine. For (the most prominent reading of) (9.28b), on the other hand, there does not have to be two glasses (maybe I only used one glass, or maybe even none!). Also, the glasses themselves (if I used them) are not in the soup!

Similarly, the following examples favour one reading over the other (for pragmatic reasons).

(9.29)a. John will carry these boxes of books upstairs. (container reading)
b. There are two beautiful bottles of wine here. (container reading)
c. I drank two cups of coffee today. (measure reading)
d. We have three bowls of soup in the pot. (measure reading)

Some examples are ambiguous:

(9.30)John bought some bottles of wine.

Also, the measure reading can be forced by the use of *-ful*(s):

(9.31)a. #I knocked off two glassful(s) of wine!
b. I poured two glassful(s) of wine in the soup.

(9.32)a. Bring two cupfuls of wine for our guests.
b. We needed three bucketfuls of cement to build that wall.
c. Three bucketfuls of mud were standing in a row against the wall.
(adapted from Rothstein 2009:110)

In order to analyse this ambiguity, we can assign two different meanings to container nouns. The first meaning, (9.33), derives the container reading.

[1] Partee & Borschev (2012) argue that there are more distinctions

(9.33) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{glass}_c]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda x \in SG. \begin{bmatrix} x \text{ is a glass in } \mathcal{M} \text{ and} \\ \text{there is } y \text{ such that } P(y) = 1 \text{ and} \\ y \text{ is in } x \text{ in } \mathcal{M} \end{bmatrix}$$

This takes the meaning of N2, P and existentially quantifies over a member of it. For example, $[\![\textbf{glass of wine}]\!]^{a,\mathcal{M}}$ is true of any glass that contains some wine. Since this is a noun, we can pluralise it and obtain *glasses of wine*, which can compose with MANY and a numeral (but notice that we have not explicitly defined pluralisation for transitive nouns like this one; see the discussion on cumulative readings from Lecture 9).

In addition to this, we postulate the following entry, where μ_{glass} returns the quantity of x in terms of the size of a (standard/contextually salient) glass.

(9.34) For any model $\mathcal{M}$ and for any assignment a,

$$[\![\textbf{glass}_m]\!]^{a,\mathcal{M}} = \lambda P \in D_{\langle e,t\rangle}.\lambda y \in D_e.\lambda D_d.\ P(y) = 1 \text{ and } \mu_{\text{glass}}(y) = d \text{ in } \mathcal{M}$$

This works exactly like measure nouns like *kilos*.

Thus, the idea here is that container nouns are ambiguous between a concrete noun use and a measure noun use. (Partee & Borschev (2012) and Scontras (2014:§3.2.3) suggest that the basic meaning is the former, concrete noun meaning (9.33) and the measure use (9.34) is derived from it).

9.1.4 Atomiser Nouns

Finally, let us briefly discuss the meanings of what Scontras (2014) calls 'atomiser nouns', e.g. *piece*, *grain*, *drop*, etc. The idea underlying this coinage is that they turn a mass noun into a count noun by specifying what atoms/singular individuals are.

Scontras (2014) develops an analysis making use of 'mereo-topological' notions, but without going into the details, the idea is that *grain of rice* is only true of each grain of rice. This is unlike the mass noun *rice*, which can be true of any combinations or sub-parts of grans of rice. A key feature of such atoms induced by atomisers is that they have spacial consistency, which should be coming from the meanings of atomisers.

For the details, see Scontras (2014:Ch.3).

9.2 Further Readings

Selkirk (1977) is one of the first studies on partitives and pseudo-partitives in formal linguistics. For the semantics of pseudo-partitives, see Schwarzschild (2006), Rothstein (2009), Partee & Borschev (2012), and Scontras (2014). All of these works are very well written and highly recommendable.

Bibliography

Abrusán, Márta & Benjamin Spector. 2011. A semantics for degree questions based on intervals: negative islands and their obviation. *Journal of Semantics* 28(1). 107–147. doi:10.1093/jos/ffq013.

Abrusán, Márta & Kriszta Szendrői. 2013. Experimenting with the king of France: Topics, verifiability and definite descriptions. *Semantics and Pragmatics* 6(10). 1–43. doi:http://dx.doi.org/10.3765/sp.6.10.

Anderson, Curt & Marcin Morzycki. to appear. Degrees as kinds. *Natural Language and Linguistic Theory* .

Bach, Emmon. 1986. Natural language metaphysics. In Ruth Barcan Marcus (ed.), *Logic, Methodology and Philosophy of Science*, vol. VII, 573–595. Amsterdam: Elsevier.

Bale, Alan. 2008. A universal scale of comparison. *Linguistics and Philosophy* 31(1). 1–55. doi:10.1007/s10988-008-9028-z.

Bale, Alan & David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26(3). 217–252. doi:10.1093/jos/ffp003.

Barker, Chris. 1992. Group terms in English: Representing groups as atoms. *Journal of Semantics* 9. 69–93.

Barker, Chris. 1998. Partitives, double genitives and anti-uniqueness. *Natural Language and Linguistic Theory* 16(4). 679–717. doi:10.1023/A:1005917421076.

Barner, David, Shunji Inagaki & Peggy Li. 2009. Language, thought, and real nouns. *Cognition* 111(3). 329–344. doi:10.1016/j.cognition.2009.02.008.

Barner, David & Jesse Snedeker. 2005. Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97(1). 41–66. doi:10.1016/j.cognition.2004.06.009.

Barner, David & Jesse Snedeker. 2008. Compositionality and statistics in adjective acquisition: 4-year-olds interpret *tall* and *short* based on the size distributions of novel noun referents. *Child Development* 79(3). 594–608. doi:10.1111/j.1467-8624.2008.01145.x.

Bartsch, REnate & Theo Vennemann. 1975. The grammar of adjectives and comparison. *Linguistische Berichte* 20. 19–32.

Beall, Jc & Michael Glanzberg. 2014. Liar Paradox. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2014 edn.

Beaver, David. 2001. *Presupposition and Assertion in Dynamic Semantics*. Stanford: CSLI.

Beaver, David & Bart Geurts. 2013. Presupposition. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, fall 2013 edn.

Beaver, David & Emiel Krahmer. 2001. A partial account of presupposition projec-

tion. *Journal of Logic, Language, and Information* 10(2). 147–182. doi:10.1023/A:1008371413822.
Beck, Sigrid. 2011. Comparative constructions. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, 1341–1390. de Gruyter.
Beck, Sigrid. 2014. Plural predication and quantified *than*-clauses. In Luka Crnič & Uli Sauerland (eds.), *The Arts and Craft of Semantics: A Festschrift for Irene Heim*, vol. 1, 91–115. Cambridge, MA: MITWPL.
Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst & Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9(1). 1–66. doi:10.1075/livy.9.01bec.
Beck, Sigrid, Toshiko Oda & Koji Sugisaki. 2004. Parametric variation in the semantics of comparison: Japanese vs. English. *Journal of East Asian Linguistics* 13. 289–344. doi:10.1007/s10831-004-1289-0.
Beck, Sigrid & Uli Sauerland. 2000. Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics* 8(4). 349–371. doi:10.1023/A:1011240827230.
Bennett, Michael. 1974. *Some Extensions of a Montague Fragment of English*: University of California Los Angeles Ph.D. dissertation.
van den Berg, Martin. 1996. *Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora*: Universiteit van Amsterdam Ph.D. dissertation.
Bhatt, Rajesh & Shoichi Takahashi. 2011. Reduced and unreduced phrasal comparatives. *Natural Language and Linguistic Theory* 29(3). 581–620.
Bierwisch, Manfred. 1988. Tools and explanations of comparison: Part 1. *Journal of Semantics* 6(1). 57–93. doi:10.1093/jos/6.1.57.
Bierwisch, Manfred. 1989. The semantics of gradation. In Manfred Bierwisch & Ewald Lang (eds.), *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*, 71–237. Springer.
Bolinger, Dwight. 1972. *Degree Words*. The Hague: Mouton.
Borer, Hagit. 2005. *In Name Only*. Oxford: Oxford University Press.
Breakstone, Micha Y. 2012. Inherent evaluativity. In Ana Aguilar Guevara, Anna Chernilovskaya & Rick Nouwen (eds.), *Proceedings of Sinn und Bedeutung 16*, vol. 1, 113–126. Cambridge, MA: MITWPL.
Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4. 275–343.
Bresnan, Joan. 1975. Comparative Deletion and constraints on transformations. *Linguistic Analysis* 1(4). 25–74.
Brisson, Christine. 2003. Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26(2). 129–184. doi:10.1023/A:1022771705575.
Bunt, Harry Cornelius. 1985. *Mass Terms and Model-Theoretic Semantics*. Cambridge: Cambridge University Press.
Büring, Daniel. 2007. Cross-polar nomalies. In *Proceedings of SALT 17*, 37–52.
Büring, Daniel. 2009. *More* or *less*. In *Proceedings of Chicago Linguistic Society 43*, .
Burnett, Heather. 2012. (A)symmetric vagueness and the absolute/relative distinction. In Ana Aguilar Guevara, Anna Chernilovskaya & Rick Nouwen (eds.), *Proceedings of Sinn und Bedeutung 16*, vol. 1, 127–141. Cambridge, MA: MITWPL.
Burnett, Heather. 2014. A delineation solution to the puzzles of asbolute adjectives. *Linguistics and Philosophy* 37(1). 1–39. doi:10.1007/s10988-014-9145-9.

Bylinina, Lisa. 2013. *The Grammar of Standards*: Universiteit Utrecht dissertation.
Carlson, Gregory. 1977. *Reference to Kinds in English*: University of Massachusetts, Amherst dissertation.
Carlson, Gregory & Jeffry Pelletier (eds.). 1995. *The Generic Book*. Chicago: University of Chicago Press.
Champollion, Lucas. 2014. Algebraic Semantics and Mereology. Lecture notes for LING-GA 3340.002, New York University.
Champollion, Lucas & Manfred Krifka. 2014. Mereology. Ms. to be published in Paul Dekker and Maria Aloni (eds.), *Cambridge Handbook of Semantics*.
Cheng, Chung-Ying. 1973. Comments on Moravcsik's paper. In Jaakko Hintikka, Gabriël Nuchelmans & Wesley C. Salmon (eds.), *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, 286–288. Springer. doi:10.1007/978-94-010-2506-5_14.
Cheng, Lisa Lai-Shen, Jenny Doetjes & Rint Sybesma. 2008. How universal is the Universal Grinder? *Linguistics in the Netherlands* 12(1). 50–62. doi:10.1075/avt.25.08che.
Cheng, Lisa Lai-Shen, Jenny Doetjes, Rint Sybesma & Roberto Zamparelli. 2012. On the interpretation of number and classifiers. *Rivista di Linguistica* 24(2). 175–194.
Cheng, Lisa Lai-Shen & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542. doi:10.1162/002438999554192.
Cheng, Lisa Lai-Shen & Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In Guglielmo Cinque & Richard Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 259–292. Oxford: Oxford University Press.
Cheng, Lisa Lai-Shen & Rint Sybesma. 2012. Classifiers and DP. *Linguistic Inquiry* 43(4). 634–650. doi:10.1162/ling_a_00109.
Cheung, Pierina, Peggy Li & David Barner. 2012. What counts in Mandarin Chinese: A study of individuation and quantification. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 210–215.
Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of 'semantic parameter'. In Susan Rothstein (ed.), *Events and Grammar*, 53–103. Dordrecht: Kluwer.
Chierchia, Gennaro. 1998b. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405. doi:10.1023/A:1008324218506.
Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening and the "Logicality" of language. *Linguistic Inquiry* 37(4). 535–590.
Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174(1). 99–149. doi:10.1007/s11229-009-9686-6.
Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. Scalar implicature as a grammatical phenomenon. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, 2297–2331. Berlin: de Gruyter.
Chomsky, Noam. 1977. On wh-movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal Syntax*, 71–132. New York: Academic Press.
Chung, Sandra & William Ladusaw. 2003. *Restriction and Saturation*. Cambridge, MA: MIT Press.
Cooper, Robin. 1983. *Quantification and Semantic Theory*. Dordrecht: Reidel.
Cresswell, Max J. 1976. The semantics of degree. In Barbara H. Partee (ed.), *Montague Grammar*, New York: Academic Press.

Cummins, Chris, Uli Sauerland & Stephanie Solt. 2012. Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy* 35(2). 135–169. doi:10.1007/s10988-012-9114-0.

Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam McHombo & Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2). 159–210.

Dayal, Veneeta. 2011. Bare noun phrases. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, 1088–1109. de Gruyter.

van Deemter, Kees. 2010. *Not Exactly: In Praise of Vaguness*. Oxford: Oxford University Press.

Delfitto, Denis. 2005. Bare plurals. In Martin Evaraert & Henk van Riemsdijk (eds.), *Blackwell Comparison to Syntax*, vol. 1, 214–259. Malden: Blackwell.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

van der Does, Jaap. 1993. The dynamics of sophisticated laziness. Ms., Universiteit van Amsterdam.

Doetjes, Jenny. 2012. Count/mass distinctions across languages. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Meaning*, vol. 3, 2559–2580. Berlin: de Gruyter.

Doetjes, Jenny, Camelia Constantinescu & Kateřina Součková. 2011. A neo-Kleinian approach to comparatives. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Proceedings of SALT 19*, 124–141.

Dotlačil, Jakub & Rick Nouwen. 2014. The comparative and degree pluralities. Ms.

Dowty, David. 1987. Collective predicates, distributive predicates and *all*. In *Proceedings of ESCOL 3*, .

Fara, Delia Graff. 2000. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics* 28(1). 45–81. doi:10.5840/philtopics20002816. Originally published under the name Delia Graff.

Farkas, Donka & Henriëtte de Swart. 2010. The semantics and pragmatics of plurals. *Semantics & Pragmatics* 3. 1–54.

Fine, Kit. 1975. Vagueness, truth and logic. *Synthese* 30. 265–300.

von Fintel, Kai. 2004. Would you believe it? The king of France is back! (Presuppositions and truth-value intuitions). In Marga Reimer & Anne Bezuidenhout (eds.), *Descriptions and Beyond*, 315–341. Oxford: Oxford University Press.

Fox, Danny. 2006. Free choice and the theory of scalar implicatures. Ms., MIT.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. doi:10.1007/s10988-006-9004-4.

Fults, Scott. 2006. *The Structure of Comparison: An Investigation of Gradable Adjectives*: University of Maryland dissertation.

Fults, Scott. 2010. Vagueness and scales. In Paul Égré & Nathan Klinedinst (eds.), *Vagueness and Language Use*, 25–50. Hampshire: Palgrave McMillan.

Gajewski, Jon. 2002. L-Analyticity and Natural Language. Ms., MIT.

Gajewski, Jon. 2008. More on quantifiers in comparative clauses. In *Proceedings of SALT 18*, .

van Geenhoven, Veerle. 1998. *Semantic Incorporation and Indefinite Descriptions*. Stanford: CSLI.

George, Benjamin. 2008. A new predictive theory of presupposition projection. In *Proceedings of SALT 18*, 358–375.

Geurts, Bart & Rick Nouwen. 2007. "At least" et al.: the sematnics of scalar modifiers. *Language* 83. 533–559.
Gillon, Brendan. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy* 10(2). 199–219. doi:10.1007/BF00584318.
Gillon, Brendan. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15(6). 597–640. doi:10.1007/BF00628112.
Gillon, Brendan. 1999. The lexical semantics of English count and mass nouns. In Evelyne Viegas (ed.), *The Breadth and Depth of Semantic Lexicons*, 19–37. Dordrecht: Kluwer.
Grimm, Scott. 2012. Degrees of countability: A mereotopological approach to the mass/count distinction. In *Proceedings of SALT* 22, 584–603.
Hackl, Martin. 2000. *Comparative Quantifiers*: Massachusetts Institute of Technology Ph.D. dissertation.
Hankamer, Jorge. 1973. Why there are two *than*'s in English. In Claudia Corum, T. Cedric Smith-Stark & Ann Weiser (eds.), *Chicago Linguistic Society* 9, 179–191.
Hausser, Roland. 1974. *Quantification in an Extended Montague Grammar*: University of Texas at Austin Ph.D. dissertation.
Hayashishita, J.-R. 2009. *Yori*-comparative: Comments on Beck et al. (2004). *Journal of East Asian Linguistics* 18(2). 65–100. doi:10.1007/s10831-009-9040-5.
Heim, Irene. 1983. On the projection problem for presuppositions. In WCCFL 2, 114–125.
Heim, Irene. 1985. Notes on comparatives and related matters. Ms., University of Texas, Austin.
Heim, Irene. 2000. Degree operators and scope. In *Proceedings of SALT* 10, 40–64. Ithaca, NY: CLC Publications.
Heim, Irene. 2006. Remarks on comparative clauses as generalized quantifiers. Ms., Massachusetts Institute of Technology.
Heim, Irene. 2008. Features on bound pronouns. In Daniel Harbour, David Adger & Susana Bejar (eds.), *Phi Theory: Phi-Features Across Modules and Interfaces*, 35–56. Oxford: Oxford University Press.
Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
Hoeksema, Jack. 1983. Negative polarity and the comparative. *Natural Language and Linguistic Theory* 1. 403–434.
Hofstetter, Stefan. 2009. Comparison constructions in Turkish. In *Proceedings of Sinn und Bedeutung* 13, 187–201.
Hyde, Dominic. 2014. Sorites Paradox. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2014 edn.
Inagaki, Shunji & David Barner. 2009. Countability in absence of count syntax: Evidence from Japanese Quantity Judgments. In Shunji Inagaki, Makiko Hirakawa, Setsuko Arita, Yahiro Hirakawa, Hiromi Morikawa, Mineharu Nakayama, Hidetosi Sirai & Jessika Tsubakita (eds.), *Studies in Language Sciences 8: Papers from the 8th Annual Conference of the Japanese Society of Language Sciences*, 111–125. Tokyo: Kurosio.
Ivlieva, Natalia. 2013. *Scalar Implicatures and the Grammar of Plurality and Disjunction*: Massachusetts Institute of Technology dissertation.
Jackendoff, Ray. 1977. *X' Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kamp, Hans. 1975. Two theories of adjectives. In Edward Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press.
Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen A. G. Groenendijk, Theo M. V. Janssen & Martin J. B. Stokhof (eds.), *Formal Methods in the Study of Language*, Amsterdam: Mathematical Center.
Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyu Oh & David Dinneen (eds.), *Syntax and Semantics 11: Presupposition*, 1–56. New York: Academic Press.
Kennedy, Christopher. 1999. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. New York: Garland Press.
Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24(1). 33–70. doi:10.1023/A:1005668525906.
Kennedy, Christopher. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45. doi:10.1007/s10988-006-9008-0.
Kennedy, Christopher. 2009. Modes of comparison. In *Proceedings of Chicago Linguistic Society* 43, .
Kennedy, Christopher. 2010. Vagueness and comparison. In Paul Égré & Nathan Klinedinst (eds.), *Vagueness and Language Use*, 73–97. Hampshire: Palgrave McMillan.
Kennedy, Christopher. 2011. Ambiguity and vagueness. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Meaning*, vol. 1, 507–535. Berlin: de Gruyter.
Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381.
Kennedy, Christopher & Jason Merchant. 2000. Attributive Comparative Deletion. *Natural Language and Linguistic Theory* 18(1). 89–146. doi:10.1023/A:1006362716348.
Khanjian, Hrayr. 2008. Classifiers and number marking. In *Proceedings of SALT* 18, 73–89.
Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4(1). 1–45.
Klein, Ewan. 1982. The interpretation of adjectival comparatives. *The Journal of Linguistics* 18. 113–136.
Klein, Ewan. 1991. Comparatives. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 673–691. Walter de Gruyter.
Kneale, William & Martha Kneale. 1962. *The Development of Logic*. Oxford: Clarendon Press.
Krasikova, Sveta. 2008. Quantifiers in comparatives. In *Proceedings of SuB12*, 337–352.
Kratzer, Angelika. 2013. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event Structure in Linguistic Form and Interpretation*, 269–300. Berlin: De Gruyter.
Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Johan van Benthem, Peter van Embde Boas & Renate Bartsch (eds.), *Semantics and Contextual Expression*, 75–115. Dordrecht: De Gruyter.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan Sag & Anna Szabolcsi (eds.), *Lexical Matters*, 29–53. Stanford: CSLI.

Krifka, Manfred. 1998. Scope inversion under risek-fall contour in German. *Linguistic Inquiry* 29(1). 75–112.

Krifka, Manfred. 1999. At least some determiners aren't determiners. In Ken Turner (ed.), *The Semantics/Pragmatics Interface from Different Points of View*, 257–291. Elsevier Science.

Krifka, Manfred. 2002. Be Brief and Vague! And how Bidirectional Optimality Theory allows for verbosity and precision. In David Restle & Dietmar Zaefferer (eds.), *Sounds and Systems*, 439–458. Berlin: Mouton de Gruyter.

Krifka, Manfred. 2007. Approximate interpretation of number words. In Joost Zwarts, Gerlof Bouma & Irene Krämer (eds.), *Cognitive Foundations of Interpretation*, 111–126. Amsterdam: Mouton de Gruyter.

Ladusaw, William. 1982. Semantic constraints on the English partitive constructions. In *Proceedings of WCCFL 1*, 231–291.

Lakoff, George. 1970. A note on vagueness and ambiguity. *Linguistic Inquiry* 1(3). 357–359.

Landman, Fred. 1989a. Groups, I. *Linguistics and Philosophy* 12(5). 559–605.

Landman, Fred. 1989b. Groups, II. *Linguistics and Philosophy* 12(6). 723–744.

Landman, Fred. 1996. Plurality. In Shalom Lappin (ed.), *Handbook of Contemporary Semantic Theory*, 425–457. Oxford: Blackwell.

Landman, Fred. 2000. *Events and Plurality*. Dordrecht: Kluwer.

Larson, Richard. 1988. Scope and comparatives. *Linguistics and Philosophy* 11(1). 1–26. doi:10.1007/BF00635755.

Lasersohn, Peter. 1995. *Plurality, Conjunction and Events*. Dordrecht: Kluwer.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75(3). 522–551.

Lasersohn, Peter. 2011. Mass nouns and plurals. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Meaning*, vol. 2, 1131–1153. de Gruyter.

Lechner, Winfried. 2001. Reduced and phrasal comparatives. *Natural Language and Linguistic Theory* 19(4). 683–735. doi:10.1023/A:1013378908052.

Lechner, Winfried. 2004. *Ellipsis in Comparatives*. Berlin: Mouton de Gruyter.

Lechner, Winfried. 2008. On binding scope and ellipsis scope. In Kyle Johnson (ed.), *Topics in Ellipsis*, 154–182. Cambridge: Cambridge University Press.

Li, Peggy, Yarrow Dunham & Susan Carey. 2009. Of substance: The nature of language effects on entity construal. *Cognitive Psychology* 58(4). 487–524. doi:10.1016/j.cogpsych.2008.12.001.

Lima, Suzi. 2014. *The Gammar Of Individuation and Counting*: University of Massachusetts, Amherst dissertation.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, Use, and the Interpretation of Language*, 302–323. Berlin: Mouton de Gruyter.

Link, Godehard. 1987. Generealized quantifiers and plurals. In Peter Gardënfors (ed.), *Generalized Quantifiers: Linguistic and Logical Approaches*, Dordrecht: Kluwer.

Lucy, John. 1992. *Language Diversity and Thought: A Reformulation of the Linguistic*

Relativity Hypothesis. Cambridge: Cambridge University Press.
Mari, Alda, Claire Beyssade & Fabio Del Prete (eds.). 2012. *Genericity*. Oxford: Oxford University Press.
Marty, Paul, Emmanuel Chemla & Benjamin Spector. 2015. Phantom readings: the case of modified numerals. *Language, Cognition and Neuroscience* 30(4). 462–477. doi:0.1080/23273798.2014.931592.
Massam, Diane (ed.). 2012. *Count and Mass Across Languages*. Oxford: Oxford University Press.
Matthieu, Éric. 2012. Flavors of division. *Linguistic Inquiry* 43(4). 650–679. doi: 10.1162/ling_a_00110.
Matushansky, Ora. 2002. Tipping the scales: The syntax of scalarity in the complement of *seem*. *Syntax* 5(3). 219–276. doi:10.1111/1467-9612.00052.
McCawley, James D. 1975. Lexicography and the count-mass distinction. In *Proceedings of Berkeley Linguistics Society* 5, 314–321.
McNally, Louise. 2011. The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In Rick Nouwen, Robert van Rooij, Uli Sauerland & Hans-Christian Schmitz (eds.), *Vagueness in Communication*, 151–168. Springer.
Merchant, Jason. 2009. Phrasal and clausal comparatives in Greek and the abstractness of syntax. *Journal of Greek Linguistics* 9(1). 134–164. doi:10.1163/156658409X12500896406005.
Moltmann, Firederike. 2007. Events, tropes, and truthmaking. *Philosophical Studies* 134(3). 363–403. doi:10.1007/s11098-005-0898-4.
Moltmann, Friederike. 2009. Degree structure as trope structure: a trope-based analysis of positive and comparative adjectives. *Linguistics and Philosophy* 32(1). 51–94. doi:10.1007/s10988-009-9054-5.
Morzycki, Marcin. 2011. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics* 19(1). 39–86. doi:10.1007/s11050-010-9063-5.
Morzycki, Marcin. 2014. Modification. Ms., Michigan State University.
Nicolas, David. 2005. Mass nouns and plural logic. *Linguistics and Philosophy* 31(2). 211–244. doi:10.1007/s10988-008-9033-2.
Nouwen, Rick. 2008. Upper-bounded *no more*: the exhaustive interpretation of non-strict comparison. *Natural Language Semantics* 16(4). 271–295. doi:10.1007/s11050-008-9034-2.
Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics & Pragmatics* 3.
Nouwen, Rick. 2011. The Semantics of Degree. Lecture notes, 24.979 Topics in Semantics, Fall 2011, MIT.
Nouwen, Rick. to appear. Plurality. In Paul Dekker & Maria Aloni (eds.), *Cambridge Handbook of Semantics*, Cambridge University Press.
Oda, Toshiko. 2008. *Degree Constructions in Japanese*: University of Connecticut, Storrs Ph.D. dissertation.
Ojeda, Almerindo E. 2005. The paradox of mass plurals. In Salikoko S. Mufwene, Elaine J. Francis & Rebecca S. Wheeler (eds.), *Polymorphous Linguistics: Jim McCawley's Legacy*, 389–410. Cambridge, MA: MIT Press.
Pancheva, Roumyana. 2006. Phrasal and clausal comparatives in Slavic. In *Proceedings of FASL 14*, 236–257.
Pancheva, Roumyana. 2007. Than. Paper presented at GLOW XXX, the University

of Tromsø.
Partee, Barbara & Vladimir Borschev. 2012. Sortal, relational, and functional Interpretations of nouns and Russian container constructions. *Journal of Semantics* 29(4). 445–486. doi:10.1093/jos/ffs009.
Pearson, Hazel. 2011. A new semantics for group nouns. In *Proceedings of WCCFL 28*, 160–168.
Pearson, Hazel, Manizeh Khan & Jesse Snedeker. 2010. Even more evidence for the emptiness of plurality: An experimental investigation of plural interpretation as a species of implicature. In *Proceedings of SALT 20*, 489–508.
Pelletier, Francis Jeffry. 2012. Lexical nouns are both +mass and +count, but they are neither +mass nor +count. In Diane Massam (ed.), *Count and Mass across Languages*, 9–26. Oxford: Oxford University Press.
Peters, Stanley. 1979. A truth-conditional formulation of Karttunen's account of presupposition. *Synthese* 40(2). 301–316.
Pinkal, Manfred. 1990. On the logical structure of comparatives. In R. Studer (ed.), *Natural Language and Logic: Lecture Notes in Artificial Intelligence*, 147–167. Berlin: Springer.
Pinkal, Manfred. 1995. *Logic and the Lexicon: The Semantics of the Indefinite*. Dordrecht: Kluwer.
Quine, Willard Van Orman. 1960. *Word and Object*. Cambridge, MA: MIT Press.
Raffman, Diana. 1994. Vagueness without paradox. *The Philosophical Review* 103(1). 41–74.
Raffman, Diana. 1996. Vagueness and context relativity. *Philosophical Studies* 81(2–3). 175–192. doi:10.1007/BF00372781.
Rett, Jessica. 2014. *The Semantics of Evaluativity*. Oxford: Oxford University Press.
Roberts, Craige. 1987. *Modal Subordination, Anaphora and Distributivity*: University of Massachusetts, Amherst Ph.D. dissertation.
van Rooij, Robert. 2008. Comparatives and quantifiers. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical Issues in Formal Syntax and Semantics* 7, 423–444.
van Rooij, Robert. 2011. Vagueness and Linguistics. In Giuseppina Ronzitti (ed.), *Vagueness: A Guide*, 123–170. Springer. doi:10.1007/978-94-007-0375-9_6.
Rothstein, Susan. 2009. Individuaating and measure readings of classifier constructions: Evidence from Modern Hebrew. *Brill's Annual of Afroasiatic Languages and Linguistics* 1(1). 106–145. doi:10.1163/187666309X12491131130783.
Rothstein, Susan. 2010. Counting and the mass-count distinction. *Journal of Semantics* 27(3). 343–397. doi:10.1093/jos/ffq007.
Rotstein, Carmen & Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12(3). 259–288. doi:10.1023/B:NALS.0000034517.56898.9a.
Rullmann, Hotze. 1995. *Maximality in the Semantics of Wh-Constructions*: University of Massachusetts, Amherst Ph.D. dissertation.
Sassoon, Galit. 2010. The degree functions of negative adjectives. *Natural Language Semantics* 18(2). 141–181. doi:10.1007/s11050-009-9052-8.
Sassoon, Galit. 2012. A *slightly* modified economy principle: Stable properties have non-stable standards. In *Proceedings of IATL 27*, Cambridge, MA: MITWPL.
Sassoon, Galit. 2013. A typology of multidimensional adjectives. *Journal of Semantics* 30(3). 335–380. doi:10.1093/jos/ffs012.

Sassoon, Galit & Natalia Zevakhina. 2012. Granularity shifting: Experimental evidence from degree modifiers. In *Proceedings of SALT 22*, 226–246.
Sauerland, Uli. 2003. A new semantics for number. In Robert B. Young & Yuping Zhou (eds.), *Proceedings of SALT 13*, 258–275. Ithaca, NY: Cornell Linguistics Club.
Sauerland, Uli. 2008. On the semantics markedness of phi-features. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi Theory: Phi-Features across Modules and Interfaces*, 57–82. Oxford: Oxford University Press.
Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reise (eds.), *Linguistic Evidence*, 409–430. Berlin: Mouton de Gruyter.
Sauerland, Uli & Penka Stateva. 2007. Scalar vs. epistemic vagueness: Evidence from approximators. In Tova Friedman & Masayuki Gibson (eds.), *Proceedings of SALT 17*, 228–245. Ithaca, NY: Cornell Linguistics Club.
Sauerland, Uli & Penka Stateva. 2011. Two types of vagueness. In Paul Égré & Nathan Klinedinst (eds.), *Vagueness and Language Use*, 121–145. Hampshire: Palgrave Macmillan.
Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen A. G. Groenendijk, Theo M. V. Janssen & Martin J. B. Stokhof (eds.), *Formal Methods in the Study of Language*, 483–512. Dordrecht: Foris.
Scha, Remko & Yoad Winter. to appear. Plurals. In Shalom Lappin & Chris Fox (eds.), *The Handbook of Contemporary Semantic Theory*, Wiley-Blackwell 2nd edn.
Schlenker, Philippe. 2008. *Be Articulate!* A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34(3). 157–212.
Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2. 1–78.
Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer.
Schwarzschild, Roger. 2005. Measure phrases as modifiers of adjectives. In *L'adjectif*, vol. 34 Recherches Linguistiques de Vincennes, 207–228. Paris: Presses universitaires de Vincennes.
Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9(1). 67–110. doi:10.1111/j.1467-9612.2006.00083.x.
Schwarzschild, Roger. 2008. The semantics of comparatives and other degree constructions. *Language and Linguistics Compass* 2(2). 308–331. doi:10.1111/j.1749-818X.2007.00049.x.
Schwarzschild, Roger. 2009. Stubborn distributivity, multiparticipant nouns and the count/mass distinction. In *proceedings of* NELS 39, 1–18.
Schwarzschild, Roger. 2014. Markers of mass and count. Talk given at NELS 45.
Schwarzschild, Roger & Karina Wilkinson. 2002. Quantifiers in comparatives: A semantics of degree based on intervals. *Natural Language Semantics* 10(1). 1–41.
Scontras, Gregory. 2014. *The Semantics of Measurement*: Harvard University dissertation.
Selkirk, Elizabeth. 1977. Some remarks on noun phrase structure. In Peter Culicover, Tom Wasow & Adrian Akmajian (eds.), *Formal Syntax*, 285–316. New York: Academic Press.
Seuren, Pieter A. M. 1973. The comparative. In Ferenc Kiefer & Nicolas Ruwet (eds.), *Generative Grammar in Europe*, 528–564. Dordrecht: Reidel.
Seuren, Pieter A. M. 1978. The structure and selection of positive and negative gradable adjectives. In *Proceedings of Chicago Linguistic Society* 14, 336–346.
Seuren, Pieter A. M. 1984. The comparative revisited. *Journal of Semantics* 3. 109–

141.
Seuren, Pieter A. M. 2005. Eubulides as a 20th-century semanticist. *Language Sciences* 27(1). 75–95. doi:10.1016/j.langsci.2003.12.001.
Shapiro, Stewart. 2006. *Vagueness in Context*. Oxford: Clarendon.
Shapiro, Stewart. 2011. Vagueness and Logic. In Giuseppina Ronzitti (ed.), *Vagueness: A Guide*, 55–81. Springer. doi:10.1007/978-94-007-0375-9_3.
Shimoyama, Junko. 2012. Reassessing crosslinguistic variation in clausal comparatives. *Natural Language Semantics* 20(1). 83–113. doi:10.1007/s11050-011-9076-8.
Solt, Stephanie. 2015. Vagueness and imprecision: Empirical foundations. *Annual Review of Linguistics* 1. 107–127. doi:10.1146/annurev-linguist-030514-125150.
Sorensen, Roy. 2013. Vagueness. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2013 edn.
Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presuppositions and Implicatures in Compositional Semantics*, 243–281. New York: Palgrave-Macmillan.
Spector, Benjamin. 2014. Plural indefinites and maximality. Talk at UCLA.
Stassen, Leon. 1984. The comparative compared. *Journal of Semantics* 3(1–2). 143–182. doi:10.1093/jos/3.1-2.143.
Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
Stassen, Leon. 2006. Comparative constructions. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 686–690. Amsterdam: Elsevier.
von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3(1–2). 1–77. doi:10.1093/jos/3.1-2.1.
Sudo, Yasutada. 2012. *On the Semantics of Phi Features on Pronouns*: Massachusetts Institute of Technology Ph.D. dissertation.
Sudo, Yasutada. 2014. Hidden nominal structures in Japanese clausal comparatives. *Journal of East Asian Linguistics* 24(1). 1–51. doi:10.1007/s10831-014-9125-7.
Syrett, Kristen. 2007. *Learning about the Structure of Scales: Adverbial Modification and the Acquisition of the Semantics of Gradable Adjectives*: Northwestern University dissertation.
Syrett, Kristen, Christopher Kennedy & Jeffrey Lidz. 2010. Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* 27(1). 1–35. doi:10.1093/jos/ffp011.
Takahashi, Shoichi. 2006. More than two quantifiers. *Natural Language Semantics* 14(1). 57–101. doi:10.1007/s11050-005-4534-9.
Toledo, Assaf & Galit Sassoon. 2011. Absolute vs. relative adjectives: Variance within vs. between individuals. In *Proceedings of SALT 21*, 135–154.
Tribushinina, Elena. 2013. Adjective semantics, world knowledge and visual context: comprehension of size terms by 2- to 7-year-old Dutch-speaking children. *Journal of Psycholinguistic Research* 42(3). 205–225. doi:10.1007/s10936-012-9217-3.
Varzi, Achille. 2015. Mereology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2015 edn.
de Vries, Hanna. 2012. Lexical distributivity with group nouns and property indefinites. In *Proceedings of NELS 43*, .
de Vries, Hanna. 2013. Distributivity and agreement: new evidence for groups as sets. In *Proceedings of the Amsterdam Colloquium 2013*, .

Williamson, Timothy. 1994. *Vagueness*. London: Routledge.
Williamson, Timothy. 1997. Imagination, stipulation and vagueness. *Philosophical Issues* 8. 215–228.
Winter, Yoad. 2000. Distributivity and dependency. *Natural Language Semantics* 8(1). 27–69. doi:10.1023/A:1008313715103.
Winter, Yoad. 2001a. *Flexibility Principles in Boolean Semantics: The Interpretation of Coordination, Plurality, and Scope in Natural Language*. Cambridge, MA: MIT Press.
Winter, Yoad. 2001b. Plural predication and the Strong Meaning Hypothesis. *Journal of Semantics* 18. 333–365. doi:10.1093/jos/18.4.333.
Winter, Yoad. 2002. Atoms and sets: A characterization of semantic number. *Linguistic Inquiry* 33(3). 493–505. doi:10.1162/002438902760168581.
Winter, Yoad. 2005. Cross-categorial restriction on measure phrase modification. *Linguistics and Philosophy* 28(2). 233–267.
Wunderlich, Dieter. 2001. Two comparatives. In István Kenesei & Robert M. Harnish (eds.), *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for Ferenc Kiefer*, 75–91. Amsterdam: John Benjamins.
Zucchi, Sandro & Michael White. 2001. Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24(2). 223–270. doi:10.1023/A:1005690022190.
Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32(4). 353–407. doi:10.1007/s10988-009-9064-3.