From Natural Logic to Natural Reasoning*

Lauri Karttunen

Stanford University

Abstract. This paper starts with a brief history of Natural Logic from its origins to the most recent work on implicatives. It then describes on-going attempts to represent the meanings of so-called 'evaluative adjectives' in these terms based on what linguists have traditionally assumed about constructions such as *NP was stupid to VP*, *NP was not lucky to VP* that have been described as factive. It turns out that the account cannot be based solely on lexical classification as the existing framework of Natural Logic assumes.

The conclusion we draw from this ongoing work is that Natural Logic of the classical type must be grounded in a more inclusive theory of Natural Reasoning that takes into account pragmatic factors in the context of use such as the assumed relation between the evaluative adjective and even the perceived communicative intent of the speaker.

1 What Is Natural Logic?

Natural Logic attempts to do formal reasoning in natural language making use of syntactic structure and the semantic properties of lexical items and constructions. It contrasts with approaches that involve a translation from a natural to a formal language such as predicate calculus or a higher-order logic.

Figure 1 sketches the history of Natural Logic as told by Johan van Benthem in [3] and in lectures.

The short version goes as follows. Natural Logic has been around over 2000 years. It started out pretty well with Aristotle and the Greeks who invented syllogisms, some two dozen valid patterns of inference in ancient Greek. In the medieval times all of this was ported into Latin and extended by people like William of Ockham and Buridan. With the waning of the Middle Ages began a decline in logic that bottoms out in the works of De Morgan in the middle of the 19th century. But soon came the rise of modern logic first with Gottlob Frege in the 1890s and on the Natural Logic side with Charles Sanders Peirce about the same time. The current revival in Natural Logic was started by Johan van Benthem [2] and his student Víctor Sánchez-Valencia [26] in the 1990s. Among the latest advances is the work by Jan van Eijck [8], Bill MacCartney [20] and the recent papers by Thomas Icard [10] and Larry Moss [11] that build on the work of MacCartney and Christopher Manning [21].

* The original work reported in this paper is part of a joint project with Cleo Condoravdi, Stanley Peters, and Annie Zaenen at the Center for the Study of Language and Information. Special thanks to Annie Zaenen for the content and form of this paper.

A. Gelbukh (Ed.): CICLing 2015, Part I, LNCS 9041, pp. 295–309, 2015.
DOI: 10.1007/978-3-319-18111-0_23

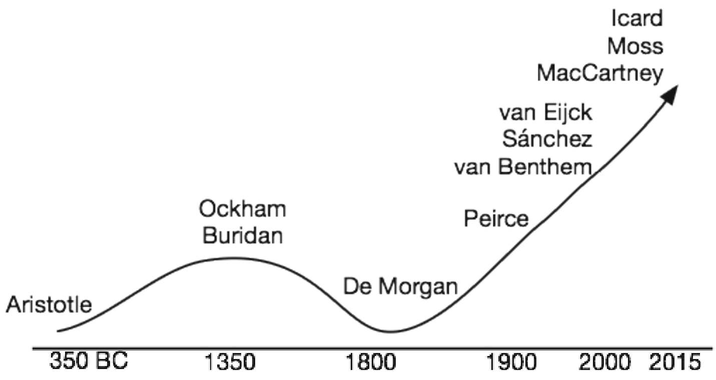

Fig. 1. A Brief History of Natural Logic

Augustus De Morgan is famous for De Morgan Laws:

The negation of a conjunction is the disjunction of two negations:
$\neg(p \wedge q) \equiv \neg p \vee \neg q$
The negation of a disjunction is the conjunction of two negations:
$\neg(p \vee q) \equiv \neg p \wedge \neg q$

Why does the does the curve of Natural Logic fall to its lowest point at his time? As van Benthem and Sánchez-Valencia point out, the laws were already clearly articulated by medieval logicians such as Ockham and Buridan. De Morgan's sole contribution is the formulation we now use. But more importantly, De Morgan was unsuccessful in his attempts to give a formal explanation of the validity and invalidity of some simple examples that medieval logicians had succeeded in explaining. For example, since horses are animals, it is obvious that (1a) is true, but since some other animals also have tails (1b) is false in our world.

(1) a. Every tail of a horse is a tail of an animal.
  b. Every tail of an animal is a tail of a horse.

The difficulty in explaining the obvious in logical terms is this. Assume that we start with a tautology such as (2) that contains two instances of the word *horse*.

(2) Every tail of a horse is a tail of a horse.

The substitution of a more general term, *animal*, for the second occurrence of *horse* is a valid inference yielding (1a). But the substitution of *animal* for the first occurrence of *horse* is an invalid inference resulting in (1b).

Suppose we start with the tautology in (3).

(3) Every tail of an animal is a tail of an animal.

Replacing *animal* with the more specific term *horse* is valid in the first instance but invalid for the second one.

Sánchez-Valencia and van Benthem report that medieval logicians, such as William of Ockham,[1] did have a right solution to the puzzle although it was wrapped up in a complex theory of *Suppositions*. One reason why the problem was difficult for them was that there was no theory to describe the syntactic structure of the Latin analogues of (2) and (3).

De Morgan thought that his rules of inference validated the correct inferences from (2) and (3) to (1a) but Sánchez-Valencia shows that they also allow the invalid inferences that yield (1b). That was the low point of Natural Logic.

The first logician in modern times who gave the right answer to the puzzle was the American Charles Sanders Peirce late in the 19th century. His system is also rather complicated but it is based on the right idea. The validity of substituting a general term like *animal* for a specific one like *horse*, or vice versa, depends on the position of the target word in the syntactic structure of the sentence.

A term that occurs in a **downward monotonic** (= **antitone**) context as the first occurrence of *animal* in in (3) can be replaced by a more specific term like *horse* as in (1a). A term that occurs in an **upward monotonic** (= **monotone**) context as the second occurrence of *horse* in (2) can be replaced by a more general term such as *animal*.

The truth of (1a) and the falsity of (1b) are obvious to any speaker of English but it is not trivial for a beginning student of logic to prove (1a) expressed in first-order logic starting with (2) and the premise that horses are animals. A proof by Natural Deduction or Sequent Calculus is a substantive homework assignment. It takes many lines of reasoning in a formal language to show the validity of such simple monotonicity inferences in ordinary English.

2 Monotonicity

The distinction between ‘more specific’ and ‘more general’ terms does not only apply to nouns like *horse* and *animal*. It is applicable to expressions of any syntactic category. If X and Y are expressions of he same syntactic type, we say that X is more specific than Y if all instances of X are instances of Y but not vice versa. In the notation introduced by Bill MacCartney [20] we write this $X \sqsubset Y$ where $\sqsubset$ is a symbol for inclusion, a generalized entailment relation. For example, Figure 2 illustrates the fact that *with* is an upward monotonic operator, $\uparrow$, and *without* a downward monotonic operator, $\downarrow$.

Figure 2 shows graphically that any action done with a knife is included in actions done with a tool: *with a knife* $\sqsubset$ *with a cutter* $\sqsubset$ *with a tool*. The preposition *without* reverses these inclusion relations: *without a tool* $\sqsupset$ *without a cutter* $\sqsupset$ *without a knife*.

Table 1 codes the monotonicity properties for some English determiners as they are traditionally described.[2] For example, as we have already seen in (2) and (3), *every* creates a downward monotonic context for its first argument, a nominal phrase, and an upward monotonic context for the second argument, a verb phrase.
Such a table is however misleading in that it suggests incorrectly that the upward/downward monotonicity can be determined locally. In fact the $\downarrow$ marks in Table 1 should be

[1] Ockham was also a pioneer of three-valued logic.

[2] Some determiners do not have any monotonicity effects. *Five* is $\uparrow\uparrow$ but *exactly five* is ==, that is, it yields no monotonicity inferences on either of its two arguments.

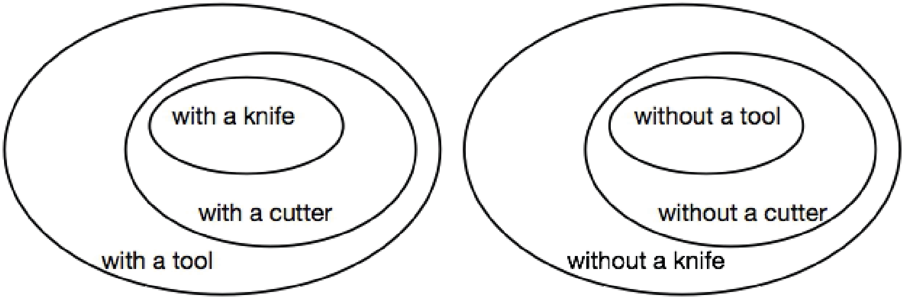

Fig. 2. With↑ NP vs. without↓ NP

Table 1. Classical monotonicity signatures for some determiners

Det	Code	Example
every	↓↑	Every house was damaged in a fire. ⊏ Every small house was damaged.
some	↑↑	Some small house was damaged in a fire. ⊏ Some house was damaged.
no	↓↓	No house was damaged. ⊏ No small house was damaged in a fire.
few	=↓	Few students have a car. ⊏ Few students have a fancy car.
many	=↑	Many professors have a fancy car. ⊏ Many professors have a car.

changed to ↑, and vice versa, if the expression appears in a negative context. If we put the construction *some NP VPs* under negation as in (4a), the valid inference pattern for *some* turns into the same as for *no* in Table 1. That is, (4a) entails (4b).

(4) a. It is not the case that some house was damaged.
b. No small house was damaged in a fire.

The same holds for *every*. In (5a) *student* is in a downward monotonic context and *cheap car* is in an upward monotonic context. Consequently, we can replace *student* by the more specific *poor student* and *cheap car* by the more general *car*. (5a) entails (5b).

(5) a. Every student has a cheap car.
b. Every poor student has a car.

But everything flips if we replace *every* by *not every*. (6a) entails (6b).

(6) a. Not every poor student has a car.
b. Not every student has a cheap car.

In (6) the first nominal argument of *every* is in an upward monotonic context that licenses the replacement of the specific term *poor student* with the more general *student*. In contrast, the phrase *has a car* in (6) is a downward entailing context that justifies replacing *car* by the more specific *cheap car*.

These complications brought in by negation have, of course, always been known to logicians and they are probably one reason why the medieval logicians, not having a syntactic structure to build on, came up with the hard-to-understand theory of suppositions. The modern approaches to the problem starting with van Benthem and Sánchez-Valencia intially took their cue from Peirce's work and set up a two-step monotonicity calculus.

In the first step the nodes in the parse tree are marked with + or - signs using the lexical signatures for determiners in Table 1 and other functors such as *with*, *without* and *not*. The result for the example (6a) is shown schematically in Figure 3 (a). Since *not* is a downward monotonic operator it's sentential argument gets a minus sign.

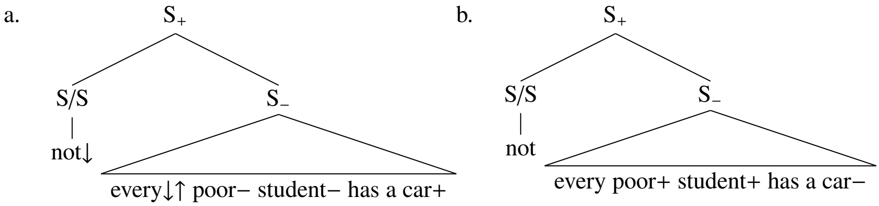

Fig. 3. Two step computation of monotonicity

The second step of the van Benthem-Sánchez algorithm traces the paths from the leaves of the parse tree to the root counting the minus signs on the path. If the number of minuses is even, the final sign is + to indicate an upward monotonic context; if the number of minuses is odd, the node is marked with - to show that it is a downward monotonic phrase. Figure 3 (b) shows that the effect in this case is to reverse the initial assignments of signs. The final marking justifies the inference from (6a) to (6b).

The two step-method of computing monotonicity is unnecessarily convoluted. David Dowty [6] describes a system that derives the marking directly in a categorial grammar. Unfortunately his bottom-up method necessitates the duplication of many lexical categories. Van Eijck [8] describes a simple top-down algorithm that computes the desired result in a single pass always starting with a positive sign on the highest node of the parse tree. That is an optimal solution to the problem.

3 Beyond Monotonicity

Historically studies of monotonicity tend to focus on the semantics of determiners and quantifier phrases. But monotonicity inferences arise with many areas of language that semanticist only recently have begun to describe such as the meaning of comparative constructions [25].

The dramatic rise in our understanding of this type of reasoning pictured in Figure 1 is not an exaggeration. Natural Logic has advanced more in the last few decades than at any time since Greek philosophers. Aristotle's 24 classical syllogisms, valid patterns of reasoning, are now understood in terms of the monotonicity properties of a few determiners and the axioms of symmetry and existential import (no empty classes) [7].

Because we now have better theories of syntax than any previous generations, we are not at all baffled by the *tail of a horse* puzzles that occupied previous generations of semanticists for centuries.

3.1 Implicatives

In this section we will focus on studies that extend the classical scope of Natural Logic from simple sentences to constructions with infinitival clauses and embedded sentences. The discussion is based on Lauri Karttunen's descriptive work on implicative constructions [15,16] and its computational implementation by Nairn et al. [22] and MacCartney [20]. The question is whether the proposition implicit in an infinitival clause is presented as true, false, or not entailed either way.

A good place to start is the example (7) in MacCartney and Manning [21]:

(7) a. James Dean refused to move without blue jeans.
b. Dean didn't dance without trousers.

Obviously—or upon reflexion at least—(7a) entails (7b). Because *without* is a downward entailing operator we would expect the entailment *without trousers* $\sqsubset$ *withouth blue jeans*. But here the entailment goes in the opposite direction, from he more specific term *blue jeans* to the more general category *trousers*. In positive contexts *dance* $\sqsubset$ *move* but in (7) the relationship is reversed because of the negative implication of *refused*.

The construction *refuse to VP* is one of the several types of implicative patterns discussed in [15] and [16]. They include two classes of **two-way implicatives** and four types of **one-way implicatives**. Table 2 contains a few examples of the first kind.[3]

Table 2. Two types of two-way implicative verbs

+ + | − − implicatives	+ − | − + implicatives
manage to	fail to
bother to	neglect to
remember to	forget to
see fit to	refrain from . . . ing
happen to	avoid . . . ing

The + + | − − implicatives are constructions that in a positive context entail the truth of the infinitival clause (++). In negative contexts they entail that the infinitival clause is false (−−). Examples in (8).[4]

[3] The examples in this section involve simple verbs. See [16] for a discussion of phrasal implicatives such as *take the time/opportunity/trouble to VP*.

[4] In addition to their entailment properties all the constructions in Table 2 suggest something else as well. For example, *remember to VP* and *forget to VP* both imply that the protagonist intended or was expected to VP. That can lead to arguments such as *I didn't **forget** to go to the party. I never intended to go there* that are not about what happened but about whether forgetting was involved.

(8) a. The culprit managed to get away. ⊏ The culprit got away.
b. She didn't bother to explain. ⊏ She didn't explain.
c. He saw fit to ask her for another chance. ⊏ He asked her for another chance.
d. Kim didn't remember to have breakfast. ⊏ Kim didn't have breakfast.

The + − | − + implicatives also yield an entailment both in positive and negative context but they reverse the polarity.

(9) a. He had failed to get into Oxford. ⊏ He had not gotten into Oxford.
b. She didn't avoid getting caught. ⊏ She got caught.
c. He didn't neglect to return her call. ⊏ He returned her call.
d. Kim forgot to have breakfast. ⊏ Kim didn't have breakfast.

Constructions such as *manage to VP* and *fail to VP* are perfectly symmetrical in that they yield an entailment both in affirmative and negative contexts. There are four types of verbs that yield an entailment about their complement clause only under one or the other polarity.

Table 3. Four types of one-way implicative verbs

++ implicatives	+− implicatives	−− implicatives	−+ implicatives
cause NP to	refuse to	can	hesitate to
force NP to	be unable to	be able to	
make NP to	prevent NP from		

The examples in (10) illustrate these one-way implicative constructions. In all cases, reversing the polarity does away with any logical entailment unlike the examples of two-way implicatives in (8) and (9).

(10) a. They forced the crowd to disperse. ⊏ The crowd dispersed.
b. Dean refused to move. ⊏ Dean didn't move.
c. He was not able to get up. ⊏ He did not get up.
d. She didn't hesitate to help him. ⊏ She helped him.

If a person says *I was able to log in* one is inclined to conclude that she did, and that may well be what the speaker intends to convey. However, it is not a contradiction for her to continue *but I did not do it*.[5]

The computation of inferences from implicative verbs has been implemented, [22], [20], as the same kind of top-down process as van Eijck's method of computing monotonicity.

Figure 4 traces the assignment of polarity marks in structure containing three stacked two-way implicatives: *not*, *fail* and *remember*. Starting with positive polarity on the top

[5] An attested example of this type: *He was able to sin, but did not; he was able to do wrong, but would not.* `http://www.liturgies.net/saints/paulinusofnola/readings.htm`. Replacing *was able* by *managed* would create a contradiction.

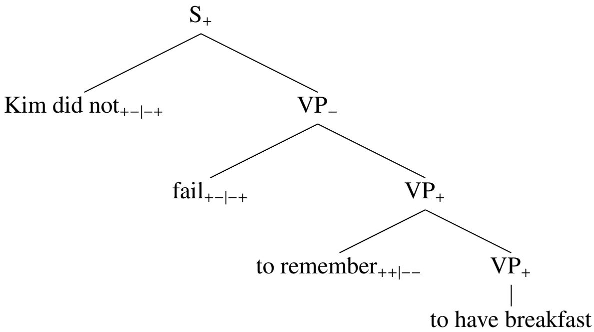

Fig. 4. Kim did not fail to remember to have breakfast $\sqsubset$ Kim had breakfast

node, the chain of VPs gets marked with + or - determined by the lexical signature of *not* or the higher verb and the polarity passed onto that clause from above.

This example entails the innermost clause, *Kim had breakfast*, because *not* and *fail* reverse the incoming polarity and *remember* preserves it. Replacing *fail* by *happen* would result in the entailment that Kim did not have breakfast. If the implicative chain is broken, say by replacing *remember* by a non-implicative verb such as *propose*, there would be no entailment about whether anyone had breakfast.

The computations with implicative verbs are a natural extension of the monotonicity calculus discussed in the previous section. This is not the case for the next topic and the subject matter on the next section. We are about to cross the boundaries of Natural Logic.

3.2 Factives

Factives and counterfactives are well-known classes of verbs that take sentential or infinitival complements, first discussed by Kiparsky and Kiparsky [19]. They were among the first types of linguistic data that sparked the debate about **presuppositions**, still inconclusive 35 years later.

Philosophers had been talking about presuppositions for much of the 20th century focusing on a few examples like *The present king of France is bald* and *Have you stopped beating your wife?* The first is due to Bertrand Russell [24], the second to Eubulides (4th century BC).[6]

When linguists got fascinated with presuppositions in the late 1960s, within a few years they made a whole zoo of ‘presupposition triggers’ that included a large collection of lexical items and constructions, factives and counterfactives being among the first. In hindsight, the fundamental error at the time was not to recognize the newly discovered ‘triggers’ were not all of the same species. They should not have all been put into the same cage. The quest for a grand unified theory of presupposition as conceived in that period has been a failure.[7]

[6] Eubulides also bequeathed us the Liar Paradox: *what I now say is false*.

[7] That is the conclusion of the article by David Beaver and Bart Geurts in the Stanford Encyclopedia of Philosophy. `http://plato.stanford.edu/entries/presupposition/`

Nevertheless, the basic observations about the meaning of factive and counteractive constructions remain unchallenged. Table 4 lists some of these constructions.

Table 4. A few factive and counterfactive constructions

factive	counterfactive
remember that	pretend that
forget that	pretend to
be bad to	
be bad that	

The difference between *remember to* vs. *remember that* is striking. In affirmative sentences and under negation *remember that* commits the speaker to the view that the embedded clause is true. With *remember to* we get a positive or negative entailment depending on the polarity of the upstairs clause, with *remember that* we only get a positive inference in (11) regardless of the polarity of the upstairs clause.

(11) a. She remembered to lock the door. $\sqsubset$ She locked the door.
b. She did not remember to lock the door. $\sqsubset$ She did not lock the door.
c. She remembered that she locked the door. $\leq$ She locked the door.
d. She did not remember that she locked the door. $\leq$ She locked the door

In the case of both (11a) and (11c) the speaker is committed to the proposition that she locked the door, but not in the same way. (11a) is a two-way implicative construction that yields a negative entailment under negation in (11b). (11c) and its negation (11d) presuppose that she locked the door. As before we use $\sqsubset$ for entailment and a new symbol, $\leq$, for presupposition.

The difference between $\sqsubset$ and $\leq$ is that presuppositions 'project' from embedded clauses in a way that entailments do not. [14] Presuppositions are impervious to negation as in (11d) and they project from the antecedent of conditionals as in (12).

(12) If she remembered that she locked the door, she did not have to drive back home. $\leq$ She locked the door.

The difference between the two-way implicative *remember to* and the factive *remember that* cannot be pinned on the complementizer, *to* vs. *that*. The constructions *be bad to* and *be bad that* are both factive, *pretend that* and *prentend to* are both counterfactve as illustrated in (13).

(13) a. It was not bad for us that we had one day of rain on our trip.
$\leq$ We had one day of rain.
b. It was not bad for us to have one day of rain on our trip.
$\leq$ We had one day of rain.
c. Kim pretended that she had everything under control.
$\leq$ Kim did not have everything under control.
d. Kim pretended to have everything under control.
$\leq$ Kim did not have everything under control.

There are no general systems that we know of for computing inferences based on factive and counterfactive constructions or many other types of 'projective meaning.' The computational tools developed in the framework of Discourse Representation Theory, [13], [12], [4], are mainly focused on anaphoric expressions and definite noun phrases, a small subset of the phenomena that have been called presuppositions.

4 Beyond Natural Logic

The progress of Natural Logic has demonstrated that it is possible to do formal reasoning on rather shallow representations of natural language sentences. For the topics covered in the previous sections it is not evident that one could do better by a translation into a formal language such as predicate calculus or a higher-order logic. The shortcomings and unsolved problems with Natural Logic that we survey in this last section are challenging in any framework for semantics. The common thread of the inference problems discussed below is the need to take into account pragmatic factors, the context of use and even the perceived intent of the speaker.

The issue we start with is that people make inferences that go beyond what the sentence logically entails or presupposes. We call them **soft inferences** because they may explicitly cancelled but, if there is no indication otherwise, they may well a be part of what the speaker intends to convey. The second problem we uncovered in our investigation of evaluative adjectives such as *stupid*, *clever*, etc. It turns out that the interpretation of expressions like *NP was not stupid to VP* as implicative or factive depends on the relationship between the evaluative adjective and the action expressed by the VP. We call it the **consonance/dissonance effect**. Finally we discuss the curious case of *lucky* revisiting the issues first uncovered in [17] highlighting the fact that the choice of meaning may depend on the perceived intent of the speaker.

4.1 Soft Inferences

One-way implicatives yield a definite entailment only under one polarity, but in many contexts they are interpreted as if they were two-way implicatives. Although *be able* is logically a $--$ implicative (see Footnote 5), in the vast majority of occurrences 'in the wild' are like (14). The intent is certainly to convey that not only was Williamson able to deliver but also did so.

(14) New Zealand called and Kane Williamson was able to deliver.

Here *be able* is clearly used to mean *manage*.

A similar observation can be made of some $+-$ implicatives like *prevent*. Examples like (15) are not contradictory. If something is not prevented it might still not happen for other reasons.[8]

(15) Her mother did not prevent her from visiting her father, but she never did.

[8] For an insightful study of *prevent* see [5].

But such examples are vanishingly rare. In the common usage *prevent* behaves like a two-way $+- \mid -+$ implicative. When there is no reason to assume otherwise, examples like (16) are understood—and undoubtedly meant to be understood—to mean that a few laughs were had even across the language barrier.

(16) The language barrier did not prevent us from having a few laughs together.

There must be some unknown pragmatic explanation why we tend to assume that when someone says that something was allowed or not prevented to occur she means that it did occur if there is nothing to suggest that it didn't. This seems to be a universal principle, not a fact about the usage of English. This may be related to the phenomenon of **conditional perfection** discussed by Michael Geis and Arnold Zwicky [9] that pushes us to interpret simple conditionals *if p then q* as biconditionals *if p and only if p then q*.

Another case of non-logical inference is illustrated in (17).

(17) a. I meant to answer your email right away.
b. I didn't mean to hurt your feelings.

The speaker of (17a) probably hasn't quickly answrered the addressee's email. The speaker of (17b) probably thinks that she has hurt the addressee's feelings. There construction *mean to VP* is of course not logically of type $+- \mid -+$. But there is a pragmatic reason why we are inclined to draw such inferences. If we feel responsible for some bad outcome, a standard way of excusing ourselves is to assert that what happened was not what we intended. The soft inference arises from the understanding the situations where the speaker would be likely to use the expressions in (17), not from any semantic relation between the sentence and the infinitival clause.

4.2 Consonance/Dissonance Effect

Most of the work descriptive work on presupposition and entailment has focused on verbs, there is very little literature on adjectives. In our joint Stanford project [18] we decided to explore the semantics of evaluative adjectives such as *stupid*, *clever*, *wise*, *brave*, *rude*, *kind*, etc. in constructions of the form *NP was ADJ to VP* and *NP was not ADJ to VP*.[9] The only substantive treatise on this topic we found is a dissertation by Neal Norrick from the 1970s [23].

According to Norrick's classification evaluative adjectives (his term) are factive. But Norrick does not discuss any examples of the type *NP was not ADJ to VP* that would bring out the difference between factives and implicatives. Norrick's judgement has been passed on from author to author including the often cited paper by Chris Barker [1] without ever having been evaluated with real data.

We decided to study the issue with due diligence, This now means asking for judgements not just of your students, friends, and colleagues but of large set of workers on the Amazon Mechanical Turk (*Turkers* they are called). We involved over nine hundred people in subsequent iterations of our crowdsourcing experiment.

[9] This section based on joint research with Cleo Condoravdi, Stanley Peters, and Annie Zaenen but my colleagues are not responsible for the views expressed here.

Because negative sentences clearly distinguish factive and implicative constructions we were interested to know the response to sentences such as (18).

(18) Paul was not smart to take the middle piece.

It turned out that the great majority of our Turkers gave this type of sentence a factive interpretation: Paul was not smart and took the middle piece. A minority of our Turkers chose the implicative reading: Paul was not smart and did not take the middle piece. This finding suggests that there is a dialect split. The majority of our subjects prefer the factive reading, a minority prefers the implicative reading.

The surprising finding was that both populations can be pushed towards their non-favored interpretation by manipulating the interplay with the adjective and the content of the VP. Running the study on (19) gave different results for the two variants.

(19) a. Paul was not smart take the best piece.
b. Paul was not smart take the worst piece.

In (19a) the adjective *smart* and the VP *take the best piece* are in a **consonant** relation: taking the best piece would be smart. In (19b) the adjective and the VP are **dissonant**: taking the worst piece would not be smart. The results of our study so far indicate that a negation of a consonant relation such as (19a) favors the implicative interpretation: Paul did not take the best piece. The negation of a dissonant relation as in (19b) biases the Turkers towards the factive reading: Paul took the worst piece.

Figure 5 summarizes our overall findings for the 21 adjectives in our study. Assuming that there in no consonance/dissonance effect in the neutral case, the middle columns give an estimate of the proportion of factive and implicative speakers.[10]

The figure shows that about 80% of the Turkers gave a factive interpretation to dissonant examples such as (19b). In a consonant case such as (19a) the majority still preferred the factive interpretation but the number of implicative interpretations doubled from the neutral case. The consonance/dissonance was particularly strong for adjectives such as *lucky* and *fortunate*.

The sobering finding of this study that we are now in the progress of replicating with a more careful experimental design suggests that some very basic inferences such as whether the event described by an infinitival complement happened or not depend on opinions that are not part of the literal meaning of the sentence. This is a difficult problem for compositional semantics and for Natural Logic as well.[11]

[10] The Either columns shows the number of subjects who said they couldn't decide between the two possible interpretations, factive or implicative.

[11] In setting up the original experiment we tried to guess what people's opinions were, say, about how lucky it would be to go to San Francisco or live in Europe. Both neutral we thought, but the results show that for our subjects *lucky to go to San Francisco* was consonant but *lucky to live in Europe* a case of dissonance.

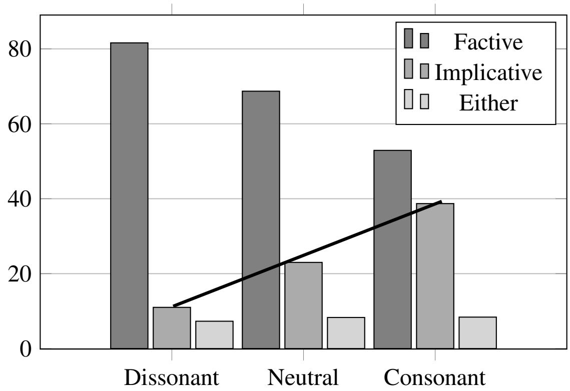

Fig. 5. Results: Percentage of Factive, Implicative, and Either choices for NP *was not* A*dj to* VP

4.3 Lucky

In addition to the strong consonance/dissonance effect, the adjective *lucky* is an interesting case for another reason. In affirmative sentence with a future tense it has two possible interpretation. A sentence such as (20) has a positive sense for most people.

(20) My future boyfriend will be so lucky to have me cook yummy food like this for him every day.

It is understood as a promise of benefits to the future boy friend. Seen as a caption to a picture of a table with delicious dishes, (20) has no other plausible interpretation.

In contrast, examples like (21) are typically interpreted as conveying a pessimistic warning: you will probably not get any return on your investments.

(21) Your will be lucky to ever get any return on your investments.

After all, what else would license the negative polarity items *ever* an *any* in the seemingly positive environments that contains none of the usual triggers of negative polarity?

An example such as (22) also suggests the pessimistic ‘probably not’ interpretation.

(22) You will be lucky to avoid a jail sentence.

The choice between the two meanings depends on many factors. With a small change the interpretation of (22) can be flipped:

(23) At least you will be lucky to avoid a jail sentence.

What *at least* least does here conversationally is to indicate that the speaker is trying to find something positive to say in an obviously bad situation, looking for a silver lining on a dark cloud. The fact that we recognize the speaker's intention to console the addressee is enough to suppress the 'probably not' interpretation. See [17] for further discussion.

5 Conclusion

We are impressed by the great progress in Natural Logic in the last few years and very aware of the need to ground it in a more comprehensive framework of Natural Reasoning that supplements logical relations with pragmatic inferences. Overall we are very optimistic about the future of this enterprise. Natural Logic is very suited for computational tasks.The improvements in hardware, the software for machine learning, and the easy way to collect data from the data on the web and by experiments with tools like AMT will advance the state of the art. The challenge is the integration of pragmatic and logical information,

References

1. Barker, C.: The dynamics of vagueness. Linguistics and Philosophy 25(1), 1–36 (2002)
2. van Benthem, J.: Language in Action: categories, lambdas and dynamic logic. Studies in Logic, vol. 130. Elsevier, Amsterdam (1991)
3. van Benthem, J.: A brief history of natural logic. In: Chakraborti, M.K., Löwe, B., Mitra, M.N., Sarukkai, S. (eds.) Logic, Navya-Nyāya & Applications, Homage to Bimal Krishna Matilal. College Publications, London (2008)
4. Bos, J.: Implementing the binding and accommodation theory for anaphora resolution and presupposition projection. Computational Linguistics 29(2), 179–210 (2003), http://dx.doi.org/10.1162/089120103322145306
5. Condoravdi, C., Crouch, D., Everett, J., Paiva, V., Stolle, R., Bobrow, D., van den Berg, M.: Preventing existence. In: Proceedings of the International Conference on Formal Ontology in Information Systems, FOIS 2001, pp. 162–173. ACM, New York (2001), http://doi.acm.org/10.1145/505168.505184
6. Dowty, D.: The role of negative polarity and concord marking in natural language reasoning. In: Harvey, M., Santelman, L. (eds.) Proceedings of SALT 4, pp. 114–144. University of Rochester (1994)
7. van Eijck, J.: Syllogistics=monotonicity+symmetry+existential import, Technical Report, vol. SEN-R0512. CWI, Amsterdam (2005), http://db.cwi.nl/rapporten/
8. Van Eijck, J.: Natural logic for natural language. In: ten Cate, B.D., Zeevat, H.W. (eds.) TbiLLC 2005. LNCS (LNAI), vol. 4363, pp. 216–230. Springer, Heidelberg (2007)
9. Geis, M.L., Zwicky, A.M.: On invited inferences. Linguistic Inquiry 2(4), 561–566 (1971), http://www.jstor.org/stable/4177664
10. Icard III, T.: Inclusion and exclusion in natural language. Studia Logica 100(4), 705–725 (2012)
11. Icard III, T., Moss, L.: Recent progress on monotonicity. In: Zaenen, A., Condoravdi, C., de Paiva, V. (eds.) Perspectives on Semantic Representations for Textual Inference. LILT, vol. 9, pp. 167–194. CSLI Publications, Stanford (2014)
12. Kamp, H.: The importance of presupposition. In: Rohrer, C., Roßdeutscher, A., Kamp, H. (eds.) Linguistic Form and its Computation. CSLI Publications, Stanford (2001)
13. Kamp, H., Reyle, U.: From Discourse to Logic: An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and DRT. Kluwer, Dordrecht(1993)

14. Karttunen, L.: Presuppositions of compound sentences. Linguistic Inquiry 4(2), 169–193 (1973)
15. Karttunen, L.: Implicative verbs. Language 47, 340–358 (2012)
16. Karttunen, L.: Simple and phrasal implicatives. In: *SEM 2012, June 7-8, pp. 124–131. Association for Computational Linguistics, Montréal (2012), http://www.aclweb.org/anthology/S12-1020
17. Karttunen, L.: You will be lucky to break even. In: King, T.H., de Paiva, V. (eds.) From Quirky Case to Representing Space. Papers in Honor of Annie Zaenen, pp. 167–180. CSLI Publications, Stanford (2013)
18. Karttunen, L., Peters, S., Zaenen, A., Condoravdi, C.: The chameleon-like nature of evaluative adjectives. In: Empirical Issues in Syntax and Semantics 10, pp. 233–250. CSSP, Paris (2014)
19. Kiparsky, P., Kiparsky, C.: Fact. In: Bierwisch, M., Heidolph, K.E. (eds.) Progress in Linguistics, pp. 143–173. Mouton, Hague (1970)
20. MacCartney, B.: Natural Language Inference. PhD thesis. Stanford University (2009)
21. MacCartney, B., Manning, D.C.: An extended model of natural logic. In: Proceedings of the Eight International Conference on Computational Semantics, pp. 140–156. Association for Computational Linguistics (2009)
22. Nairn, R., Condoravdi, C., Karttunen, L.: Computing relative polarity for textual inference. In: ICoS-5, pp. 67–76 (2006)
23. Norrick, N.R.: Factive adjectives and the theory of factivity. Niemeyer, Tübingen (1978)
24. Russell, B.: On denoting. Mind 14(56), 479–493 (1905)
25. Smessaert, H.: Monotonicity properties of comparative determiners. Linguistics and Philosophy 8(3), 295–336 (1996)
26. Sánchez-Valencia, V.: Studies on Natural Logic and Categorial Grammar. PhD thesis. University of Amsterdam (1991)