How Much Logic is Built into Natural Language?

Ed Keenan
Linguistics, UCLA

10/18/05

Query Does knowing a natural language (English, Japanese, Swahili,...) imply knowing any logic?

The Query is reasonable (First Order) Predicate Logic ($PL_=$) is a "Universal Grammar" for the languages of Elementary Arithmetic, Euclidean Geometry, Set Theory, Boolean Algebra, ... It defines their expressions, their semantic interpretations, and texts, called proofs, that syntactically characterize the boolean semantic entailment relation: P entails Q iff Q is true whenever P is.

1 Properties of PL overtly present in Natural Language (NL)

1.1 Function Symbols (F1s, F2s,...) and Naming Expressions (F0s)

PL $+$ and $\times$ are F2s, squaring ^{2} is an F1: $2, 3, 3^2, (3^2+2), (3^2+2)^2, (3^2+2)^2+2, \ldots$

NL kin terms are F1's: *the dean, the mother of the dean, the mother of the mother of the dean...*
These are easier to understand if we vary the function expression: *the employer of the mother of the dean, the wife of the employer of the mother of the dean...*

Recursion = the values of a function lie in its domain, so its application iterates Not limited to possessive constructions. In childrens rhymes and songs:

Relative clauses *This is the house that Jack built, This is the malt that lay in the house that Jack built, This is the rat that ate the malt that lay in the house that Jack built...*

Prepositional phrases *There's a hole in the bottom of the sea, There's a log in the hole in the bottom of the sea, There's a bump on the log in the hole in the bottom of the sea...*

Compositionality meaning of a derived expression a function of those it is derived from: '$(2+3)$' denotes the value of the function denoted by '$+$' at the numbers denoted by '2','3'.

A Fundamental Similarity PL and NL are recursive, compositional systems. They build infinitely many non-synonymous expressions from a small finite word list (See Section 5)

- Leading Question of Md Linguistics: Account for how we produce and understand arbitrarily many novel expressions in NL. Recursion + Compositionality a partial answer
- Lakoff and Núñez (2000): repetition is the linguistic basis of recursion.
  Keenan: Recursion is self-application. Dripping faucets are not recursive.
- Recursion (self application) is a "statistical accident." Most functions don't iterate:
  The height of the dean, #the height of the height of the dean, ...

Two place functions are less naturally typified in NLs: *The taller of Bo and Joe, The taller of Sam and [the taller of Bo and Joe]...* So this F2 iterates, as do *the offspring of Joe and Flo, the canal between my house and yours.* Other F2s don't iterate: *the distance from* x *to* y, **the distance from* x *to the distance from* y *to* z

1.2 Predicate-Argument Formulas (FMs) /Sentences (Ss)

PL Simple FMs = Predicate + Names. '$2 > 1$', $2^2 = (3+1)$.

NL abundant: P1s $\approx$ *sleeps, laughs;* P2s $\approx$ *praises, describes;* P3s $\approx$ *gives, hands;*

- Arguments are often (not always) **asymmetrically** related: In PL $2 > 1$ and $1 > 2$ both make sense (but differ in truth value). *I wrote that poem* is natural, *That poem wrote me* is nonsense. The first argument of *write* is the **Agent**, the second its **Patient**.
- The second argument of a P2 may be referentially bound to the first, but not conversely:

  Ben washed/punished himself — *Himself (Heself) washed/punished Ben

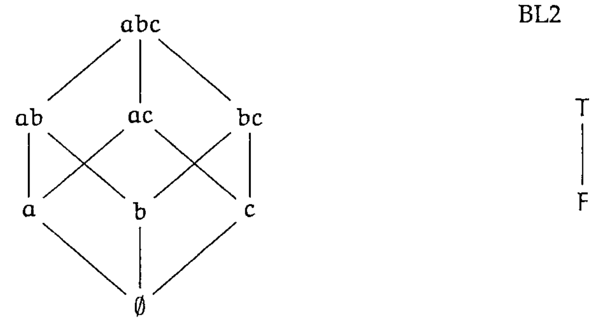
- The existence of a referent of the second argument of a P2 may be conceptually dependent on the action expressed by the P2: In *The boy set a fire*—the existence of the fire (not the boy) depends on the act of setting. Similarly: *He made a mistake / committed a crime.* The first argument of a P2 is never dependent on it in this way. It may be with a P1: *A light breeze sprang up, A fire broke out, An accident happened*
- P2s in NL may fail to be isomorphic. *Ben washed the car* passivizes to *The car was washed by Ben.* But *Ben has a car* does not passivize: **The car is had by Ben.* A binary relation in PL always has a converse.

1.3 Boolean operations

In PL *and, or,* and *not* build FMs from FMs; they denote boolean functions: *and* is a binary greatest lower bound (glb) operator, noted $x \wedge y$; *or* a binary least upper bound (lub) operator, noted $x \vee y$, *not* is a complement operator, noted $\neg x$.

Boolean Lattices (B,$\leq$) are distributive, complemented lattices. $\leq$ is a boolean partial order:

BL8 BL2



Here $x \leq y$ iff $x = y$ or you can move up along edges from x to y. $x \wedge y$ is the "highest" element which is $\leq x$ and $\leq y$. $x \vee y$ is the lowest element that both x and y are $\leq$ to. BL2 is the boolean lattice of truth values, which FMs denote. Writing $TV(\phi)$ for the truth value of ϕ, we have $TV(\phi \& \psi) = TV(\phi) \wedge TV(\psi)$, $TV(\phi \text{ or } \psi) = TV(\phi) \vee TV(\psi)$ and $TV(\text{not } \phi) = \neg(TV(\phi))$.

Quantification in PL $Qx\phi$ is a FM, where ϕ a FM, x a variable and Q is either the universal quantifier all, noted $\forall$, or the existential quantifier some/there exists, noted $\exists$. For example:

$\forall x(x^2 \geq x)$	'For any number x, its square is greater than or equal to x'
$\exists x(Even(x) \& Prime(x))$	'There is a number x which is both even and prime'

Semantically $\forall$ an arbitrary glb operator, and $\exists$ an arbitrary lub operator. E.g. $TV(\forall x(x^2 \geq x)) = TV(0^2 \geq 0) \wedge TV(1^2 \geq 1) \wedge (2^2 \geq 2) \wedge \ldots$ Writing $TV(\phi[^x/_b])$ for the truth value of ϕ when the variable x is set to denote b, we see that $\forall$ = "AND writ big"; $\exists$ = "OR writ big."

$TV(\forall x\phi) = \bigwedge\{TV(\phi[^x/_b]) \mid b \in E\}$ $TV(\exists x\phi) = \bigvee\{TV(\phi[^x/_b]) \mid b \in E\}$

PL ties variable binding to quantification. It is enlightening to separate them, as in $ALL(x.\phi)$ where $(x.\phi)$ is a P1 built from a FM (P0) by prefixing the variable x. Then Qs combine directly with P1s to form FMs (P0s). (Habitués can read $(x.\phi)$ as $(\lambda x.\phi)$.

NL Negation Present in all NLs (Dryer 2005), often with categories other than Sentence.
Bill *isn't* a linguist; *Not* more than six students laughed; *Not* a creature was stirring

Universal Quantification Present in all NLs (knowledgeable conjecture, kc; Gil 2005). *All* cats are black in the dark; The students have *all* left for summer break;

Existential Quantification All NLs may assert and deny existence (kc):
Yesh / Ein yeladim ba gan 'There are / (aren't any) children in the park' (Hebrew)
Misy / Tsy misy totozy eo 'There are / (aren't any) mice there (Malagasy)

Conditional Ss (Traugott 1985; Khrakovskii 1998) present in all NLs (kc). If it doesn't rain tomorrow the crops will die; Fred should be punished if he's guilty

'and', 'or' and 'neither ... nor ...' are ubiquitous in NL; in PL, they only connect FMs;

Neither John nor Bill / Either John or Bill came to the party early	NP
He neither laughed nor cried / either laughed or cried	P1
He lives neither in nor near / either in or near New York City	Preposition
He neither praised nor criticized / either praised or criticized each student	P2
He is neither industrious nor intelligent / either industrious or intelligent	Adjective
Neither did John come early nor did Fred leave late	S (P0)
Either John came early or Fred left late	

Gen 1 *The set in which expressions of a category C denote is a boolean lattice (B, $\leq$), supporting that the boolean operations are "properties of mind" (Boole 1854).*

Gen 2 *Modifiers are usually restricting:* tall student $\leq$ student, *that is, all tall students are students, all skillful doctors are doctors, etc.*

Variable usage in NL: *and* may = *and then*, as in *Flo got married and got pregnant* $\neq$ *Flo got pregnant and got married*, or *and as a result*, as in *John drank too much and got sick*. But not always: *Flo is 6 feet tall and studies biology* = *Flo studies biology and is 6 feet tall.* Usage in logic abstracts from this variation to yield (P&Q) is true iff P is and Q is, whence the semantic symmetry: P&Q = Q&P. Similarly with *or*, which is sometimes intended as exclusive, as in *John either laughed or cried* (? but not both). But not always: *Do you have two nickels or a dime?* must be answered 'Yes' if you have both two nickels and a dime.

2 Logical Properties covertly present in NL (Whorf-Sapir Hypothesis)

Knowing English implies knowing the distribution of NPI's (negative polarity items)—e.g., *ever* and *any*, whose presence is licensed by overt negation, as in (1), but also by certain NPs in subject position, as in (2):

(1) a. John hasn't ever been to Pinsk a'. John didn't see any birds on the walk
b.*John has ever been to Pinsk b'.*John saw any birds on the walk.

(2) a. No student here has ever been to Pinsk a'. Neither John nor Mary knew any Russian
b.*Some student here has ever been to Pinsk b'.*Either John or Mary knew any Russian.
c. Fewer than five / *More than five students here have ever been to Pinsk
d. At most / *At least two students here have ever been to Pinsk

query Which NPs license NPI's? What if anything do they have in common with negation?

Gen 3 *NPI licensers are expressions which denote monotone* ***decreasing*** *functions*

Def Let $(A,\leq)$ and $(B,\leq)$ be posets (e.g. boolean lattices), F a function from A into B. Then

a. F is *increasing* (order preserving, isotone) iff for all $x, y \in A$, if $x \leq y$ then $F(x) \leq F(y)$.

b. F is *decreasing* (order reversing, antitone) iff for all $x, y \in A$, if $x \leq y$ then $F(y) \leq F(x)$.

Test for Increasingness (↑): if all Ps are Qs and X is a P, therefore X is a Q. Ex: 'some poet' is ↑: Suppose all Londoners drink stout and some poet is a Londoner. Therefore some poet drinks stout.

Gen 4 *Virtually all syntactically underived NPs are ↑: Proper Names (*Ned, Gail*), pronouns (*he, she, they*), demonstratives (*this, those*).*

Gen 5 *The closure of Proper Name denotations under the (complete) boolean operations is the denotation set for all quantified NPs (*No/Most/All students,...*). In fact they are a set of* ***free*** *generators for that set—so they determine its boolean structure uniquely.*

Test for Decreasingness (↓): All Ps are Qs and X is a Q, therefore X is a P.
'No poet' is ↓: if all Londoners drink stout but no poet drinks stout then no poet is a Londoner
Negation is ↓: if Londoner → drinking stout then not drinking stout → not being a Londoner

Gen 6 *The major ways of building NPs from NPs preserve or reverse monotonicity:*

a. *Conjunctions and disjunctions of ↑ NPs are ↑; analogously for ↓ NPs.*

b. *Possessive NPs have the monotonicity value of the possessor: X's doctor is ↑(↓) if X is.*

c. *Negation reverses monotonicity: not more than two boys is ↓ since more than two boys is ↑*

query Which NPs occur naturally in the post-*of* position of partitives, as in Two of _?
yes: Two of *those students*, two of *John's/the ten/John's ten/my students*
no: *two of *most students*, *two of *no students*, *two of *more male than female students*

Gen 7 *Post-*of *DPs of the form Det + Noun denote proper principal filters (=for some* $p > 0$, $F(q) =$ *True iff* $p \leq q$*). Linguists call the Dets which build such NPs* definite

query Which NPs occur naturally in Existential There (ET) contexts, as in:
Aren't there *at most four undergraduate students* in your logic class
Weren't there *more students than teachers arrested* at the demonstration?
Just *how many students* were there at the party?
Aren't there *as many male as female students* in the class?
There was *no student but John* in the building
Wasn't there *only one student besides John* at the lecture?
There are *most students* in my logic class
*Isn't there *the student who objects to that*?
*Isn't there *every student who gave a talk at the conference*?
*Was there *neither student* arrested at the demonstration?

Gen 8 *Just NPs built from intersective Dets and their boolean compounds (modulo pragmatic factors, partially excluded by negative or interrogative contexts) occur in ET contexts.*

Intersective (Generalized Existential) Dets are ones whose values at a pair A,B of properties just depends on $A \cap B$. Formally, they satisfy the invariance condition below:

(3) $D(A)(B) = D(X)(Y)$ whenever $A \cap B = X \cap Y$.

some intersective Dets some, a/an, no, several, more than six, at least / exactly / fewer than / at most six, between six and ten, just finitely many, infinitely many, about / nearly / approximately a hundred, a couple of dozen, practically no, not more than ten, at least two and not more than ten, either fewer than five or else more than twenty, that many, How many?, Which?, more male than female, just as many male as female, no... but John

Co-intersective Dets *every, all but two,...* which satisfy (3) with $-$ for $\cap$, are not intersective. Nor are proportionality Dets, such as most, less than half, seven out of ten, 10% of the, ...

3 Properties of PL not present in NL

Precision NL, not PL, is structurally ambiguous

1. John didn't leave because the children were crying
   R1: That's why he stayed [not leave][because the children were crying]
   R2. He left for some other reason [not [leave because the children were crying]]
   Compare in PL: $\neg(P \rightarrow Q)$ versus $(\neg P \rightarrow Q)$
2. Every student read a Shakespeare play (over the vacation)
   R1: For every student there was a play he read—maybe different students reaD different plays
   R2: There was one Sh. play that every student read (maybe Hamlet, maybe Lear,...)
   Compare in PL: $\forall x \exists y(x < y)$ vs $\exists y \forall x(x < y)$ They have different truth values
3. John told Bill that he had the flu. John said: "I have the flu", "You have the flu", or Henry (identified in context) has the flu. Compare: john_x(x told bill that x had the flu), bill_y(john told y that y had the flu), john told bill that z had the flu.
4. John thinks he's clever and so does Bill [think that John is clever, think that he himself is clever]
   John_x(x think x is clever & Bill think that x is clever)
   John_x(x think x is clever) & Bill_y(y think that y is clever)

PL is transparent, NL is not (Fine 1992, Kracht 2003) *The clown laughed* is an expression of English. It is a subsequence *and* a meaningful part of *They believed that **the clown laughed***; but in *The children who watched **the clown laughed*** it is subsequence but not a meaningful part. So English is not transparent.

4 Logical resources of NLs not present in PL

NL quantifiers map pairs of properties to truth values, the first restricting the domain of quantification, as in *Most **poets** daydream.* PL quantifiers have just one property argument:

a. Some poets daydream = $\exists x(P(x) \& D(x)) \equiv \text{SOME}(\lambda x(P(x) \& D(x))$

b. All poets daydream = $\forall x(P(x) \rightarrow D(x)) \equiv \text{ALL}(\lambda x(P(x) \rightarrow D(x))$ $[(P \rightarrow Q) = (\neg P \text{ or } Q)]$

Theorem (Keenan 1992) The domain eliminable NL quantifiers are just the (co)-intersective ones (generalized universal and generalized existential)

Gen 9 *All PL quantifiers are domain reducible; not so in NL. E.g. Proportionality Quantifiers*

Def If Det is proportional then the truth of *Det poets daydream* depends on the proportion of poets that daydream. ($DAB = DXY$ whenever $|A \cap B| / |A| = |X \cap Y| / |X|$)

Examples: most, seven out of ten, less than half, not one... in ten
Most poets daydream does not mean either (For most objects x (Poet(x) & Daydream(x)) or (For most objects x, if Poet(x) then Daydream(x)). BUT

Gen 10 *a. NL Quantifiers are domain independent: Blik defined by* $\mathrm{BLIK(A)(B)} = T$ *iff* $|\neg A| = 2$ *is not a possible English determiner.* Blik cats are black *would be true iff the number of non-cats is two.*

b. NL Qs are overwhelmingly ***conservative****: Det As are Bs cannot depend on Bs which are not As, so* $\mathrm{DAB} = \mathrm{D(A)(A \cap B)}$ *NB: Conservativity (CONS) and Domain Independence (DI) are independent. (BLIK is CONS but not DI;* F *in* $\mathrm{FAB} = T$ *iff* $|A| = |B|$ *is DI but not CONS)*

Gen 11 *Proportionality Quantifiers determine reasoning paradigms not present in PL*

- *Exactly half the students passed. Therefore, Exactly half the students didn't pass*
- *Between a third and two thirds of the students passed the exam.*
  Therefore, between a third and two thirds of the students didn't pass the exam.
- $Qx\phi$ *never entails* $Qx\neg\phi$*, for* Q *='all' or 'some'*

Gen 12 *Non-trivial Proportionality quantifiers are "logical" (= their denotations are permutation invariant) but* ***not definable*** *in PL. Similarly with cardinal comparatives, of type* $((1,1),1)$*:*
More *poets* ***than*** *priests daydream;*
Fewer *boys* ***than*** *girls,* ***More than twice as many*** *girls* ***as*** *boys;*
Half again as many *girls* ***as*** *boys*
These quantifiers may have multiple occurrences:
Fewer *boys* ***than*** *girls read* ***more*** *poems* ***than*** *plays.*
And in different form,
Jack read ***more poems than Jill***
A certain number of *students applied for* ***a smaller number*** *of scholarships*

Gen 13 *PL quantifiers are* ***extensional****, NL ones may not be. In a situation in which the doctors and the lawyers are the same individuals,* Every doctor attended (the meeting) *and* every lawyer attended... *have the same truth value, but* Not enough doctors attended *and* not enough lawyers attended *may have different values. All PL quantifiers are like* every *here.* too many, surprisingly many, ... *are like* not enough.

Gen 14 *In PL multiple quantification into different arguments of a predicate is just iterated quantification of type (1) quantifiers, (4) But NL presents unreducible type <2> quantifiers, (5):*

(4) Every student read at least one poem = $\forall x(S(x) \rightarrow \exists y(Py \& xRy))$

(5) **Different people** like **different things**
Jack danced with Jill but **no one else** danced with **anyone else**
Which students read **which plays?**
Jack and Jill support **rival political parties** / **the same candidates** / **each other**

5 Intuitions of NL "Structure"

Isomorphic expressions are ones treated the same by the structure building functions. And, via **Compositionality** they are meaningful in the same way. This often enables us to tell when two expressions differ in structure, despite appearing similar.

1. The (x, x') P1s below seem structurally similar, but are meaningful in different ways. In (a)

   a. That fish is good to eat — a′. That child is reluctant to eat
   b. That teacher is easy to please — b′. That teacher is eager to please

   we talk about eating the fish, but in (a′) the child is the one who is eating (or not) and the objects to be eaten aren't specified. And, as expected, they are structurally dissimilar. For example: the a,b cases can modify an N by framing it, but not the a′,b′ ones.

   a. a good fish to eat — a′. *a reluctant child to eat (a child who is reluctant to eat)
   b. an easy teacher to please — b′. *an eager teacher to please

2. Consider:

   (a) The juror expected to punish him / himself

   (b) The judge who the juror expected to punish him / himself

   In (a) himself = the juror, him = someone else; in (b) the opposite pattern obtains.

3. Ambiguity (= has two structures): disambiguated, some operations just apply to one of them. Ex: 'flying planes' = 'planes which are flying' or = 'the act of flying planes'

   AGREEMENT distinguishes them *Flying planes **is** dangerous* vs *Flying planes **are** dangerous* Modals (can, should,...) neutralize number agreement: *Flying planes can be dangerous* is ambiguous.

References

Boole, G. 1954. *An Investigation of the Laws of Thought.* Macmillian & Co. Cambridge

Dryer, M. 2005. Negative morphemes. *World Atlas of Language Structures.* Haspelmath et al

Fine, Kit. 1992. Transparency Part I: Reduction. Ms. UCLA

Gil, D. 2005. Conjunction and universal quantifiers. In Haspelmath et al.

Haspelmath, M., M. Dryer, D. Gil, and B. Comrie (eds). 2005. World Atlas of Language Structures OUP.

Keenan, E.L.1992. Beyond the Frege Boundary *L&P* 15:199–221

Keenan, E.L. and E.P. Stabler. 2003. *Bare Grammar* CSLI, Stanford

Khrakovskii, V. S. (ed). 1998. Tipologiia uslovnyhkh konstruktsii. Nauka Saint-Petersburg.

Kracht, Marcus. 2003. *The Mathematics of Language* Mouton de Gruyter

Lakoff, G. and R. Núñez 2000. *Where Mathematics Comes From* Basic Books

Traugott, E.C. 1985. Conditional Markers. In *Iconicity in Syntax* J.Haiman (ed) Benjamins.