Induction and Comparison

PAUL M. PIETROSKI

Abstract

Frege proved an important result, concerning the relation of arithmetic to second-order logic, that bears on several issues in linguistics. Frege's Theorem illustrates the logic of relations like PRECEDES(x, y) and TALLER(x, y), while raising doubts about the idea that we *understand* sentences like 'Carl is taller than Al' in terms of abstracta like heights and numbers. Abstract paraphrase can be useful—as when we say that Carl's height exceeds Al's—without reflecting semantic structure. Related points apply to causal relations, and even grammatical relations like DOMINATES(x, y). Perhaps surprisingly, Frege provides the resources needed to recursively characterize labelled expressions *without* characterizing them as sets. His theorem may also bear on questions about the meaning and acquisition of number words.

Introduction

Suppose that Carl is just barely taller than Bob, who is just barely taller than Al. It follows that Carl is taller than Al, but *not* that Carl is just barely taller than Al. Yet if Carl is much heavier than Bob, who is heavier than Al, then Carl is much heavier than Al. And if Al is Bob's father, while Bob is Carl's father, then Al is Carl's *grand*father. Though if Al is Bob's brother, while Bob is Carl's, then Al and Carl are brothers. Speakers of English recognize such inferential relations. This raises questions about (i) how the relevant sentences are understood, and (ii) the kinds of inference that are germane. Perhaps 'Carl is taller than Al' means roughly that Carl's height exceeds Al's height, and speakers make relevant inferences about heights and exceeding.[1] But I'll urge a different perspective, based on Frege's (1879, 1884, 1892, 1893) reasons for inventing the modern logic—presupposed by most semanticists—that lets us deal with relations like SUCCESSOROF(x, y) and GREATERTHAN(x, y). This requires a

[1] Cresswell (1976) offers a semantics of this sort, though in a framework that makes the relation to psychology not entirely clear. Klein (1980) articulates and then argues against such hypotheses about natural language meaning. But while I agree with many of Klein's critical remarks, my use Frege is rather different than his. For more recent "degree semantics" approaches, see Kennedy (1999a, 1999b), Beck *et.al.* (2004), and references there. Schwartzschild (2002), who speaks in terms of "intervals," arguably has a different view. And it may be that the view I have in mind is more presupposed, in discussions of related topics, than explicitly defended in published work.

©2007 Paul Pietroski. *University of Maryland Working Papers in Linguistics 15,* ed. A. Conroy, C. Jing, C. Nakao and E. Takahashi, pp. 154-188. College Park, MD: UMWPiL.

review of "Frege's Theorem" about the foundations of arithmetic. But linguists should know about this theorem in any case.

It is well known that Frege tried and failed to reduce arithmetic to logic. It is less well known, outside a community of specialists, that Frege did establish a foundational result. Given his second-order logic, and suitable definitions, Frege reduced the axioms of arithmetic to a single principle: $\forall F \forall G[(\#F = \#G) \equiv \text{OneToOne}(F, G)]$; whatever the Fs and Gs may be, the *number of* the Fs is also the *number of* the Gs iff the Fs *correspond one-to-one with* the Gs. Or as Frege would have put it: for any Concepts F(x) and G(x), the number of things falling under the former is the number of things falling under the latter iff a one-to-one function associates the things falling under F(x) with those falling under G(x). Frege showed how to derive all of arithmetic from this generalization, now called "Hume's Principle." Frege's error was to supplement his otherwise consistent logical system with an allegedly logical axiom that let him derive Hume's Principle: as Russell famously noted, the resulting logic proved too much—e.g., that there was and was not a set of all the sets that do not contain themselves. But despite this failure, Frege did reduce arithmetic to a principle that links our notions of number and one-to-one correspondence, given a consistent fragment of the second-order logic he invented in order to formulate an interesting but false Logicist thesis.[2]

For present purposes, it is especially important that Frege derived the axiom of mathematical induction from a purely *logical* principle. In arithmetic proofs, one often relies on the following conditional premise whose antecedent is complex: *if* the first number (zero) has some property P, and every number is such that if it has P then so does its successor, *then* every number has P. (Here and throughout, 'number' should be understood as 'natural number', a predicate satisfied by all and only the nonnegative whole numbers.) Frege showed that this axiom of arithmetic was a *theorem* of his logic, given requisite definitions of 'number' and 'successor'. The relevant principle governs inferences involving any first-order relation like FATHEROF(x, y), and a corresponding "ancestral" relation like FOREFATHEROF(x, y); where the ancestral relation can be defined, given second-order quantification, in terms of the first-order relation. Frege was most interested in the relation

---

2 See Parsons (1965), Wright (1983), Boolos (1998), and the essays in Demopolous (1994). Given Aristotelian logic, even as developed by the medievals, Kant was surely right to conclude that (knowledge of) arithmetic was not reducible to (knowledge of) logic plus definitions. But Frege effectively reopened the question by changing the conception of logic, much as questions about the relation of chemistry to physics changed as conceptions of physics changed. See Heck (1999) for a helpful introduction to these issues.

PREDECESSOROF(x, y) and its "transitive closure" PRECEDES(x, y). But his reasoning applies equally well to relations like IMMEDIATELYDOMINATES(α, β) and DOMINATES(α, β), PROXIMALLYCAUSED(c, e) and CAUSED(c, e), MINIMALLYTALLER(x, y) and TALLER(x, y). Frege's Theorem may also bear on questions about how children understand numeric words like 'three' and 'seven'. But in any case, I think we should look for a *semantics* according to which 'taller' is understood in terms of second-order quantification over concrete individuals, and a dovetailing conception of natural *logic* according to which speakers appreciate inferential relations involving ancestrals.

It can be tempting instead to specify the meaning of 'Carl is taller than Al' as follows, abstracting away from various details: H(Carl) > H(Al); where 'H' indicates a function from individuals to numbers that correspond to ranks on an appropriate "height scale." But prima facie, humans can judge that one thing is taller/heavier/closer/funnier than another *without* being able to associate the things compared with abstracta ordered by a relation indicated with '>'. Likewise, it seems that a speaker can infer that Carl is taller than anyone who is shorter than Bob, yet be unable to judge that Carl's height *exceeds* the heights of those whose heights are exceeded by Bob's. Of course, these appearances may be deceiving. Perhaps we do understand sentences of the form '_ is taller than _' in a way that relies on a capacity to compare numbers, or other abstracta like degrees of height. But this is hardly obvious. Our ability to judge one number greater than another may be parasitic on the very capacities that let us understand comparative constructions; where in paradigmatic cases, involving nonarithmetic words like 'taller', the relevant first-order relations hold between perceptible entities like Carl and Al.

This matters, in part because we must eventually face the question of *what it is* for Carl's height to exceed Al's. And this is presumably not a matter of Carl's height-number having a greater number-number than Al's height-number. So if we say that speakers represent the world in terms of some heights exceeding others, we need to say what this amounts to, without falling back on the claim that some individuals are taller than others. Notation like 'H(Carl) > H(Al)' may let us describe truth conditions in ways that are fine for certain purposes. But if '>' means what it means in arithmetic, we cannot just assume that such notation provides a good account of how ordinary speakers *understand* expressions like 'taller'. Extant theories can, no doubt, be rewritten along lines suggested here. My aim is to motivate such rewriting, not to criticize particular theories, and certainly not to offer a complete theory of comparative constructions. But I will conclude by using a Fregean proposal to help explain the otherwise puzzling absence of monomorphemic predicates with relevant comparative meanings.

We can imagine a language in which ‘Carl owtites Al’ means that Carl is taller than Al. So why don’t we often see noncomposite lexical items with meanings like λy.λx.TALLER(x, y)? Drawing on Pietroski (2005, 2006), my suggestion will be that relational *concepts* are lexicalized as *monadic* predicates; where some predicates are satisfied by things like events, in which individuals “participate,” and some predicates are essentially *plural*—in the sense of being satisfied by some things without being satisfied by any one of them. (For example, ‘formed a circle’ might be satisfied by some people, no one of whom formed a circle.) The idea will be that Carl is taller than Al iff: there are some things ordered by the relation TALLER(x, y), with Carl as the external/“outermost” thing, and Al as the internal/“innermost” thing; in which case, Al and Carl exhibit an ancestral relation definable in terms of MINIMALLYTALLER(x, y).

1 Validity and Arithmetic

Let me ease into the details with some reminders about the role and importance of second-order quantification in Frege’s (1879) logic, from which much of contemporary semantics descends.

1.1 Quantifiers and Predicates

The first-order predicate calculus, which is a fragment of Frege’s logic, certainly has virtues. Inventing it, without also inventing a formal system that allowed for quantification into positions occupiable by predicates, would have been impressive enough—especially in combination with Frege’s idea that that propositions (Gedanken, “things” that exhibit logical relations) have *function-argument* structures. Still, endlessly many expressions of natural language are not firstorderizable. Boolos (1984) reviews examples involving ‘most’ and ‘only’, but also stresses sentences like ‘For every pleasure, there is a pain’, with the implication that there are as many pains as pleasures. We may have thoughts of the form ‘$\forall$x[Fx $\supset$ $\exists$y(Gy)]’, or using restricted quantifiers ‘$\forall$x:Fx($\exists$y:Gy)’. But we also have thoughts, expressed with words like ‘every’, that cannot be captured in first-order terms.[3] Recognizing this, at least implicitly, Frege invented a logic that validated conditionals like ‘Fa $\supset$ $\exists$x(Fx)’ *and* the second-order variant ‘Fa $\supset$ $\exists$X(Xa)’. For his purposes, biconditionals like ‘Fa

---

3

See Rescher (1962), Wiggins (1983). Imagine a company that advertises ‘For every dollar we receive, a penny will go to charity’, and then sends exactly one penny to charity after receiving a million dollars—saying that this satisfies the only plausible first-order regimentation of the English sentence. This seems indefensible, *pace* Quine (1950). Here and throughout, I use modern *notation*, instead of Frege’s. But the logic is his.

$\equiv \exists X \exists x[Xx \,\&\, (x = a) \,\&\, \forall y(Xy \equiv Fy)]$' were crucial.

Frege also offered a distinctive interpretation of predicative variables. This interpretation is not required, nor is it without difficulties; see §2 below. But for now, let's follow Frege and say that capitalized variables range over Concepts: functions, from entities to truth or falsity, that reflect a kind of abstraction—starting with a thought "about" a particular thing, and abstracting away from the thing thought about, leaving an "unsaturated" thought-component. Consider, for example, the thought we express with 'Euclid was clever'. Using '**u**' as a name for the geometer in question, we can represent this thought as follows: **Clever**(**u**). Ignoring the specific contribution of the name leaves a first-order Concept, **Clever**( )$_x$, that maps each entity x to truth iff x is clever. Ignoring the specific contribution of the Concept-expression leaves a second-order Concept, X(**u**), that maps each Concept to truth iff that Concept maps Euclid to truth.

Letting '**C**y' abbreviate '**Clever**( )$_y$', '$\exists X \exists x[Xx \,\&\, (x = \mathbf{u}) \,\&\, \forall y(Xy \equiv \mathbf{C}y)]$' means that there is some Concept X and some entity x such that: X maps x to truth; and x is identical with **u**; and for each entity y, X maps y to truth iff **C** maps y to truth.[4] From this perspective, quantification into positions occupiable by predicates is not essentially different than quantification into positions occupiable by names. Starting with a thought analyzable as a Concept **C** saturated by an entity **u**, we can abstract a "thought-frame" corresponding to the open sentence '**C**y', or a thought-frame corresponding to the open sentence 'X**u**'. And as Frege noted, variables can be encoded in ways that make such abstraction more explicit, at the cost of typographic convenience. We can, if we like, depict the thought that Euclid is clever as follows.

```
       x__________              y___________
       |    |      |            |    |      |
∃_∃_{ _ ( _ ) & [( _ ) = u] & ∀_ [ _( _ ) ≡ C( _ )]}
   |____|______________________|
   X
```

The novelty and advantage of such analysis was intertwined with Frege's

4 Using '≡' as the biconditional highlights, as Frege did, parallels between material equivalence and identity: if $\forall x(Hx \equiv Px)$, then in Frege's logic, 'P' can be substituted for 'H' *salva veritate*; likewise, if h = p, 'p' can be substituted for 'h'. As we'll see, Frege construed talk of numbers in terms of a third kind of equivalence: if the things that "fall under" one Concept correspond 1-1 with the things that fall under another Concept, the Concepts are numerically equivalent. Or put another way, two Concepts "have the same number" if their extensions are equinumerous.

associated conception of logical structure, and its potential divergence from grammatical structure. While medieval logicians had greatly improved Aristotelian logic, extending its scope and reducing the number of primitive inferential patterns, well-known problems beset the underlying idea that propositions have subject-copula-predicate structure. Sentences with transitive verbs, quantificational direct objects, and relative clauses—as in (1) and (2)—

(1) Every politician deceived some voter who trusted him
(2) Every politician who trusted himself deceived some voter

indicate thoughts whose logical relations cannot be captured in these traditional terms. One can say that the corresponding propositions are both of the form [(Every Φ) is Ψ], with the monadic predicates being: 'Politician' and 'DeceivedSomeVoterWhoTrustedHim' in (1); 'PoliticianWhoTrustedHimself' and 'DeceivedSomeVoter' in (2). But without a systematic way of reflecting logical structure *within* the complex monadic predicates, endlessly many implications go unexplained.[5] By contrast, Frege could represent the corresponding propositions as having (agrammatical) constituency structures like those shown in (1a) and (2a).

(1a) $\forall x[Px \supset \exists y(Vy \;\&\; Tyx \;\&\; Dxy)]$
(2a) $\forall x[(Px \;\&\; Txx) \supset \exists y(Vy \;\&\; Dxy)]$

Though given his interests in the foundations of arithmetic, Frege cared more about (3) and (3a).

(3) Every number has a successor
(3a) $\forall x[Nx \supset \exists y(Syx)]$

The idea was that a natural language sentence like (3) could be used to express a potential premise/conclusion more perspicuously represented with (3a), which reflects the logically significant structure. But by themselves, first-order "logical forms" would have been inadequate for Frege's purposes. He wanted to represent *all* the axioms of arithmetic in a way that made it possible to offer rigorous proofs of interesting theorems, while also revealing the axioms themselves as *consequences of*

---

5 See Pietroski (2003) for an overview and references; see Ludlow (2002) for illuminating discussion and potential connections with a minimalist syntax.

more basic principles. As noted above, while Frege didn't reduce arithmetic to logic alone, he did represent the axioms of arithmetic in a way that let him derive these axioms from logic and a single basic principle. But what is most important, for our purposes, is Frege's subproof that the "axiom" of mathematical induction is really a special case of a (second-order) logical theorem that is not essentially concerned with numbers. This is what bears most directly on semantic hypotheses about words like 'taller' and 'father'. But it will also be useful to have, in the background, a sketch of Frege's larger result.

Whatever the relative merits of (3a), compared with earlier analyses, it does not yet capture the idea that each number has its own "unique" successor—or that there is at least one number, or that there is a "first" number, or that there is no last number. We can add first-order representations of three more Dedekind-Peano axioms, as shown below.

(DP:i)	Zero is a number
(DP:I)	N0
(DP:ii)	Zero is not the successor of any number
(DP:II)	$\forall x[Nx \supset \neg S0x]$
(DP:iii)	No two numbers have the same successor
(DP:III)	$\forall x \forall y[(Nx \ \& \ Ny \ \& \ x \neq y) \supset \neg \exists z(Szx \ \& \ Szy)]$

But for Frege, such representations frame the real questions. Are these independent axioms, which seem obvious only because we somehow intuit these fundamental arithmetic truths? Or do they follow from "deeper" truths? Can we define 'zero', 'number', and 'successor'—'0', 'Nx', and 'Szy'—in a way that reveals arithmetic axioms as theorems of a more general theory? How should we formalize the axiom of induction, stated roughly as the conditional claim (DP:iv)?

(DP:iv) *if* zero has a "property" such that
whenever a number has that property, its successor has that property
then every number has that property

And how are these four axioms related to (3) and (3a), renumbered below?

(DP:v)	Every number has a successor
(DP:V)	$\forall x[Nx \supset \exists y(Syx)]$

In particular, can (DP:iv) be formulated in a way that does *not* presuppose (DP:V)?

Frege realized that he could state the axiom of mathematical induction as (DP:IV)

(DP:IV) $\forall X[X0 \ \& \ \forall x \forall y(Xx \ \& \ Syx \supset Xy) \supset \forall x(Nx \supset Xx)]$

and still capture its special role in proofs of arithmetic theorems. Given Frege's interpretation of second-order variables, (DP:IV) says for every Concept_X: *if* zero falls under it_X, and each thing_x that falls under it_X is such that each successor_y (if any) of that_x thing falls under it_X, *then* every number falls under it_X; where subscripts track intended referential dependence of pronouns as used. Frege was also able to define 'zero', 'number', and 'successor' in terms of 'predecessor'. And this let him reveal (DP:IV) as a theorem of *logic*, usable in a proof of (DP:V). So at least in this sense, the Dedekind-Peano axioms are not five independent and fundamental truths. It is worth being clear about all this, and the role of second-order quantification in Frege's ingenius definitions, since the defined arithmetic terms exhibit logical relations like those exhibited by 'taller' and 'father'. Indeed, Frege took the ordinary language terms as models, and tried to generalize in a way that would cover the arithmetic cases.

1.2 A Sketch of Frege's Theorem

Frege assumed that we can use words like 'zero', 'one', and 'two' as entity-designators as well as second-order predicates. There is a sense in which he takes the latter use—as in 'There are two apples' and 'The apples are two', with the logical form '$\exists X[\text{Two}(X) \ \& \ \forall x[X(x) \equiv \text{Apple}(x)]$'— as the basic use. But this doesn't tell us what the first-order variables in generalizations like (4) and (5) range over.

(4) For every prime, there is a greater one

(4a) $\forall x\{\text{Prime}(x) \supset \exists y[\text{Prime}(y) \ \& >(y, x)]\}$

(5) There are infinitely many primes

(5a) $\exists X\{\text{InfinitelyMany}(X) \ \& \ \forall x[X(x) \equiv \text{Prime}(x)]\}$

And of course, Frege wanted to make sense of formulae like 'Successor(1) = 2', with '1' and '2' interpreted as entity-designators.

With this in mind, he defined the arithmetic *entity* zero as *the number of* nonselfidentical things. In terms of Concepts, zero is the number of things "falling under" (mapped to truth by) the Concept expressed with '$x \neq x$'. Given this starting point, other numbers can be defined recursively: one is the number of things identical with zero; two is the number of things identical with zero or one; etc. If only for

convenience, let ‘#’ indicate a higher-order Concept that maps each first-order monadic Concept, like Prime(x) or Frog(x), to the number that is the number of things falling under the Concept. Then we can represent Frege’s idea as follows: $0 = \#(x \neq x)$; $1 = \#[(x = 0)]$; $2 = \#[(x = 0) \vee (x = 1)]$; $3 = \#[(x = 0) \vee (x = 1) \vee (x = 2)]$; etc.[6] This isn’t yet a general definition of ‘number’. But for Frege, the key arithmetic notion is that of precedence; and given this notion, along with zero, a general definition is available. The numbers are said to be zero and *the things that zero precedes*. So given a defintion of ‘precedes’, each number can be identified as the number of numbers preceding it. More formally:
$\forall x\{\text{Number}(x) \equiv [(x = 0) \vee \text{PRECEDES}(0, x)]\} \;\&\; \forall y\{\text{Number}(y) \supset y = \#[\text{PRECEDES}(x, y)]\}$.

It is important to distinguish ‘PRECEDES(x, y)’ from ‘PREDECESSOROF(x, y)’. The former implies that x is less than y, but not that $y = x + 1$. Zero precedes endlessly many numbers, but is the predecessor of exactly one. Of course, the notions are related: x is a/the predecessor of y iff x *immediately* precedes y. But the trick was to define ‘PRECEDES(x, y)’ in terms of ‘PREDECESSOROF(x, y)’, and offer an independent definition of the latter.

Here too, Frege employed second-order quantification, revealing the power of his logic.

$\text{PREDECESSOROF}(x, y)$ iff $\exists X \exists z\{Xz \;\&\; (y = \#X) \;\&\; [x = \#(w \neq z \;\&\; Xw)]\}$
That is, x is a predecessor of y iff there is a Concept X and an entity z such that: z falls under X; y is the number of X—i.e., the number of things that fall under X; and x is the number of things apart from z that fall under X. For example, 2 is a predecessor of 3 iff there is some Concept_X under which three things fall, and some entity_z falling under it_X, such that (exactly) two *other* things fall under it_X. There is sure to be such a Concept and entity, given the following Concept: $(x = 0) \vee (x = 1) \vee (x = 2)$. Once defined, predicates like ‘PREDECESSOROF(x, y)’ and ‘PARENTOF(x, y)’ can be used to express relations such that: one entity bears the relation to another entity, which bears the relation to a third entity; the first does not bear the relation to the third; yet the first entity does, as a matter of logic, bear a corresponding “ancestral” relation to both the second entity and the third. Each of my ancestors bears the relation ANCESTOROF(x, y) to me. And intuitively, my ancestors are those individuals such that: each of my parents is one of them; for every one of them, each of his or her parents is also one of them; and nobody else is one of them. (This leaves

6 There are complications here, due to Frege’s (1892) insistence that Concepts are not objects. But given the lambda-calculus, we can say: $0 = \#(\lambda x.x \neq x)$, $1 = \#[\lambda x.(x \neq x) \vee (x = 0)]$, *etc*. See Zalta (2003) for an elegant and accessible presentation of the details.

room for the possibility that one of my ancestors was parentless.[7]) Likewise, the numbers that precede seven are those numbers such that: the predecessor of seven is one of them; for each of them that has a predecessor, its predecessor is one of them; and nothing else is one of them. (This leaves room for zero, which has no predecessor.) And given Frege's logic, one can define an "ancestralizing" function that maps relations like PREDECESSOROF(x, y) to "transitive closure" relations like PRECEDES(x, y).

In the biological case, x is an ancestor of y iff x had at least one child, and every Concept X satisfies the following inductive conditional: *if* each child of x falls under X, and a child c falls under X whenever a parent of c falls under X, *then* y falls under X. Informally, the idea is that x is an ancestor of y iff y has every "hereditary property" that x "passed on;" where a given individual, Eve, may have passed on properties like *being human*—or more to the point here, *being descended from or identical with Eve*. Generalizing with numbers in mind, let's say that for each relation R: each Concept X is "R-hereditary" iff $\forall w \forall z\{Xw \,\&\, R(w, z) \supset Xz\}$; and x is an "R-ancestor" of y iff x bears R to something, and y falls under every R-hereditary Concept under which x falls.[8] The relation $[\mathbf{ANC}(R)](x, y)$ can thus be defined as follows: $\forall x \forall y\{[\mathbf{ANC}(R)](x, y) \equiv_{df} \exists z[R(x, z)] \,\&\, \forall X[Xx \,\&\, \forall w \forall z\{Xw \,\&\, R(w, z) \supset Xz\} \supset Xy]\}$.

That is, x is an R-ancestor of y iff x bears R to something, and for every Concept X: *if* x falls under X, and an entity_z falls under X whenever something_w that falls under X is R-related to that_z entity, *then* y falls under X. Put another way (see §2.1), x is an R-ancestor of y iff x bears R to something, and whatever the Xs may be:

7

Readers who cannot describe parents (or grandparents) as ancestors can substitute suitably restricted definitions for their idiolects.

8

See Zalta (2003). We could also replace the condition that x bears R to something, and that x falls under X, with the condition that *anything* to which x bears R is something that falls under X. This leaves room for the possibility that x doesn't fall under X, even if falling under X is always "passed on," say from father to son. (Imagine a past Adam, and the Concept *being nonhuman or having had a human father*. Adam would not fall under this Concept, even though each of his male descendants would.) We could also say that x is a forefather of y iff y had a father, and every Concept is such that: *if* each father of y falls under the Concept, and anybody who falls under the Concept had a father who falls under the Concept, *then* x falls under the Concept. This leaves room for the possibility that y doesn't fall under the Concept, even if falling under the Concept is always "inherited" from fathers. (Imagine a future Joseph with exactly one son, who is nonhuman, and the Concept *being nonhuman or having had a human son*. While each of Joseph's forefathers would fall under this Concept, Joseph might not.)

y is one of them *if* x is one of them, and an entity_z is one of them whenever one of them is R-related to it_z. Initially, this sounds complicated. But upon reflection, it comes to seem as truistic as it is.

Frege treats ‘PRECEDES(x, y)’ as an abbreviation of ‘[**ANC**(PREDECESSOROF)](x, y)’. So by definition, PRECEDES(x, y) iff: x is a number;[9] and y falls under each Concept X such that (i) whatever x is a predecessor of falls under X, and (ii) whenever something falls under X, whatever it is a predecessor of falls under X. We can thus define ‘<’ as desired: <(x, y) iff PRECEDES(x, y). Given a relation R, we can define ‘$\underline{R}$’ so that: $\underline{R}(x, y) \equiv R(x, y) \lor (x = y)$. In particular, $\underline{\text{PRECEDES}}$(x, y) iff PRECEDES(x, y) or (x = y); hence, $\underline{\text{PRECEDES}}$(x, y) iff $x \leq y$. And for each number z, $\underline{\text{PRECEDES}}$(0, z). So the numbers just *are* zero and the things zero precedes.

Frege’s definitions thus make it explicit that the numbers are things that support mathematical induction. So the axiom of induction

(DP:IV) $\quad \forall X[X0 \;\&\; \forall x\forall y(Xx \;\&\; Syx \supset Xy) \supset \forall x(Nx \supset Xx)]$

need not be viewed as a special arithmetic law concerning numbers. On the contrary, it can be viewed as a special case of a more general *logical* truism—concerning any things that form a series ordered by a relation, and Concepts that are hereditary with regard to that relation. The “Logical Induction” principle (6), once understood, sounds obvious and is.

(6) If an entity falls under a Concept that is hereditary with regard to a relation,
and the entity is the initial thing in a series of things so related,
then each thing in the series of things so related falls under that Concept.

Frege formalizes this principle by first defining a technical notion. Given a relation R, a Concept X is *hereditary on the R-series starting with entity e* iff the following condition obtains: $\forall x\forall y\{[\underline{\mathbf{ANC}(R)}](e, x) \;\&\; [\underline{\mathbf{ANC}(R)}](e, y) \;\&\; R(x, y) \supset (Xx \supset Xy)\}$. Given the relation PREDECESSOROF(x, y), ‘$[\underline{\mathbf{ANC}(R)}](e, x)$’ says that e is less than or equal to x, and the whole conditional says: *if* e is less than or equal to both x and y, and PREDECESSOROF(x, y)—that is, *if* x is the predecessor of y, and e precedes

---

9 By definition, $\forall y[\text{Number}(y) \equiv \text{Precedes}(0, y) \lor (y = 0)]$. And by definition, Precedes(0, y) iff: $\exists z[\text{PredecessorOf}(0, z)] \;\&\; \forall X\{\forall z[\text{PredecessorOf}(0, z) \supset Xz] \;\&$ $\forall w\forall z[Xw \;\&\; \text{PredecessorOf}(w, z) \supset Xz] \supset Xy\}$. So given that 0 is the predecessor of 1, y is a number greater than zero iff: $\forall X\{X1 \;\&\; \forall w\forall z[Xw \;\&\; \text{PredecessorOf}(w, z) \supset Xz] \supset Xy\}$.

y—*then* y falls under X if x does. If e is the number zero, e must precede y, given that y has a predecessor. So in the special case of the relation that matters for proofs by mathematical induction, with zero as the base case, a Concept is hereditary on the relevant R-series iff that Concept is hereditary on the numbers. Correspondingly, (6) can be formalized in a way that leaves it general enough to be provable as a theorem of logic.

(6a) Given a relation R(x, y), a Concept X, and an entity e:
if X is hereditary on the R-series starting with e, and e falls under X,
then $\forall x[\underline{R}(e, x) \supset Xx]$.

And given Frege's defintions, (DP:IV) is a special case of this theorem: *if* X is hereditary on the numbers—i.e., the PREDECESSOROF-series, starting with zero—and zero falls under X, *then* any number (i.e., anything greater than or equal to zero) falls under X. This makes it clear that (6) is no more *about* numbers than it is about parents.

1.3 Digression

It may help, in this regard, to think about other cases of relations with "transitive closures" that we care about. For example, one event may be a "distal" cause of another: c may cause e *via* some intermediate event d, which causes e and is caused by c; and while c may also be a distal cause of d, there presumably are "proximal" causes with "immediate" effects. Given the relation PROXIMALCAUSE(d, d'), we can define the corresponding ancestral relation CAUSE(c, e).

We can also define a more constrained ancestral relation to capture a now familiar idea discussed by Hart and Honoré (1959): a PROXIMALCAUSE-series, starting with an *action* of some person, in which the chain of responsibility is not "broken" by the intentional action of another person—not even if this second action was somehow caused by the first. As an approximation, consider the following stipulations: PROXIMALCAUSE*(d, d') iff PROXIMALCAUSE(d, d') & ¬Action(d'); ACTCAUSE(a, e) iff Action(a) & [**ANC**(PROXIMALCAUSE*)](a, e). It follows that if ACTCAUSE(a, e), there is a causal chain from the action to the effect that does *not* go through a second action that causes the effect and is caused by the initial action. Some such proposal might be *part* of a reply to Fodor (1970), Fodor and Lepore (1998, 2002), in defense of certain theories of causative constructions; see Pietroski (1998, 2000, 2005) and references there.

In §2.3, I note that at least in principle, ancestral relations might be used to define labelled expressions recursively *without* identifying such expressions with sets.

And if we let ‘MT’ abbreviate ‘MINIMALLYTALLER’, the corresponding ancestral might be defined as follows: $\forall x\forall y\{\text{TALLER}(x, y) \equiv \exists z[\text{MT}(x, z)]\ \&\ \forall X[Xx\ \&\ \forall w\forall z\{Xw\ \&\ \text{MT}(w, z) \supset Xz\} \supset Xy]\}$. Perhaps x is taller than y iff: x is minimally taller than something—maybe x itself, minus a smidgeon; and for each Concept X, *if* x falls under X, and an entity_z falls under X whenever something_w that falls under X is minimally taller that_z entity, *then* y falls under X. If we can spell out ‘MINIMALLYTALLER’ in terms of smallest differences that are not contextually irrelevant, for purposes of ordering things in a certain way, ‘TALLER’ can be viewed as an ancestral predicate that indicates a transitive relation: $\forall x\forall y\forall z[\text{TALLER}(x, y)\ \&\ \text{TALLER}(y, z) \supset \text{TALLER}(x, z)]$. And a Concept like BEINGASTALLAS(x, Carl) would be hereditary on the MT-series starting with the given entity. This does not presuppose that ‘TALLER(Carl, Al)’ is a claim to the effect that Carl’s height-number is greater than Al’s height-number. On the contrary, ‘>(3, 1)’ is equivalent to ‘PRECEDES(1, 3)’, which does *not* say that 3’s number-number exceeds 1’s number-number. Second-order quantification might thus render appeal to heights superfluous. But this is getting ahead, since there is still a little more of semantic interest to extract from Frege’s result.

1.4 Back to the Theorem

Two of the Dedekind-Peano axioms follow almost immediately from Frege’s defintions of zero and number.

(DP:I) $N0$
(DP:II) $\forall x[Nx \supset \neg S0x]$

Like (DP:IV), these are *logical* consequences of the definitions. The third axiom,

(DP:III) $\forall x\forall y[(Nx\ \&\ Ny\ \&\ x \neq y) \supset \neg\exists z(Szx\ \&\ Szy)]$

according to which no two numbers have the same successor can be rewritten as follows: $\forall x\forall y\forall z[\text{PREDECESSOROF}(x, z)\ \&\ \text{PREDECESSOROF}(y, z) \supset x = y]$. And this is a logical consequence of Hume’s Principle, repeated below.

(HP) $\forall F\forall G[(\#F = \#G) \equiv \text{OneToOne}(F, G)]$

A pair of Concepts correspond one-to-one, and are equinumerous in this Cantorian sense, iff some function $\mathfrak{F}$ meets the condition below.

$$\forall x\{Fx \supset \exists y[Gy\ \&\ \supset \mathfrak{F}(x) = y]\}\ \&\ \forall x\{Gx \supset \exists y[Fy\ \&\ \supset \mathfrak{F}(x) = y]\}$$

Given (HP), together with the fact that each number w numbers the Concept of being a number that precedes w, it follows that PREDECESSOROF(x, y) is a one-to-one relation; in which case, the rewrite of (DP:III) follows immediately. Frege proves that PREDECESSOROF(x, y) is one-to-one, given (HP), by establishing an intuitive principle: *if* F and G are equinumerous Concepts, and x falls under F while y falls under G, *then* the Concepts $[Fz \mathbin{\&} \neg(x = z)]$ and $[Gz \mathbin{\&} \neg(y = z)]$ are also equinumerous. In short, subtracting one thing from each side of a one-to-one correspondence leaves a one-to-one correspondence; hence, PREDECESSOROF(x, y) is one-to-one. So while logic alone is not enough to prove (DP:III), Frege's second-order logic does provide a framework for deriving the axiom from (HP) without any further assumptions about numbers.

Likewise, the "generative" axiom that every number has a successor

(DP:V) $\forall x[Nx \supset \exists y(Syx)]$

follows from (HP) without further nonlogical/nondefinitional assumptions. Rewriting in terms of Frege's central notion, the axiom says: ∀x{Nx ⊃ ∃y[Ny & PREDECESSOROF(x, y)]}. This follows from a claim that will by now be familiar: ∀w[PREDECESSOROF(w, #{<u>PRECEDES</u>(z, w)})]; each number w is the predecessor of the number of numbers that precede or are identical with w. The proof of this claim, which draws together much of what Frege had already established at this stage, is not hard. But the details are sufficiently many that I refer interested readers to Zalta (2003) initially, and then the essays in Demopolous (1994), especially Heck (1994).

For our purposes, it should be sufficiently clear that given Frege's definitions, it follows from (HP) that the number three is the *predecessor of* the number of things that *precede or are identical with* the number three. These things are none other than the numbers zero, one, two, and three; and these things number four, which is by definition, the number of things identical with zero, one, two, or three. Likewise, four is the predecessor of five, and so on.

This is at least suggestive of how one might bootstrap from a language in which 'two' and 'three' figure as second-order predicates, as in 'There are three apples', into a language in which such words can also be *mentioned* as elements of a list—'one', 'two', 'three', ...—with each word ***w*** being a predicate (that when used is) satisfied by the words in the list up to and including ***w***; cf. Benacerraf (1965), Hurford (1987), Gallistel and Gelman (1991, 2000), Gallistel, Wiese (2003), Gelman and Cordes (2005).

The idea would be that the "primary" meanings of number-words are given roughly as follows: a first-order Concept X falls under the second-order Concept indicated with 'three' iff three things fall under X. There are various ways of saying what it is for three things to fall under X—or for the Xs to be three, or for 'Three(X)' to be true—including the familiar first-order characterization: $\exists x \exists y \exists z \forall w\{Xx \& Xy \& Xz \& [Xw \supset (w = z) \vee (w = y) \vee (w = x)]\}$. Though at some point, Fregean recursion will presumably be necessary: $Four(X) \equiv \exists Y \exists z\{Three(Y) \& \neg Yz \& \forall x[Xx \equiv Yx \vee (x = z)]\}$. But however one specifies the meanings of words like 'three', such words can be listed. And if 'three' is a second-order predicate, then even when mentioned in a list, it remains a predicate satisfied by *any* three things—including the words 'one', 'two', and 'three'. This invites the hypothesis that at least for children at a certain stage of acquiring English, number-words *are* things that can number Concepts. And perhaps in particular: 'three' = #[(x = 'one') v (x = 'two') v (x = 'three')]. A child might use the *word* 'three' to number this special metalinguistic Concept precisely because 'three' retains its predicative meaning when listed third.[10] And given that 'three' numbers this doubly disjunctive Concept, knowledge of Hume's Principle might let the child conclude that 'three' also numbers any equinumerous Concept. In which case, the child could know that 'three' numbers any Concept$_X$ such that the things falling under it$_X$ correspond one-to-one with the following number-words: 'one', 'two', 'three'. But I will not pursue this Fregean speculation any further here.

2. Comprehension Without Sets

It is more important, in the context of stressing that induction is independent of quantification over abstracta, to introduce an alternative to Frege's interpretation of second-order quantification as effectively first-order quantification over Concepts. And here, it is useful to think about the following comprehension schema: $\exists X \forall x(Xx \equiv \Phi x)$; where the schematic predicate 'Φx' can be replaced by any well-formed open sentence, like 'Prime(x)' or 'Prime(x) & >(x, 2) ⊃ Odd(x)', with 'x' as the unbound variable.

2.1 Avoiding Russell's Paradox

10 If it helps, we can distinguish adjectival from nominative uses of 'three' as follows, using the subscripts 'A' and 'N' to disambiguate: the adjective 'three$_A$' indicates a higher-order Concept such that a Concept X falls under this higher-order Concept iff three things fall under X; and this adjective, which can be *mentioned*, is the thing denoted by uses of the name 'three$_N$'. Indeed, if nominative reflecting a kind of quotational use, perhaps the adjective 'three$_A$' just is the semantic value of the mentioned expression '*three*$_A$'. Perhaps an entity/word x is a value of '*three*$_A$' iff x = 'three$_A$', while a Concept X is a value of 'three$_A$' iff three things fall under X.

Frege's intent was that all instances of the schema would be theorems of logic. His idea was that given any predicate 'Φx', there would be a Concept X—a function from entities to truth or falsity—such that X maps each entity to truth iff that entity falls under the Concept indicated with 'Φx'. This is trivial, as Frege intended it to be, *if* each predicate indicates a Concept. But Frege also held that each Concept has an extension: for each Concept, there is the (perhaps empty) set of entities of entities that the Concept maps to truth. And as Russell noted, this leads to contradiction, given open sentences like '$x \notin x$'.

According to Frege, if $\exists X \forall x(Xx \equiv x \notin x)$, then some Concept$_X$ is such that each entity$_x$ falls under it$_X$ iff that$_x$ entity is not an element of itself$_x$. But any such Concept has an extension,

$\{x: x \notin x\}$, which we can call 'Bert'. And this alleged set, Bert, either is or is not an element of itself. Now if Bert *is* an element of itself, then Bert is an element of $\{x: x \notin x\}$, and so Bert *is not* an element of Bert; yet if Bert *is not* an element of itself, then Bert is an element of $\{x: x \notin x\}$, and so Bert *is* an element of Bert. The supposition that Bert exists thus implies a contradiction. So it's false that Bert exists. So '$x \notin x$' does not yield a true instance of Frege's comprehension schema. Likewise, the formal sentence '$\exists X \forall x(Xx \equiv x \notin x)$' is false if 'X' is interpreted as ranging over *sets* of entities that 'x' ranges over, and the sets are among the things that 'x' ranges over. This spelled doom for Frege's attempt to derive Hume's Principle from logic alone.

One can respond by restricting the domains for variables of various types. But as Boolos (1984, 1998) shows, there is another coherent gloss of second-order variables according to which '$\exists X \forall x(Xx \equiv x \notin x)$' is true—or at least compatible with the absence of any set that is the set of nonselfelemental things. We can interpret second-order variables as *plural* variables, each of which has *more than one* value relative to each assignment of values to variables. On this construal, '$\exists X(Xe)$' means that *one or more things*, the Xs, *are* such that entity e is *one of them*. And '$\forall X(Xe)$', which is logically equivalent to '$\neg \exists X \neg(Xe)$', means that there are no(t one or more) things such that e is one of them. We can also gloss '$\forall X(Xe)$' as follows: whatever the Xs are, e is one of them; see Lewis (1991), Linnebo (2004) and references there. In terms of open sentences, the idea is that 'Xx' is true relative to an assignment **A** of values to variables iff the entity that **A** assigns to 'x' is one of the (one or more) entit*ies* that **A** assigns to 'X'.

At least prima facie, this does not just restate a Fregean interpretation, according to which 'Xx' is true relative to **A** iff the entity that **A** assigns to 'x' is an element of the extension (of the Concept) that **A** assigns to 'X'. For on the Boolos construal, '$\exists X \forall x[Xx \equiv (x \notin x)]$' means that there are one or more things$_X$ such that for each thing$_x$, it$_x$ is one of them$_X$ iff it$_x$ is not an element of itself$_x$. And at least prima facie, this claim is true. There are indeed some things, like you and me, that are not selfelemental; and the nonselfelemental things all exist, even if no *set* is such that its members are them—i.e., even if no set s is such that $x \in s$ iff $x \notin x$. But if a formal sentence is true on one interpretation and false on another, the interpretations differ. Skeptics may suspect that '$\exists X \forall x[Xx \equiv (x \notin x)]$' remains false with 'X' interpreted as a plural variable.[11] But suspicion is not an argument. And the plural interpretation certainly seems to be a coherent, apparently distinctive construal of second-order variables.

Consider a domain with exactly five entities: a, b, c, d, and e. Given such a domain, there seem to be thirty-one possibilities for assigning values to variables, as depicted below.

—	a	b	ba	c	ca	cb	cba
d	da	db	**dba**	dc	dca	dcb	dcba
e	ea	eb	eba	ec	eca	ecb	ecba
ed	eda	edb	edba	edc	edca	edcb	edcba

If we think about assignments as ways of modelling potential acts of demonstration, then the blank corresponds to cases of demonstrating nothing, while other "cells of the lattice" correspond to cases of demonstrating one or more entities in the stipulated domain. Initially, one might be inclined to say that each of nonempty cells indicates exactly one set-like entity with one or more elements of the five-membered domain, {a, b, c, d, e}; see Link (1983). On this essentially singularist conception of assignments, according to which each assignment assigns at most *one* value to each variable, we don't really have a domain with exactly five entities. If the first-order variables range over a domain of five things, the second-order variables range over a domain of at least thirty-one.[12] But this familiar construal of the lattice is not

11 Perhaps whenever we judge that some entity is *one of* some things we can think about, we must thereby think of the entity as satisfying some condition like the following: it falls under a Concept which applies to all and only those things; or it is an element of a set such that each of those things is an element of the set.

12 The empty set would make thirty-two. And what about entities like {{a}, a, b, c, d, e}? Schwartzschild (1996) argues against expanding the domain in this way for the plural variables of natural language. But formally, many options are available at this point. Note, however, there is nothing puzzling about assigning more than one value to a variable. Assigning exactly one entity to a singular variable, like 'it', is akin to an act of demonstrating that entity alone. Likewise, an act of demonstrating several

mandatory.

Consider the eleventh cell, indicated with '**dba**'. Instead of thinking about assigning the set {d, b, a} as *the* value of a variable, we can think about three entities—d, b, and a—as the value*s* of that variable. To highlight this contrast, imagine binary numerals, with our five entities numbered as follows: a, 1; b, 10; c, 100; d, 1000; and e, 10000.

—	00001	00010	00011	00100	00101	00110	00111
01000	01001	01010	**01011**	01100	01101	01110	01111
10000	10001	10010	10011	10100	10101	10110	10111
11000	11001	11010	11011	11100	11101	11110	11111

Now the eleventh cell is indicated with '**01011**', which designates the sum of three entity correlates: 01011 = 1000 + 10 + 1. One can still hypothesize that this arithmetic relation reflects a set-forming, or perhaps merelogical operation. From this perspective, '01011' stands for a plural entity x_{pl} such that y is an element of x_{pl} iff y is identical with d or b or a. But we can also read '01011' as five answers to yes/no ($\top/\bot$) questions about whether a certain entity, perhaps with others, is assigned to a given variable: (e, $\bot$), (**d**, $\top$), (c, $\bot$), (**b**, $\top$), (**a**, $\top$). This construal may also require a slightly enlarged domain, in so far as it requires the "sentential" values $\top$ and $\bot$. But appeal to *this* Fregean trick need not be combined with the *further* trick of associating each plural variable with a plural entity (perhaps via some Concept); cf. Link (1983).

In this regard, it is worth recalling that Frege appealed to plural entities—like nonempty, nonsingleton extensions—as entity-correlates of unsaturated Concepts. And if we set aside his attempt to derive Hume's Principle, Frege appealed to Concepts as a way of implementing his idea that a "gappy" sentence like 'Prime(x)' associates each element of the domain in question with a sentential value. But this idea is better than Frege's implementation of it. Treating each meaningful predicate as an indicator of some Concept, which maps each entity to $\top$ or $\bot$, has a problematic implication: the mere existence of the predicate guarantees the existence of the corresponding mapping/extension. We can say instead that '**Clever**(**u**)' does indeed

---

things is akin to assigning more than one entity to a plural variable. Given a tendentious semantic theory, one might insist that what we call an act of demonstrating several things is really an act of demonstrating a plural thing (with elements). But prima facie, this is the fancy idea in need of theoretical support. And there is much to be said in favor of the hypothesis that human languages employ plural variables, each of which can have many values relative to an assignment; see Boolos (1998), Schein (1993, 2006, forthcoming), Higginbotham (1998).

imply '$\exists X\forall x\{[Xx \equiv$ **Clever**(x)] & X**u**$\}$', but that this does *not* logically imply the existence of a correlated set, $\{x$: **Clever**(x) $= \top\}$.[13] On the Boolos construal, '**Clever**(**u**)' and '$\exists X\forall x\{[Xx \equiv$ **Clever**(x)] & X**u**$\}$' both mean that **u** is one of the (one or more things that are) clever.

As Sainsbury (1990) stresses, there may not *be* a set whose elements are all and only the clever. Unlike 'prime', 'clever' is vague; and there is no set x and entity e such that it is vague whether or not e is an element of x. But even if there is a set of the clever, there is no Zermelo-Frankl set of all and only the nonselfelemental. Yet there surely are one or more nonselfelemental things. And at least prima facie, we can express this thought with '$\exists X\forall x[Xx \equiv (x\notin x)]$', letting the singular variable 'x' and the plural variable 'X' range over the things—with 'X' having more than one value relative to each assignment of values to variables. Readers who worry about unrestricted quantification over everything can replace 'things' in the previous sentence with 'Zermelo-Frankl sets'. At this point, I think the burden of argument lies with skeptics who think that the Boolos-construal is really just Frege's construal in disguise. Elsewhere, I have argued that the former construal is preferable with respect to the second-order variables in theories of meaning for natural languages; see Pietroski (2005, 2006), drawing on Schein (1993, 2001, forthcoming). But for now, my point is only that Frege's interpretation in terms of Concepts is not mandatory. Boolos provides a coherent alternative with some prima facie attractions.

2.2 Formalism and Induction (Reprise)

In an important sense, a variable that can be assigned value*s* adds nothing new, while a variable that is always assigned a set of one or more values may add paradoxically much. As Boolos (1998, p.72) says, "We need not construe second-order quantifiers as ranging over anything other than the objects over which our first-order quantifiers range...a second-order quantifier needn't be taken to be a kind of first-order quantifier in disguise, having items of a special kind, collections in its range." This matters in part because genuinely plural variables make room for essentially plural predicates. Some things can *plurally* satisfy an essentially plural predicate even if no one thing can satisfy the predicate. Boolos (1984) offers, among others, the example 'rained down'; some rocks can rain down even if no thing can. Schein (1993) offers

---

13 Boolos (1985) asks, reporting conversation with a skeptic, is anyone who believes that Napoleon was *not* one of his ancestors thereby committed to the existence of sets? As Boolos notes, while "Frege's definition, whose logical utility, fruitfulness, and interest have been established beyond doubt, cannot be dismissed for such an utterly crazy reason, it is not at all easy to see what a good answer to [the skeptic's] question might be."

'clustered'; some elms can be clustered in the middle of the forest even if no single thing can be clustered anywhere. And importantly, given some things, they are sure to be plural in way that no thing can be. Unsurprisingly, 'plural' is a plural predicate *par excellence*.

So we can introduce a pair of restricted quantifiers, '$\exists$X:Plural(X)' and '$\exists$X:$\neg$Plural(X)'; where the latter is equivalent to '$\exists$x', and $\exists$X:Plural(X)[$\forall$x:Xx(Fx)] iff $\exists x \exists y$[Fx & Fy & $x \neq y$]. By contrast, $\exists$X:$\neg$Plural(X)[$\forall$x:Xx(Fx)] iff one or more things *such they are not more than one* are such that each of them is an F. So if $\exists$X:Plural(X)[$\forall$x:Xx(Fx)], $\exists$X:$\neg$Plural(X)[$\forall$x:Xx(Fx)]. As noted above, we can also treat numerical predicates as second-order, with the first few understood in terms of more basic notions, but eventually by means of recursion. Whatever the one or more Xs may be: One(X) $\equiv$ $\neg$Plural(X); Two(X) $\equiv$ $\exists x \exists y$[Xx & Xy & ($x \neq y$)]; AtLeastTwo(X) $\equiv$ Plural(X); MoreThanTwo(X) $\equiv$ Plural(X) & $\neg$Two(X); etc.

Given this understanding of the formalism, the mathematical axiom of induction

(DP:IV) $\quad \forall X[X0 \mathbin{\&} \forall x \forall y(Xx \mathbin{\&} Syx \supset Xy) \supset \forall x(Nx \supset Xx)]$

says that whatever the Xs may be: *if* zero is one of them, and whenever something is one of them, its successor is also one of them, *then* every number is one of them. And this is still a special case of the corresponding principle of logic.

(6a) Given a relation R(x, y), a Concept X, and an entity e:
if X is hereditary on the R-series starting with e, and e falls under X,
then $\forall x[\underline{R}(e, x) \supset Xx]$.

Reading this formalism with the second-order variables interpreted plurally, it says that given one or more ordered pairs, the Rs, and one or more things, the Xs, and an entity e: *if* the Xs are hereditary on the Rs starting with (e, y) for some entity y, and e is one of the Xs, *then* every entity x is one of the Xs if (e, x) is one of the ordered pairs defined ancestrally in terms of the Rs. Frege's quantification over Concepts is thus dispensible, at least for cases of induction involving relations that can be captured in terms of predicates of ordered pairs.[14]

14

In the Appendix to Lewis (1991), cowritten with Burgess and Hazen, the authors explore an alternative to actually quantifying over pairs: namely, quantifying over individuals in monadic third-order logic. See Hazen (1997a, 1997b, 2000); see also Linnebo (2005). But for purposes of natural language semantics, as opposed to pure logic, quantification over ordered pairs is relatively innocent. See

Induction is not essentially about numbers, any more than it is about parents. And given the Boolos interpretation of second-order variables as plural variables, neither is induction covertly about Concepts or their extensions. When we reason by induction about people, while thinking about who begat who (or who is taller than who), we are reasoning in a second-order way that can be captured with variables ranging plurally over "the objects over which our first-order quantifiers range." Our second-order thoughts and inferences need not be construed as first-order thoughts and inferences that range, in a disguised way, over items of a special kind. Induction is *not* a special kind of first-order inference over abstracta that are somehow related to elements in a more basic domain. Induction is a kind of second-order inference that may, in special cases like arithmetic, be applied to abstract domains.[15]

2.3 Case Study: Labelled Phrase Markers

Imagine a language with primitive expressions of three kinds: alphas, betas, and gammas. Combining a gamma with an alpha or a beta creates a complex gamma. Combining a beta with an alpha creates a complex beta. (Think of alphas as adjuncts, and gammas as predicates that take betas as arguments.) This allows for expression types like the one indicated below.

```
        γ
       / \
      γ   α
     / \
    γ   β
       / \
      β   α
```

Such a language might have endlessly many expressions, even if there are finitely many alphas, betas, and gammas. Whatever expressions *are*, they can and often must be specified recursively. This makes it tempting to identify "labelled" expressions with sets. If [blob$_\beta$ alla$_\alpha$]$_\beta$ is an expression, one might take it to *be* the set {blob$_\beta$ {blob$_\beta$,

---

Pietroski (2005) for further discussion in the context of neo-Davidsonian approaches to semantic composition.

15

See Boolos (1987) for an argument that we can (quickly and easily) recognize as valid *many* inferences that are firstorderizable "in principle," but only in proofs whose steps would *far* outnumber the subatomic particles in the known universe.

$alla_\alpha\}\}$, conventionally identified with the ordered pair <$blob_\beta$, $alla_\alpha$>; cf. Chomsky (1995). This encodes the fact that 'blob alla' is, for purposes of further combination, like 'blob'. Still, the idea of identifying expressions with sets rankles. Identifying numbers with sets is bad enough, since as Benaceraff (1965) notes, any particular identification seems arbitrary: is the number two identical with $\{\{\varnothing\}\}$, or $\{\varnothing, \{\varnothing\}\}$, or some other set? And while we can represent certain properties of expressions by correlating expressions with sets, it seems even more gratuitous to identify particular expressions with particular sets. Of course, one can ignore this if one isn't really worried about what expressions *are*. But one might have thought that we should be able to characterize expressions recursively without tendentious claims about what expressions are. And there is indeed another option, given Frege's treatment of induction and Boolos' construal of variables.[16]

Let's say that x ImmediatelyDominates y—for short, ID(x, y)—iff y is a constituent of x, *and* no constituent of x has y as a constituent. Eventually, 'constituent of' will be defined. But for now, an intuitive grasp of the relation ID(x, y) will suffice. If the only way to form complex expressions in our imagined language is by concatenating simpler ones, then every complex expression is of the form x^y, for some pair of expressions x and y. In which case, ID(x, y) $\equiv \exists z(x = z$^$y)$. For illustration, suppose that 'Glug blob alla' is of the form $Glug_\gamma$^($blob_\beta$^$alla_\alpha$), and hence of the form $[Glug_\gamma\ [blob_\beta\ alla_\alpha]_\beta]_\gamma$.[17] The primitive constituents, which are marked as expressions of certain types, immediately dominate nothing. But the complex expression $blob_\beta$^$alla_\alpha$ ImmediatelyDominates both $blob_\beta$ and $alla_\alpha$. Likewise, $Glug_\gamma$^($blob_\beta$^$alla_\alpha$) ImmediatelyDominates both the primitive gamma $Glug_\gamma$ and the complex beta $blob_\beta$^$alla_\alpha$. But $Glug_\gamma$^($blob_\beta$^$alla_\alpha$) does not ImmediatelyDominate $blob_\beta$ or $alla_\alpha$. If 'Glug blob alla affa' is of the form ($Glug_\gamma$^($blob_\beta$^$alla_\alpha$))^$affa_\alpha$, then it is of the form $[[Glug_\gamma\ [blob_\beta\ alla_\alpha]_\beta]_\gamma\ affa_\alpha]_\gamma$. This complex gamma ImmediatelyDominates $affa_\alpha$ and $[Glug_\gamma\ [blob_\beta\ alla_\alpha]_\beta]_\gamma$, but nothing else.

There are obvious analogies between the intransitive relation ID(x, y) and the relation SUCCESSOROF(x, y). Intuitively, there is also a transitive relation

16

Brody (2000) develops a similar line of thought internal to a certain theory of syntax. But my point here is simply to illustrate the power of Frege's logic with an explicitly linguistic example.

17

In terms of Chomsky's (1957) *is-a* relation, suppose that the string of words (or word-sounds) 'Glug blob alla' is-a $[\ldots_\gamma\ [\ldots_\beta \ldots_\alpha]]$. Then given the rules for determining the type of a complex expression 'Glug blob alla' is-a $[\ldots_\gamma\ [\ldots_\beta \ldots_\alpha]_\beta]_\gamma$. This would be an idealized claim about how certain linguistic signals are classified by certain speakers.

DOMINATES(x, y) that $Glug_{\gamma}$^($blob_{\beta}$^$alla_{\alpha}$) bears to each of its constituents—$Glug_{\gamma}$, $blob_{\beta}$^$alla_{\alpha}$, $blob_{\beta}$, and $alla_{\alpha}$—analogous to the transitive relation GREATERTHAN(x, y). So unsurprisingly, DOMINATES(x, y) iff [**ANC**(ID)](x, y). And so in particular, DOMINATES([$Glug_{\gamma}$ [$blob_{\beta}$ $alla_{\alpha}]_{\beta}]_{\gamma}$, $blob_{\beta}$) iff: $\forall$X{X([$Glug_{\gamma}$ [$blob_{\beta}$ $alla_{\alpha}]_{\beta}]_{\gamma}$) & $\forall$w$\forall$z[X(w) & ID(w, z) ⊃ X(z)] ⊃ X($blob_{\beta}$)}; whatever the Xs may be, *if* the complex expression is one of them, and whenever one of them immediately dominates something, that (simpler) expression is one of them, *then* $blob_{\beta}$ is one of them. Since this second-order universal generalization is true, [$Glug_{\gamma}$ [$blob_{\beta}$ $alla_{\alpha}]_{\beta}]_{\gamma}$ sominates $blob_{\beta}$, albeit mediately. Or putting the point in the other direction: IMMEDIATECONSTITUENTOF(x, y) iff $\exists$z(y = x^z); and CONSTITUENTOF(x, y) iff [**ANC**(IMMEDIATECONSTITUENTOF)](x, y).

We can now characterize the gammas simply as follows.

$\forall$x{Gamma(x) ≡ $\exists$y{PrimitiveGamma(y) & [CONSTITUENTOF(y, x) v (x = y)]}}

This formulation, which mimics Frege's definition of number, implies that x is a *non*primitive Gamma iff x has some primitive gamma as a constituent. Spelling out 'CONSTITUENTOF(y, x)', which abbreviates a second-order generalization, would reveal the underlying inductive truism. An expression y is a constituent of a distinct expression x iff: whatever expressions you choose, x is sure to be one of them *if* those expressions are such that (i) they include every expression that has y as an immediate constituent, and (ii) for each of them, anything that has it as an immediate constituent is also one of them. Put yet another way, the biconditional above says that x is gamma iff: x is a primitive gamma, *or* for some primitive gamma y, x is among any expressions that include y and the ImmediateDominater of each of those expressions.

Given a characterization of the gammas, the betas can be described in similar fashion.

$\forall$x{Beta(x) ≡ ¬Gamma(x) &
$\exists$y{PrimitiveBeta(y) & [CONSTITUENTOF(y, x) v (x = y)]}}

That is, x is a beta iff: x is not a gamma; and for some primitive beta y, x is a constituent of or identical to y. And in the imagined language, x is an alpha iff x is neither a beta nor a gamma. For purposes of characterizing expressions in this language, we needn't think of alphas as having labels of their own; alphas are simply expressions that extend a beta or a gamma. But it is important that the relevant expression types exhibit a compositional hierarchy. Suppose that combining a gamma with a beta yielded a gamma, and combining a beta with alpha yielded a beta, while combining a gamma with an alpha yielded an alpha. (Think of rock, scissors, paper.) Then we couldn't characterize the expressions by characterizing the gammas, as above, and then the nongammas. But if the expression types do exhibit an appropriate hierarchy, recursive descriptions of the expressions are easily provided, given a logic that allows quantification into positions occupiable by predicates.

Let me summarize by reformulating the main point. Without introducing labels for complex expressions, we can say that [Glug$_\gamma$ [blob$_\beta$ alla$_\alpha$]] is the result of combining blob$_\beta$ with alla$_\alpha$, and then combining the result with Glug$_\gamma$. And given the following principle,

$\forall$x{Gamma(x) $\equiv$ $\exists$y{PrimitiveGamma(y) & {x $\neq$ y $\supset$
$\exists$z[ID(z, y)] & $\forall$X{$\forall$z:ID(z, y)[Xz] & $\forall$z$\forall$w[Xz & ID(w, z) $\supset$ Xw] $\supset$ Xx}}}}

it follows that [Glug$_\gamma$ [blob$_\beta$ alla$_\alpha$]] is a gamma. And suppose we are also given that [blob$_\beta$ alla$_\alpha$] is not a gamma, because it is either a beta or an alpha. Then the following principle

$\forall$x{Beta(x) $\equiv$ $\neg$Gamma(x) & $\exists$y{PrimitiveBeta(y) & {x $\neq$ y $\supset$
$\exists$z[ID(z, y)] & $\forall$X{$\forall$z:ID(z, y)[Xz] & $\forall$z$\forall$w[Xz & ID(w, z) $\supset$ Xw] $\supset$ Xx}}}}

ensures that [blob$_\beta$ alla$_\alpha$] is beta. We can encode these consequences with further subscripts, as in [Glug$_\gamma$ [blob$_\beta$ alla$_\alpha$]$_\beta$]$_\gamma$. But none of this implies that labelled expressions are sets with elements corresponding to the labels. For we can intepret the capitalized variables as ranging plurally over the expressions that the first-order variables range over—without quantifying over *sets* of expressions, or taking expressions to *be* sets. We can think about expression*s* without mischaracterizing them, or our thoughts about them, in terms of abstracta distinct from the expressions themselves. This invites the thought that for purposes of stating theories of natural language *syntax*, we should employ a second-order logic with the predicative variables interpreted *a la Boolos*.

3 Comparing the Perceptible

Suppose there are, as there seem to be, distinctively human capacities that let us recursively generate labelled phrase markers as we do. And suppose, more tendentiously, that these capacities also underlie our ability to reason inductively in the ways illustrated above. In particular, let's speculate that a competent speaker of a natural language can represent the world in second-order terms—though perhaps only in a restricted way, via the logical resources of a *monadic* predicate calculus whose predicative variables are understood as plural variables. This is, in effect, to posit a capacity to generate "logical skeletons" that are more Fregean than the subject-predicate structures envisioned by medieval logicians, but perhaps more constrained and reflective of grammatical structure than Frege envisioned. If humans have such a capacity, we might understand words like 'tall' and 'taller' as monadic predicates used to express relational thoughts *indirectly*, in second-order terms.

Elsewhere, I have argued that for purposes of compositional semantics, the resources of a Boolos-style second-order monadic predicate calculus are sufficient, required, and yet limited enough to help explain various constraints on natural language; see Pietroski (2005, 2006). To make a long story very short, all the usual textbook cases and more can be handled by supposing that speakers can do the following: *lexicalize* mental representations as (potentially plural) monadic predicates—like 'Red(X)', 'Ball(X)', and predicates like 'Stab(E)' that can be satisfied by things like events, which can have "participants;" *conjoin* monadic predicates, as in 'Red(X) & Ball(X)'; introduce certain *thematic predicates*, like 'Agent(E, __)' and 'Theme(E, __)', when combining grammatical predicates with arguments as in 'They_1 stabbed them_2'—thereby creatating complex monadic predicates like 'Agent(E, they_1) & PastStab(E) & Theme(E, them_2)'; and occasionally introduce *existential closure* of a variable.[18] On this view, lexicalization is a matter of imposing a common "monadic format" on mental representations—which have whatever (psychological) adicities they have—in order to combine linguistic correlates of these mental representations via the essentially conjunctive combinatorics provided by the human language faculty. While this conception of lexicalization is certainly controversial, let me repeat that it is compatible with the usual textbook cases and more. Indeed, I have argued that it helps explain some otherwise puzzling facts about causative constructions, propositional attitude reports, and natural language quantification. Here, though, I want to sketch the implications for comparative constructions, in the context of Frege-Boolos accounts of inductive inference.

3.1 Big Ants are Bigger than Small Ones

If we know that Adam is a big ant, we know something about Adam *and* the ants. To a first approximation, we know that the ants are such that Adam is bigger than most of them—and hence, that most of the ants are smaller than Adam. If we take our size concepts to be comparative/relational, not monadic, this invites a logical paraphrase of (7) along the lines of (8).

(7) Adam is a big ant

(8) $\exists X{:}\forall x[Xx \equiv \text{Ant}(x)]\{X(\text{adam})\ \&\ \text{MOST}(X){:}\text{B}{\small\text{IGGER}}(\text{adam}, X)]\}$

18

More precisely, 'Agent(E, they_1,)' should be spelled out as follows: Agent(E, they_1, **A**) iff $\exists X{:}\forall x[Xx \equiv \text{Assigns}(\mathbf{A}, x, 1)]\{\text{Agent}(E, X)\}$. That is, the Agents of the Es are they_1 relative to assignment **A** iff the things that **A** assigns to the first index are the Agents of the Es. And likewise for 'Theme(E, them_2, **A**)'.

Adam is a big ant iff the ants are such that Adam is one and bigger than most. We can abbreviate 'MOST(X):BIGGER(adam, X)]', which says that most Xs are such that Adam is bigger, as follows: BIG-ONE(adam, X); or [BIG-ONE(X)](adam). It may well be that 'most' is not quite what is wanted for these purposes. But given a better way of specifying how the ants must be related to the things/ants that are bigger than Adam, in order for (7) to be true, we can presumably make an appropriate substitution for 'MOST' in (8).

In terms of the syntax and compositional semantics underlying (7), we might imagine a variant on Higginbotham's (1983) appeal to autonymous theta-marking. Suppose that 'big' combines with a covert pronoun conindexed with 'ant', as in (7a).

(7a) Adam is a big-~~one~~$_1$ ant$_1$

Then we can specify the semantic values of 'ant$_1$', relative to assignments of values to variables, as follows: Val(X, 'ant$_1$', **A**) iff $\exists$Y:$\forall$y[Yy $\equiv$ Ant(y)]{$\forall$x:Xx(Yx) & $\forall$x[Assigns(**A**, x, 1) $\equiv$ Yx]}; where 'Assigns(**A**, x, 1)' means that **A** assigns x as one of the one or more values of the first indexed variable. On this view, the Xs are values of 'ant$_1$' relative to **A** iff: **A** assigns the ants to the first index, *and* each of the Xs is one of the things assigned to that index. The semantic values of 'big-~~one~~$_1$'can be specified correlatively.

Val(X, big-~~one~~$_1$, **A**) iff each of the Xs is
bigger than most of things that **A** assigns to the first indexed variable

The idea is that 'big' combines with 'one$_1$', or at least that 'big' is understood as 'big-one$_1$', because each terminal node in the syntax must be interpreted as a monadic predicate. Likewise, I suggest, 'bigger' is understood as a monadic predicate. But 'bigger' is more like 'stabbed', which can take an internal and external argument, and thereby spread the relationality of the underlying concept across a sentence in which *arguments of* the predicate are interpreted thematically.

We know, from developments of Davidsonian event analyses (see, e.g., Taylor [1985]), that 'stabbed' can be treated as a monadic predicate of events as in (9);

(9) $\exists$e[External(e, Brutus) & Event(e) & PastStab(e) & Internal(e, Caesar)]

where by stipulation, the external participant of an event is its Agent, and the internal participant of an event is its Theme. Event variables can also be plural, as in (10),

(10) $\exists$E[External(E, They$_1$) & Event(E) & PastStab(E) & Internal(E, them$_2$)]

which can be spelled out as in (11), relativizing to an assignment **A**;

(11) $\exists E\{\iota X{:}\forall x[Xx \equiv \text{Assigns}(\mathbf{A}, x, 1)]\{\text{External}(E, X)\}\ \&\ \forall e{:}Ee[\text{Event}(e)]\ \&$
$\forall e{:}Ee[\text{PastStab}(e)]\ \&\ \iota X{:}\forall x[Xx \equiv \text{Assigns}(\mathbf{A}, x, 2)]\{\text{Internal}(E, X)\}\}$

where 'External(E, X)' can be firstorderized as '$\forall e{:}Ee\{\exists x{:}Xx[\text{External}(e, x)]\}\ \&\ \forall x{:}Xx\{\exists e{:}Ee[\text{External}(e, x)]\}$', and likewise for 'Internal(E, X)'. So construed, (10) says that one or more things satisfy four conditions: their external participants were the things assigned to the first indexed variable; each of them was an event—a thing whose external participant is an Agent, and whose internal participant is a Theme; each of them was a stab; and their internal participants were the things assigned to the second indexed variable. So more briefly, (10) says that one or more events of stabbing were such that their Agents were the things assigned to the first indexed variable, and their Themes were the things assigned to the second indexed variable.

This provides an attractive analysis of (12); see Schein (1993), Pietroski (2005).

(12) They stabbed them

And given the Boolos construal of '∃E...E...', (9) can be written as (13),

(13) $\exists E[\text{External}(E, \text{Brutus})\ \&\ \text{Event}(E)\ \&\ \text{PastStab}(E)\ \&\ \text{Internal}(E, \text{Caesar})]$

thus providing a unified analysis for both (12) and 'Brutus stabbed Caesar'. Moreover, each stabbing is something like a causal process, starting with the action of an Agent and ending with the motion of an implement presumably used to stab something. As discussed above, this invites appeal to the relation ACTCAUSE(c, e), defined ancestrally in terms of PROXIMALCAUSE(d, d'). But my main point here is that parallel analyses for (14) and (15) are available,

(14) Carl is bigger than Adam
(15) They are taller than them
as shown in (16) and (17).

(16) $\exists O[\text{External}(O, \text{Carl})\ \&\ \text{OrderedPair}(O)\ \&\ \text{Bigger}(O)\ \&\ \text{Internal}(O, \text{Adam})]$
(17) $\exists O[\text{External}(O, \text{They})\ \&\ \text{OrderedPair}(O)\ \&\ \text{Taller}(O)\ \&\ \text{Internal}(O, \text{them})]$

Note that (15) can be used to correctly report that some basketball players (they, the ones wearing red) are taller than some others (them, the ones wearing blue), even if the relevant comparisons must be position by position. Suppose the shorter center is taller than every player apart from the taller center; the shorter left guard is taller than every player apart from the two centers and the taller left guard; etc.[19] One can still capture the truth of (15), in such a context, by treating 'taller' as a monadic predicate. The predicate is satisfied by some ordered pairs, the Os, iff for each$_o$ of them: ∃x∃y{External(o, x) & Internal(o, y) & TALLER(x, y)}; or put another way, ∃x∃y{External(o, x) & Internal(o, y) & [**ANC**-MINIMALLYTALLER](x, y)}. This is, by no means, a *theory* of comparative constructions. But it is worth noting that on this kind of view, (14) and (15) do not involve covert quantification over numbers or heights. And even the quantification over ordered pairs may, in the end, be dispensible.

3.2 Series of Things

Suppose that at least within a given context, we can make sense of the idea that some things can (together) exhibit a MINIMALLYTALLER-series. This does not require that in every context, one thing is minimally taller than another iff the difference between the two is just barely noticeable. In some contexts, differences of an inch might be well above the threshold of discriminability, yet small enough to ignore for the conversational purposes at hand. But at least prima facie, in any given context, one thing counts as minimally taller than another only if the difference between the two is not below the threshold of discriminability for that context. (Though in special contexts, the threshold of discriminability might depend on apparatus external to our own perceptual systems.) And for these purposes, let's not worry about whether there are enough "things" to have a MINIMALLYTALLER-series for any two things we might want to compare.

If only for simplicity, let's suppose that given any thing that can be taller than something, we can inductively specify some things: the initial object; that object minus a contextually determined smidgeon ("off the top"); the previous object minus a contextually determined smidgeon; and so on, until we have an object such that no object is it minus a contextually determined smidgeon. This supposition, ugly though it is, makes presentation of the ideas below easier. If these ideas have merit, they can

19 Perhaps (15) is compatible with situations in which one or two players on the shorter team are taller than their positional opponents. But from a theoretical perspective, this may be "noise" due to genuine *group* comparisons. I hear (15) as clearly true given a one-to-one correspondence that pairs each of "they" with a shorter one of "them." But in the absence of such a correspondence, I am not sure that (15) itself is true, even they are "mostly" taller than them.

be recast in other terms. For example, whatever things are being compared, given any two of them x and y: if neither is taller than the other relative to the context at hand, then ¬MINIMALLYTALLER(x, y) & ¬MINIMALLYTALLER(y, x); and if x is taller than y, relative to the context at hand, then [**ANC**-MINIMALLYTALLER](x, y). This at least imposes substantive constraints on the relation MINIMALLYTALLER(x, y). Schwartzschild (2002) provides helpful and compatible discussion of scales, intervals, and their properties. But here, I'll operate with the idea of subtracting smidgeons, while recognizing the need for something better.

If some things, the Xs, exhibit a MINIMALLYTALLER-series, then presumably: one_x of $those_X$ things is such that it_x is not taller than any of $them_X$; and one_x of $those_X$ things is such that none of $them_X$ are taller than it_x. So it seems that we can talk, sensibly enough, about the "innermost" one of the Xs and the "outermost" one of the Xs. Thus, we might gloss (18) as (19).

(18) Carl is taller than Al
(19) ∃X[Outermost(X, Carl) & Taller(X) & Innermost(X, Al)]

The idea here would be that 'taller' is a plural predicate, satisfied by some thing*s* iff *they* exhibit a series of a certain sort. There is nothing especially new or exciting about this idea. My point is simply that the unexciting idea can be implemented, in second-order terms, without quantifying over things beyond things that *have* heights. Quantification over such things *and* their heights goes farther, at least if such quantification is supposed to reflect how speakers understand sentences like (18); cp. Klein (1980).

Given a (contextually determined) decision about what counts as *minimally* taller, and thus what counts as a smidgeon, we can effectively determine (up to vagueness) how *many* elements there have to be in a MINIMALLYTALLER-series that links Carl to Al. And one might hope to exploit this fact with regard to sentences like (20).

(20) Carl is much taller than Al

To be sure, 'much' is itself context sensitive, like 'many'. But (20) requires that Carl *not* be minimally taller than Al, nor even just a smidgeon or two more than minimally taller. So perhaps (20) is like (18), with the added condition that the Xs are many. Even if few "ordinary" things are taller than Al but shorter than Carl, the interveners may include many "Carl-minus-smidgeons."

Some such proposal might also help explain why (20) implies (18), *and* why ‘much taller’ is transitive, while ‘just barely taller’ is not. Perhaps even complicated examples like (21) can be dealt with in this way.

(21) Carl is so much taller than Bob than Bob is taller than Al

Note that (21) is true iff: some Xs exhibit a MINIMALLYTALLER-series with Carl and Bob as the outermost and innermost Xs; some Ys that exhibit a MINIMALLYTALLER-series with Bob and Al as the outermost and innermost Ys; and the Xs correspond *many-to-one* with the Ys.

Many analyses of (22) are compatible with the discussion here.

(22) Carl is tall

But one obvious starting point, suggested by the earlier discussion of (7),

(7) Adam is a big ant

would be to treat (22) like ‘Carl is a tall-~~one~~$_1$’, with the covert prominal understood contextually. More interestingly, in (23),

(23) Carl is six feet tall

‘tall’ apparently serves as a device that lets us express a relation between Carl and six feet. So maybe the logical form of (23) is, at one level of abstraction, ‘SixFeetTall(Carl)’; where SixFeetTall(x) iff x is *as tall as*—or perhaps iff x is *EquiTallWith* six (stacked) feet. Then (24)

(24) Carl is a foot taller than Al

might be analyzed along the following lines: there are some things$_X$ such that they$_X$ exhibit a (MINIMALLY)TALLER-series with Carl and Bob as the outermost and innermost of them$_X$, respectively, and *they*$_X$ measure a foot; where some things that exhibit a series can be *like* (the markings on) a ruler, which measures a foot by indicating—in accordance with a certain method of projection—points separated by twelve inches.

A series of individuals can also indicate points separated by twelve inches, given the right method of projection, if the last individual in the series is a foot taller than the first. If the top of Al’s head indicates a “zero” point, the top of Carl’s head can indicate a point that is one foot higher. Again, this is trivial. But as Frege's treatment of induction shows, representing certain truisms as such often requires second-order resources. And it is all too easy to *mis*characterize certain truisms by thinking exclusively in singularist/first-order terms, since this often leads to analyses that *distort* the thoughts in question by "paraphrasing" ordinary claims in terms of covert quantification over abstracta not implicated by the ordinary claims.

These are, to be sure, imprecise speculations as opposed a serious theory. But if second-order quantification and plural variables and ancestral relations are relevant in *other* constructions, it may be appropriate to think about comparative constructions in connection with these other constructions. Just as we have learned that theories of quantification and plurality can and should hang together, perhaps theories of comparatives and plural variables can and should hang together—and likewise for theories of quantification and number. In which case, we should assess theories in any particular subdomain with a clear view about what the logic underlying all of these subdomains does and does not require, in terms of quantification over abstracta. The demands of theory construction in semantics make it temping to appeal to sets, plural objects, numbers, heights, etc. But the history of the subject suggests that caution is always in order when making such appeals. Frege himself erred in this arena, by assuming more abstracta than logic requires or permits. That should make us mortals pause, especially since Frege's successes have opened up *various* ways of thinking about how his logic relates to natural language semantics and variables in natural languages.

4 Conclusion

Logical induction may be important for theoretical linguistics, even if children do not induce languages from experience. Either our human capacities for inductive reasoning lie near the heart of our capacity to generate and understand expressions of a human language, or not. If they do, then theoretically minded linguists should try to understand human inductive capacities and the kinds of understanding they make possible, independent of other cognitive capacities. If not, then we should be clear about this, and not pretend otherwise—say, by adopting semantic theories that exploit the full resources of the logic that Frege used to reduce arithmetic to Hume's Principle.

But suppose our best theories of language do presuppose that speakers have inductive capacities. Then considerations of theoretical parsimony suggest that we theorists should squeeze as much as we can from our representations of human inductive capacities, before adding controversial assumptions about how speakers understand expressions. This leaves room for hypotheses according to which speakers understand certain sentences in terms of covert quantification over abstracta. But when advancing such hypotheses, we should be cautious. And we should consider more than one way of interpreting our theoretical formalism. Here, as elsewhere, Frege provides a model from which we can still learn.

Acknowledgments
For helpful comments and discussion, my thanks to: Alexander Williams, Cedric Boeckx, Richard Heck, Norbert Hornstein, and Juan Uriagereka.

References

Beck, S., Oda, T., and Sugisaki, K. 2004. Parametric Variation in the Semantics of Comparison:Japanese vs. English. *Journal of East Asian Linguistics* 13: 289-344.

Benacerraf, P. 1965. What Numbers Could Not Be. *Philosophical Review* 74: 47-73.

Brody, M. 2000. Mirror Theory: Syntactic Representation in Perfect Syntax. *Linguistic Inquiry* 31: 29-56.

Boolos, G. 1975. On Second-Order Logic. *Journal of Philosophy* 72: 509-27.

Boolos, G. 1984. To be is to be the Value of a Variable (or the Values of Some Variables). *Journal of Philosophy* 81: 430-50.

Boolos, G. 1985. Nominalist Platonism. *Philosophical Review* 94: 327-44.

Boolos, G. 1987. A Curious Inference. *Journal of Philosophical Logic* 16:1-12.

Boolos, G. 1998. *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cresswell, M. 1976. The Semantics of Degree. In B. Partee (ed.), *Montague Grammar*. New York: Academic Press.

Demopolous, W. (ed.) 1994. *Frege's Philosophy of Mathematics*. Cambridge, MA: Harvard.

Fodor, J. 1970. Three Reasons for not Deriving 'Kill' from 'Cause to Die'. *Linguistic Inquiry* 1: 429-38.

Fodor, J. and Lepore, E. 1998. The Emptiness of the Lexicon. *Linguistic Inquiry* 29: 269-88.

Fodor, J. and Lepore, E. 2002. *The Compositionality Papers*. Oxford: Oxford University Press.

Frege, G. 1879. *Begriffsschrift*. Halle: Louis Nebert. English translation in J.van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical*

Logic, 1879-1931 (Cambridge, MA: Harvard University Press, 1967).
Frege, G. 1884. Die Grundlagen der Arithmetik. Breslau: Wilhelm Koebner. English translation in J. L. Austin (trans.), *The Foundations of Arithmetic* (Oxford: Basil Blackwell, 1974).
Frege, G. 1892. Function and Concept. In Geach and M. Black (trans.), *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell (1980).
Frege, G. 1893, 1903. Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet, 2 vols. Jena: Pohle. English translation in M. Furth (trans.), *The Basic Laws of Arithmetic* (Berkeley: University of California, 1967).
Gallistel, C.R. and Gelman, R. 1991. Subitizing: The preverbal counting process. In W. Kessen, A. Ortony, and F. CraiK (eds.), *Memories, Thoughts, and Emotions:essays in honor of George Mandler*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Gallistel, C.R. and Gelman, R. 2000. Non-verbal Numerical Cognition: From Reals to Integers. *Trends in Cognitive Science* 4: 59-65.
Gallistel, C.R., Gelman, R., and Cordes, S. 2005. The Cultural and Evolutionary History of the Real Numbers. In *Evolution and Culture*, S. Levinson and P. Jaisson (eds.), MIT Press.
Hart, H. and Honoré, A. 1959. *Causation and the Law*. Oxford: Oxford University Press.
Hazen, A. 1997a. Relations in Monadic Third-Order Logic. *Journal of Philosophical Logic* 26: 619-628.
Hazen, A. 1997b. Relations in Lewis's Framework without Atoms. *Analysis* 57: 243-8.
Hazen, A. 2000. Relations in Lewis's Framework without Atoms: a Correction. *Analysis* 60: 351-3.
Heck, R. 1993. The Development of Arithmetic in Frege's Grundgesetze der Arithmetik. *Journal of Symbolic Logic* 58:579-601. Reprinted with postscript in Demopolous (1994).
Heck, R. 1994. Definition by Induction in Frege's Grundgesetze der Arithmetik. In Dempolous (1994).
Heck, R. 1999. Frege's Theorem: An Introduction. *The Harvard Review of Philosophy* 7: 56-73.
Higginbotham, J. 1998. On Higher-Order Logic and Natural Language. *Proc. of the British Academy* 95: 1-27.
Hurford, J. 1987. *Language and Number*. Oxford: Basil Blackwell.
Keefe, R. and Smith, P. 1996. (eds.): *Vagueness: A Reader*. Cambridge, MA: MIT

Press.
Kennedy, C. 1999a. *Projecting the Adjective: The Semantics of Gradability and Comparison*. New York: Garland.
Kennedy, C. 1999b. Gradable Adjectives Denote Measure Functions, Not Partial Functions. *Studies in the Linguistic Sciences* 29.1.
Klein, E. 1980. A Semantics for Positive and Comparative Adjectives. *Linguistics and Philosophy* 4:1-45.
Lewis, D. 1991. *Parts of Classes*. Oxford: Blackwell.
Link, G. 1983. The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretic Approach. In R. Bäuerle, C. Schwarze, and A. von Stechow (eds.), *Meaning, Use, and Interpretation of Language* (Berlin, de Gruyter).
Linnebo, Ø. 2004. Plural Quantification. *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2004/entries/plural-quant/>.
Ludlow, P. 2002. LF and Natural Logic. In Preyer and Peters (2002).
Parsons, C. 1965. Frege's theory of number. In Max Black (ed.) Philosophy in America, Cornell University Press, 1965, pp. 180-203; reprinted, in Demopoulous (1994).
Pietroski, P. 1998. Actions, Adjuncts, and Agency. *Mind* 107: 73-111.
Pietroski, P. 2000. *Causing Actions*. Oxford: Oxford University Press.
Pietroski, P. 2003. Quantification and Second-Order Monadicity. *Philosophical Perspectives* 17: 259-98.
Pietroski, P. 2005. *Events and Semantic Architecture*. Oxford: Oxford University Press.
Pietroski, P. 2006. Interpreting Concatenation and Concatenates. *Philosophical Issues* 16:221-45.
Preyer, G. and G. Peters, G. (eds.) 2002. *Logical Form and Language*. Oxford, OUP.
Quine, W.V.O. 1950. *Methods of Logic*. New York: Henry Holt.
Rescher, N. 1962. Plurality Quantification. *Journal of Symbolic Logic* 27: 373-4.
Sainsbury, M. 1990. Concepts without Boundaries. King's College London, Inaugural lecture. Reprinted in Keefe and Smith (1996).
Scha, R. 1981. Distributive, Collective, and Cumulative Quantification. In J. Groenendijk *et.al.* (eds.), *Formal Methods in the Study of Language*. Amsterdam: Mathematisch Centrum.
Schwarzschild, R. 1996. *Pluralities*. Dordrecht: Kluwer.
Schwarzschild, R. 2002. The Grammar of Measurement. In B. Jackson (ed), *Proceedings of Semantics and Linguistic Theory XII*. Ithaca: CLC Publications, Cornell University.

Schein, B. 1993. *Plurals*. Cambridge, MA: MIT Press.
Schein, B. 2001. Adverbial, descriptive reciprocals. In R. Hastings *et. al.*, *Proceedings of Semantics and Linguistic Theory XI.* Ithaca: CLC Publications.
Schein, B. 2002. Events and the Semantic Content of Thematic Relations. In Preyer and Peters (2002).
Schein, B. 2006. Plurals. In E. Lepore and B. Smith (eds.), *The Oxford Handbook in the Philosophy of Language* (Oxford: Oxford University Press).
Schein, B. forthcoming. *Conjunction Reduction Redux*. Cambridge, MA: MIT Press.
Taylor, B. 1985. *Modes of Occurrence*. Oxford: Blackwell.
Wiese, H. 2003. *Numbers, Language, and the Human Mind*. Cambridge: Cambridge University Press.
Wiggins, D. 1980. 'Most' and 'All': Some Comments on a Familiar Programme and on the Logical Form of Quantified Sentences. In M. Platts (ed.), *Reference, Truth, and Reality*. London: Routledge and Kegan Paul.
Wright, C. 1983. *Frege's Conception of Numbers as Objects*. Scots Philosophical Monographs, vol 2. Aberdeen: Aberdeen University Press.
Yi, B. 1999. Is Two a Property? *Journal of Philosophy* 96: 163-90.
Zalta, E. (2003): Frege (logic, theorem, and foundations for arithmetic). *The Stanford Encyclopedia of Philosophy* (Fall 2003 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2003/entries/frege-logic.