The Role of Negative Polarity and Concord Marking in Natural Language Reasoning

David Dowty Ohio State University dowty@ling.ohio-state.edu

1 Sánchez Valencia's Natural Logic

1.1 Semantic vs. Deductive Treatments of Inference

For better or for worse, most of the large body of research on natural language semantics done in the past two decades has employed semantic (model-theoretic) methods but has ignored deduction. Of course, deductive and semantic analysis of the same 'logic' have long been viewed as intrinsically complementary methods by logicians, for well-known reasons. It is also a truism that humans do not carry anything analogous to infinite sets of possible worlds or situations around in their heads, so the study of deduction—inferential relations based on syntactic properties of some kind of "representations" of denotations—are potentially of relevance to the psychology of language and to computational processing of meaning in a way that model-theoretic semantics alone is not.

The construction of deductive systems is sometimes said to involve "making certain semantic distinctions visible in the syntax", i.e. as syntactic properties that deductive rules can be made sensitive to. For example, consider the typed lambda calculus; though our semantics comprehends all possible operations of types (t,t) and ((t,t),t) (i.e. truth-functions) already, it is only by allowing deductive rules to single out $\neg, \land, \lor, \rightarrow$, etc. syntactically that we can formulate the deductive rules of the propositional calculus.

But exactly which semantic distinctions are relevant? Many alternatives are possible, depending on our goals, and many semantic distinctions do not seem to be of broad relevance for any goal (e.g. that between different individual predicates, like horse vs.cow). From the beginnings of the history of logic (and particularly with Leibniz), up through the recent past (the generative semanticists), and present, one pervasive goal has been the construction of a NATURAL LOGIC, one which permits the description of the most important common and important inferences of natural language, in terms of syntactic forms which have as much formal similarity to natural language sentences as possible (and/or as much justification as abstract "logical forms" in an

empirical linguistic theory as possible). Without accepting the very strong assumption of the generative semanticists—that there is a *unique* Natural Logic that captures all and only the linguistically expressible valid inferences (Lakoff 1972)—one can still agree that there is interest in studying deductive systems which approximate the class of common linguistic inferences to some interesting degree or in some interesting way.

1.2 The Linguistic Significance of Upward and Downward Monotone Inferences

Jack Hoeksema (1986), Johan van Benthem (1991), and Víctor Sánchez Valencia (1991) have all called attention to the significance and prevalence of upward and downward monotone inference patterns in natural language inference. (I henceforth write " \uparrow M" for 'upward monotone' and " \downarrow M" for 'downward monotone'.) Some examples from Hoeksema involving lexical hyponyms are in (1) (\uparrow M contexts) and (2) (\downarrow M contexts):

A	A Y	Nina has a bulldog bulldog < dog Nina has a dog	kiss < t	ssed Nina ouch uched Nina
A	A Y > B B Y	She didn't give him a fl flower > rose She didn't give him a ro		Every carp is a fish carp > koi Every koi is a fish

In his PhD thesis, Sánchez Valencia (1991) observes that the numerous rules of traditional Aristotelian Syllogistic, which is a logic based on common nouns and on verb phrases, can all be viewed as either as (i) an inference replacing one predicate with another denoting a superset of the first, in one class of environments—the 1M ones; or else as (ii) an inference replacing a predicate with another denoting a subset of the first, in a different class of environments— hand, Sánchez points out that the following two rules are also in effect \(\gamma \) and ↓M inferences: (i) substituting one formula for another (when these stand in a conditional relationship) in a formula where the first is within the scope of an even number of negations (including of course zero negations)—modus ponens and transitivity are two instances of this rule; and (ii) a similar substitution with the converse conditional relationship in a context where the replaced formula is in the scope of an odd number of negations—a rule having modus tollens as an instance. (Sánchez also investigates some less than successful attempts to build general monotonicity rules into a logic made by De Morgan, Leibnitz and Ockham.)

1.3 $\uparrow M$ vs. $\downarrow M$ Contexts and Existence Entailments

Before going on to Sánchez' system itself, I might mention two other linguistically significant manifestations of monotonicity differentiation in natural

.

language semantics.

One is the now very familiar observation by Ladusaw (1979) that negative polarity items in English and many languages are (with certain pragmatically-governed exceptions) limited to $\downarrow M$ contexts.

Another interesting linguistic correlation with monotonicity, not pointed out explicitly as such apparently, involves differential existence entailments from NPs and their relationship to anaphoric possibilities. It is easy to show that: if Det is a determiner that is $\uparrow M$ or non-monotone for both its restriction argument (R) and nuclear scope argument (NS), and ϕ is an $\uparrow M$ or non-monotone context (formal characterizations of these monotonicity notions can be given in the system presented below), then

```
\phi(Det(R)(NS)) entails (except in one trivial case)<sup>1</sup> R \cap NS \neq \emptyset
```

However, if (i) Det is $\downarrow M$ for either its R or NS argument, and ϕ is an $\uparrow M$ context, or (ii) Det is $\uparrow M$ (or non-monotone) and ϕ is a $\downarrow M$ context, then

```
\phi(Det(R)(NS)) never entails R \cap NS \neq \emptyset (and in the case of the Det no, entails in fact that R \cap NS = \emptyset.)
```

Thus for example, Sam baked some (many, several, at least a dozen, etc.) brownies yesterday entails that there is a non-empty intersection between the set of brownies and the set of things Sam baked yesterday, while all of Sam didn't bake brownies yesterday. Sam didn't bake any brownies yesterday and Sam baked no brownies yesterday lack this entailment, entailing in fact that this intersection is empty. Given this observation, it makes sense to entertain the following DISCOURSE ANAPHORA HYPOTHESIS: The prototypical ("core") cases of discourse anaphora are those where a pronoun refers to a set (possibly singleton) entailed to exist by a previous sentence in accord with the afore-mentioned principle for $\uparrow M$ versus $\downarrow M$ determiners and contexts, where Det(R)(NS) corresponds to the "linguistic antecedent" of the pronoun. Call such cases SEMANTICALLY ENTAILED DISCOURSE REFERENTS. It follows from the above observation that there can be no semantically entailed discourse referents with $\downarrow M$ Dets (or other Dets in sentential $\downarrow M$ contexts). So for example, definite discourse anaphora is possible in (3a) but not (3b) or (3c):

- (3) a. Sam baked some brownies yesterday. They were delicious.
 - b. Sam didn't bake brownies yesterday/Sam didn't bake any brownies yesterday. *They were delicious.
 - c. Sam baked no brownies yesterday. *They were delicious.

To be sure, not all cases of "discourse reference" involve semantically entailed discourse referents as characterized above; such reference is possible when contextual information or pragmatic reasoning of various kinds often allows

¹This case is the hypothetical determiner 'zero or more', mentioned on occasion by Ed Keenan.

a hearer to infer a non-empty intersection of R and NS and thus a discourse antecedent. The determiners every and few are logically consistent with both empty and non-empty intersections. To see an example, consider first "existential import", the inference that in $\mathbf{every}(R)(NS)$, this intersection is non-empty — now generally accepted to be a pragmatic rather than logical inference. Such cases naturally lead to acceptable discourses like Every first-year student; in the department came to my party last night. They, left earlier than the other students, though. If such an existential import implicature is canceled in one way or another, then anaphora is no longer possible. Although a great variety of cases of pragmatically "assisted" anaphora undoubtedly exist, too varied to survey here (see Dowty 1993, to appear for some discussion), it does seem that when all contextual and pragmatic factors are carefully excluded, the above principle correctly delimits the cases where definite discourse anaphora is possible.²

1.4 Cross-Categorial Definition of Monotonicity

Sánchez Valencia's own monotonicity logic, which can be sketched only briefly here, is based on a version of the Lambek calculus (a kind of highly general categorial grammar). He begins by defining (recursively) a partial ordering among denotations of each logical type (Sánchez 1991:67-94, 95-100)

- (4) a. If $c, d \in D_e$, then $c \leq_e d$ iff c = d.
 - b. If c, $d \in D_t$ then $c \le_t d$ iff c = 0 or d = 1 (i.e., \le_t is the truth table for " \rightarrow ", or 'less-than-or-equal-to' on the set $\{0,1\}$).
 - c. If c, $d \in D_{(\alpha,\beta)}$, then $c \leq_{(\alpha,\beta)} d$ iff for each $a \in D_{\alpha}$, $c(a) \leq_{\beta} d(a)$.

For example, for expressions of type (e,t), i.e. sets, it follows from (4) that $a \leq_{(e,t)} b$ means that a is a subset of b. If p and q are formulas, $p \leq_t q$ is the same as $(p \to q)$.

The second step is to define upward and downward monotonicity across all types in terms of this ordering (Sánchez 1991:101):

- (5) a. A function $z \in D_{(\alpha,\beta)}$ is UPWARD MONOTONE iff for every $x,y \in D_{\alpha}$, $x \leq_{\alpha} y$ entails $z(x) \leq_{\beta} z(y)$.
 - b. A function $z \in D_{(\alpha,\beta)}$ is DOWNWARD MONOTONE iff for every $x,y \in D_{\alpha}$, $x \leq_{\alpha} y$ entails $z(y) \leq_{\beta} z(x)$.
 - c. A function $z \in D_{(\alpha,\beta)}$ is NON-MONOTONE iff it is neither upward monotone nor downward monotone.

²This is not to deny however that a fairly complex theory, such as Discourse Representation Theory, is well-motivated, in order to subsume more unusual discourse phenomena such as modal subordination (Roberts 1989) and sentence-internal anaphora such as bound variable anaphora of both regular and donkey varieties; rather, the point is that it is the fact of semantically-entailed reference that makes basic discourse anaphora for quantificational antecedents possible at all.

A function which is $\uparrow M$ is thus one which preserves the relative ordering of any two of its arguments (in their own domain) in the relative ordering of the respective values it maps these arguments into (in their domain). A $\downarrow M$ function, by contrast, maps arguments into values which are ordered in the opposite way from the arguments. Hence $\uparrow M$ functions are also known as MONOTONICITY-PRESERVING functions, while $\downarrow M$ functions are also known as MONOTONICITY-REVERSING functions, a terminology which we will find useful later on.

To illustrate with some linguistic examples: in the garden is $\uparrow M$ in ((e,t),(e,t)) (predicate modifiers denotations) because $\llbracket \text{dance} \rrbracket \leq \llbracket \text{moves} \rrbracket$, and "Sam [dances in the garden]" entails "Sam [moves in the garden]" (and, crucially, for ALL predicates α and β such that $\llbracket \alpha \rrbracket \leq \llbracket \beta \rrbracket$, "Sam $\llbracket \alpha \rrbracket$ in the garden]" entails "Sam $\llbracket \beta \rrbracket$ in the garden]." However, didn't is $\downarrow M$ in ((e,t),(e,t)), because "Sue $\llbracket \text{didn't} \text{ move} \rrbracket$ " entails "Sue $\llbracket \text{didn't} \text{ dance} \rrbracket$ " (and similarly for other predicate pairs). without is $\downarrow M$ in ((e,t),(e,t)), since "Sam talked [without moving]" entails "Sam talked [without dancing]", etc.

1.5 Lexical Monotonicity Marking

Since basic expressions differ with respect to their monotonicity entailments (for example **some** vs. **every** vs. **no** differ, as is familiar from both generalized quantifier and Ladusaw's work on polarity items), it is necessary to have such items entered as basic expressions in categories which are somehow syntactically distinct. Sánchez (1991:110) thus introduces new logical types for functor expressions which are $\uparrow M$ (respectively $\downarrow M$) for their arguments:

```
(6) If (\alpha, \beta) is a category, then (\alpha^+, \beta) and (\alpha^-, \beta) are also categories.
```

(The 'plain' category (α, β) remains a well-formed category name alongside the other two; it is the category of non-monotone functors.)

A reasonable assignment of lexical determiners to such categories (here, as functions from sets to functions from sets to truth values) would thus be as follows:

```
(7) a(n) is assigned to: ((e,t)^+,((e,t)^+,t))

every is assigned to: ((e,t)^-,((e,t)^+,t))

no is assigned to: ((e,t)^-,((e,t)^-,t))
```

Such lexical monotonicity assignments do not of course suffice to correctly predict the inferences that can be drawn in all sentences, of course. The above type for $\mathbf{a}(\mathbf{n})$ would, for example, predict $\uparrow \mathbf{M}$ inferences for both Restriction and Nuclear Scope arguments, and this is correct for some cases (e.g. that from An armadillo danced one can infer both An animal danced and An armadillo moved). We do not however get a $\uparrow \mathbf{M}$ inference from the noun armadillo in Sam didn't catch an armadillo but instead $\downarrow \mathbf{M}$ inferences such as Sam didn't catch a female animal. Similarly, from Sue sang without catching every armadillo

we get not an $\downarrow M$ inferences from this same noun but $\uparrow M$ inferences such as $Sue\ sang\ without\ catching\ every\ animal$. Obviously, compositional rules are required which take account of the monotonicity effects of other functors in the sentence.

1.6 Step II: External Monotonicity Marking:

Two more steps are employed, the first of which is (EXTERNAL) MONOTONIC-ITY MARKING (Sánchez 1991:110)³, which adds '+' and '-' to functor and argument nodes in a derivation tree:

To make this and the next step intuitive without going into details, it will help to look at examples. For purposes of this paper, I will use skeleton derivations of English sentences, using only logical types like ((e, t), t) as categories, whereas a linguistically adequate categorial grammar would enrich these with numerous syntactic features (as in HPSG): the reader may imagine these as added. Of the various ways of treating quantificational noun phrases in object position that have been proposed in the recent categorial literature, I will adopt the method of treating such NPs as functors which combine with transitive verbs to form (intransitive) verb phrases; though this "duplicates" determiners as well as NPs in different object and subject categories, it permits me to use simpler derivation trees. The reader who is not familiar with this approach is asked to take it on faith that the assignments of "+" and "-" for the object-category determiners are the appropriate one, given the subject-determiner assignments in (7). Finally, I will ignore the problem of accounting formally for correct word order; functors will sometimes be placed to the left of their arguments, sometimes to the right, as needed for English word order. Here then is one derivation before and after the Monotonicity Marking Rules have been applied ((9) vs. (10)), and another derivation shown after Marking, (11):

³Again, a number of complications are ignored here; for example, this and the next definition are actually more complex because the Lambek calculus is used, which allows also withdrawal of terms; also, the semantics of monotonicity in the lambda calculus must be treated, because of its use in interpreting withdrawal of Lambek terms

(9) Sue caught every armadillo
$$\frac{((e,t),t)}{((e,t),t)} = \frac{(e,(e,t))}{((e,(e,t))^+,(e,t))} = \frac{((e,t)^-,((e,(e,t))^+,(e,t)))}{((e,t)^-,((e,(e,t))^+,(e,t)))}$$

(10) Sue caught every armadillo
$$((e,t),t)$$
 $(e,(e,t))$ $((e,t)^-,((e,(e,t))^+,(e,t)))$ (e,t) $+$ $+$ (e,t) $+$ $+$ t

(11)

1.7 Step III: Determining the Polarity of Each Constituent

To determine the actual direction of inference ($\uparrow M$, $\downarrow M$, or neither) for each constituent requires a third step, called POLARITY DETERMINATION, which is accomplished by these rules:

- (12) If D is a syntactic derivation (tree) with root node α , then:
 - a. A node γ has polarity in D iff all the nodes in the path from γ to α are marked, and
 - b. a node γ is POSITIVE iff γ has polarity and if the number of nodes in this path marked by '-' is even.
 - c. a node γ is NEGATIVE iff γ has polarity and if the number of nodes in this path marked by '-' is odd.

(Some nodes will fail to have a polarity, viz., those that are in the scope of a non-monotone function; I will ignore non-monotone functors in the rest of this paper, for the sake of brevity.)

The polarities derived by these rules for the two previous examples can be illustrated in the POLARITY SUMMARIES (13) and (14): these indicate the polarity of each word and constituent determined by the rules:

- (13) $Sue^+(caught^+(every^+(armadillo^-)^+)^+)^+$
- (14) $Sam^+(didn't^+(catch^-(every^-(armadillo^+)^-)^-)^+)$

In (13) for example, armadillo is marked with polarity '-' because there is one '-' in the path from this word to the root in the derivation (10), while this same word is marked '+' in (14), because the path to the root in (11) contains two '-' signs; both the constituents every armadillo and catch every armadillo have one '-' sign in their paths in (11), hence are marked negative in (14).

The rules of inference for this logic (which are proved by Sánchez to be valid for the semantics associated with the system) are the following, where "occurs positively/negatively" means of course positively (or negatively) according to the polarity determination defined as above:⁴

(15) Where expression M occurs positively in derivation N(M):

$$\frac{[\![M]\!] \leq [\![M']\!] \quad N(M)}{N(M')}$$

Where expression M' occurs negatively in derivation N(M'):

$$\frac{[\![M]\!] \leq [\![M']\!] \quad N(M')}{N(M)}$$

2 An alternative formulation of Sánchez' Monotonicity Logic

Although the format of Sanchez' deductive system (and a similar, less detailed one in van Benthem 1991) is no doubt highly appropriate for the logical studies he developed it for, it is unsuited to the linguistics applications I am interested in. First, it involves three separate steps (syntactic derivation, monotonicity marking, polarity determinations), which would results in a rather indirect linking between logic and negative polarity licensing. Second, notice that the

⁴Although "[M] \leq [M]" is, strictly speaking, a meta-language semantic statement rather than an object-language formula it can as far as I can see be replaced by an object-language formula schema $I_{\alpha}(M, M')$, meaning the denotation of M (of category α) is included in that of M', a notation that Sanchez himself uses, though as an abbreviation; cf. p. 125.

symbols "+" and "-" have a different significance in monotonicity-marked tree from that in polarity summary (e.g. a constituent can have a "+" monotonicity mark but a "-" polarity, or vice versa).

I will at this point present an alternative formulation, which I believe is almost certainly equivalent to Sanchez' (though I do not have a proof of this at present). The goal is to "collapse" the independent steps of Monotonicity Marking and Polarity Determination into the syntactic derivation itself, so that words and constituents are are generated with the markings already in place that they would receive in Sánchz' polarity summaries. The symbols "+" and "-" will be used, unambiguously, only to indicate the (final) logical polarity.

The following types will be employed, and here I now introduce rudimentary syntactic category distinctions which go beyond logical types, namely the familiar categorial grammar symbols S, NP, CN, etc.:

- (16) a. NP (= type e), S (= type t), and CN (= type (e,t)) are (primitive) categories.
 - b. If A and B are any categories, so is A/B.
 - c. If A/B is a category, so are A^+/B^+ , A^+/B^- , A^-/B^+ , and A^-/B^- .
 - d. Parallel definitions to be given for left-leaning-slash categories $B \setminus A$ (functor combining with B to yield A).⁵

We will in addition need to invoke polarity marking on complex categories themselves, e.g. $(NP \setminus S)^+$ vs. $(NP \setminus S)^-$. It turns out, however, that we do not really need to define these separately as categories; rather, it suffices to let the +/- value of the result-category of the functor indicate the +/- of the functor as a whole; thus we can use the following as a notational abbreviation:

(17)
$$(A/B)^+ =_{def} (A^+/B)^+ =_{def} (A^+/B)$$

 $(A/B)^- =_{def} (A^-/B)^- =_{def} (A^-/B)$

Since one and the same word (or constituent) can appear with positive polarity in one derivation and negative polarity in another (cf. catch and armadillo in the examples above), most lexical items will, in this formulation, be entered in two categories, a "+"-marked category and its "-"-marked counterpart (e.g. $catch \in (NP^+\backslash S^+)/NP^+$ and $catch \in (NP^-\backslash S^-)/NP^-$), though with the same semantic interpretation in each case. $\uparrow M$ and $\downarrow M$ functors also appear

⁵NB $B\setminus A$ will not, as in some notational usages, indicate a functor seeking an A on the left to form a B.

⁶The reason this is possible is that, as Hoeksema (1986) observes, and Sánchez proves, in a function-argument combination the function expression always has the same polarity as the combination as a whole (while the polarity of the argument expression, on the other hand, depends on whether the particular function is an ↑M or ↓M one); this is in fact the motivation for "+" invariably assigned to the function expressions in the Marking rules in (8)—note that a "+" at the step of Monotonicity Marking results in polarity preservation in the ultimate polarity determination, not necessarily positive polarity per se.

in two categories, but now the terms "preserving" and "reversing" are more appropriate:

- (18) a. Lexical items in general appear in both a "+"-marked and a "-"-marked category though with the same interpretation.
 - b. $\uparrow M$ functors (= MONOTONICITY PRESERVING FUNCTORS) appear in a pair of categories of the forms A^+/B^+ and A^-/B^- .
 - c. \downarrow M functors (= MONOTONICITY REVERSING FUNCTORS) appear in a pair of categories of the forms A^+/B^- and A^-/B^+ .

The categorial Slash-Elimination rules (or "Functional Application" Rules) must appropriately respect +/- marking:

(19) Polarity-Preserv. /-Elimin.: (20) Polarity-Revers. /-Elimin.:
$$\frac{A^{+}/B^{+} \quad B^{+}}{A^{+}} / E \qquad \qquad \frac{A^{+}/B^{-} \quad B^{-}}{A^{+}} / E$$
$$\frac{A^{-}/B^{-} \quad B^{-}}{A^{-}} / E \qquad \qquad \frac{A^{-}/B^{+} \quad B^{+}}{A^{-}} / E$$

Rather that literally specify the four separate rules indicated above however, we might regard "+" and "-" as syntactic features, as features are treated elsewhere in linguistics (but here with deductive significance), and these four rules would be simply instances of a single slash-elimination schema, where A and B are variables for complexes of syntactic features:

$$\frac{A/B \quad B}{A} / E$$

For various purposes (cf. below), one might want to adopt a notation for category schemas of the preserving and reversing categories, insofar as the same expressions appear in both members of such pairs:

(22) a.
$$A/PB =_{def} \{A^{+}/B^{+}, A^{-}/B^{-}\}$$

b. $A/RB =_{def} \{A^{+}/B^{-}, A^{-}/B^{+}\}$

To complete this reformulation of Sánchez' system, we need a final definition: a grammar as defined so far will generate sentences of the two subcategories S^+ and S^- , but it is only the former which should count as independent sentences (though the latter is also needed, for clauses embedded inside $\downarrow M$ functions)

(23) If ϕ is of category S^+ , ϕ is a well-formed (non-embedded) sentence.

Here are the category assignments we would make for determiners in this system: ("VP" is used as an abbreviation for $(NP \setminus S)$.)

$$\begin{array}{lll} \text{(24)} & \text{a} & \text{is assigned to: } \left\{ \begin{array}{l} (S^+/VP^+)/CN^+ \\ (S^-/VP^-)/CN^- \end{array} \right\} &= (S/_PVP)/_PCN \\ & \text{every is assigned to: } \left\{ \begin{array}{l} (S^+/VP^+)/CN^- \\ (S^-/VP^-)/CN^+ \end{array} \right\} &= (S/_PVP)/_RCN \\ & \text{no} & \text{is assigned to: } \left\{ \begin{array}{l} (S^+/VP^-)/CN^- \\ (S^-/VP^+)/CN^+ \end{array} \right\} &= (S/_RVP)/_RCN \end{array}$$

As direct object quantificational NPs are treated here as functions from transitive verbs (category $(NP \backslash S)/NP$, abbreviated VP/NP or simply TV) to VP), each determiner in category $(S^{\alpha}/VP^{\beta})/CN^{\gamma}$ above is assumed to have an object-NP counterpart in $(TV^{\beta} \backslash VP^{\alpha})/CN^{\gamma}$. Once again, the operation of the system is best understood through a few simple examples:

$$\begin{array}{c|c} & & & & & & & \\ & & & & & & \\ \hline Sue & & & & & \\ \hline S^+/VP^+ & & & & & \\ \hline S^+/VP^+ & & & & & \\ \hline \end{array}$$

If a sentence is well-formed, it must be derivable in category S^+ , so we can use an informal procedure analogous to Genzen-Sequent Parsing (cf. Moortgat 1991) to determine what polarity marking must appear on each category. So in (25), Sue is the main functor, and as it is polarity-preserving, it will appear only in the categories S^+/VP^+ and S^-/VP^- . Of these, only the former could result in an S^+ , so that is the category for Sue. This implies in turn that the VP is marked VP^+ , else it could not combine with S^+/VP^+ . The verb catch is a monotonicity-preserving function, and if it had had a simple NP object, we could deduce it must be marked VP^+/NP^+ by parallel reasoning. Instead, it is the argument of the quantificational NP functor, but it turns out that all object NPs in every will be monotone-preserving, hence of either $TV^+ \backslash VP^+$ or $TV^- \backslash VP^-$, but again only the former could result in VP^+ after combining with an argument, so this determines both the category of the object NP and of the verb, TV^+/NP^+ . Given the lexical choices for object every, $(TV^+\backslash VP^+)/CN^-$ and $(TV^-\backslash VP^-)/CN^+$, only the former gives the needed result category, thus fixing the polarity for it and also determining that the CN argument of every appears in CN^- . Thus there is a unique polarity assignment possible for this example, marked straightforwardly in the derivation tree, and it is in fact the same as the one determined by Sánchez' method in (9) and (13).

The next example adds a polarity reversing auxiliary verb didn't, which will be entered in categories VP^+/VP^- and VP^-/VP^+ . In order for the whole VP didn't catch every armadillo to be in VP^+ here, the VP argument of didn't must be in VP^- , thus the words and sub-constituents of catch every armadillo must have the opposite of their polarities in (25):

3 NPIs and Negative Concord as Facilitating Inference Processing by Overtly Marking ↓M Contexts

From the extensive body of recent research on the many manifestations of negative polarity and concord, it appears such words and phrases have various historical sources (e.g. first as denoting 'minimal amounts', then as NPIs via conventionalization or lexicalization of the "pragmatic scales" first described by Fauconnier) and serve various functions for speakers (e.g. a means of 'strengthening' negation syntactically, or lending vividness, etc.) The main hypothesis of this paper is that another, very important function of these items has been overlooked:

(27) Hypothesis: Given that (i) ↑M and ↓M inferences are a very significant pattern of natural language reasoning, and (ii) the distribution of NPIs (and NC) is (almost) coextensive with logically ↓M contexts, we can hypothesize that one important reason for the existence of NPI and NC marking is to directly mark positions syntactically which are subject to ↓M inferences.

Another way of viewing this hypothesis is as suggesting that this direct morphological marking of $\downarrow M$ positions, also a way of "making certain semantic distinctions visible in the syntax", may in effect provide the natural language hearer with a "shortcut" to Monotonicity Inferences.

(28) a. NPI RULE: If expression M occurs as the argument of a NPI α in derivation N(M), then:

- An inference to N(M') is valid, where $[M'] \subseteq [M]$
- An inference to N(M') is not valid, where $[\![M]\!] \subseteq [\![M']\!]$
- b. DEFAULT RULE:⁷ If expression M occurs as the argument of a (lexical) expression α in derivation N(M), where α is not an NPI but is a member of a lexical category which has NPIs as members, then:
 - An inference to N(M') is valid, where $[\![M]\!] \subseteq [\![M']\!]$
 - An inference to N(M') is not valid, where $[\![M']\!] \subseteq [\![M]\!]$

Recalling our earlier observation about \$\Delta\$M NPs and existence entailments, a further processing role for noun phrase NPIs (any CN, anybody, etc.) can be suggested: since we have observed that there is no possible semantically-entailed discourse referent for a quantificational NP with \$\Delta\$M Det (or otherwise in a \$\Delta\$M context), the occurrence of a NPI-marked noun phrase may immediately tell the hearer, in advance of understanding a whole clause compositionally, that there is no possible semantically entailed discourse referent for that NP. (Again, contextually or pragmatically "assisted" discourse reference is to be excepted.)

3.1 Using a Monotonicity-Based Deductive System to Describe NPI and NC distribution

Since a monotonicity-based deductive system involves treating ↑M and ↓M properties as syntactic properties, the possibility of describing natural language NPI and NC (Negative Concord) distribution directly in terms of a logic like the preceding one naturally suggests itself. In an earlier paper (Dowty 1993). I constructed such an analysis on the original Sánchez formulation with its three-step polarity determination. As a consequence of the indirect way in which logical polarity is determined in this version, the licensing of NPIs and

⁷Rule (28b) is called a default rule because non-monotone contexts also exist, but these are relatively rare in natural language (e.g. the scope of quantifiers like *exactly ten*), with †M contexts being by far the most common.

NC was likewise quite indirect: NPIs (and 'concordant' negative morphemes in NC languages) were marked with a special syntactic label, and an "output filter" was imposed which stated that a derivation was to be considered well-formed if and only if all labeled morphemes turned out to occur in a logically negative context (i.e. terminated a path to the root of the sentence in which there was an odd number of "—" markings).

In the alternative formulation presented above, syntactic NPI licensing can be made much more direct and simple: we can take advantage of the fact that lexical expressions here appear in categories "already marked" for logical polarity. Up to now, I have assumed each word occurs in both a "+"-marked and a "-" -marked category variant, but we now modify this assignment for polarity items, both positive polarity items (PPIs) and NPIs, as follows:

- (29) a. NPIs are entered ONLY in "-"-marked categories.
 - b. PPIs are entered ONLY in "+"-marked categories.

The consequence of course is that NPIs, unlike words in general, can now *only* occur where their logical context is $\downarrow M$, and PPIs where the context is $\uparrow M$. Here are some sample category assignments:

(30)
$$\mathbf{a(n)} \qquad \text{is assigned to:} \quad \left\{ \begin{array}{l} (S^+/VP^+)/CN^+ \\ (S^-/VP^-)/CN^- \end{array} \right\}$$

$$\mathbf{any} \qquad \text{is assigned to:} \qquad (S^-/VP^-)/CN^- \\ \mathbf{several} \qquad \text{is assigned to:} \qquad (S^+/VP^+)/CN^+ \\ \mathbf{no} \qquad \text{is assigned to:} \qquad \left\{ \begin{array}{l} (S^+/VP^-)/CN^- \\ (S^-/VP^+)/CN^+ \end{array} \right\}$$

$$\mathbf{once} \qquad \text{is assigned to:} \qquad \left\{ \begin{array}{l} VP^+\backslash VP^+ \\ VP^-\backslash VP^- \end{array} \right\}$$

$$\mathbf{ever} \qquad \text{is assigned to:} \qquad \left\{ \begin{array}{l} VP^+\backslash VP^+ \\ VP^-/VP^- \end{array} \right\}$$

$$\mathbf{never} \qquad \text{is assigned to:} \qquad \left\{ \begin{array}{l} VP^-/VP^+ \\ VP^-/VP^- \end{array} \right\}$$

A derivation involving a NPI will look like those already shown,

$$\frac{\text{Sam}}{S^{+}/VP^{+}} = \frac{\frac{\text{didn't}}{VP^{-}/NP^{-}} = \frac{\frac{\text{catch}}{VP^{-}/NP^{-}} = \frac{(TV^{-}\backslash VP^{-})/CN^{-}}{TV^{-}\backslash VP^{-}}}{\frac{VP^{-}}{VP^{-}}/E} = \frac{\frac{\text{Sam}}{S^{+}/VP^{+}} = \frac{VP^{-}}{S^{+}}}{\frac{VP^{+}}{S^{+}}} = \frac{\frac{\text{Any}}{(TV^{-}\backslash VP^{-})/CN^{-}}}{\frac{(TV^{-}\backslash VP^{-})/CN^{-}}{TV^{-}\backslash VP^{-}}} = \frac{VP^{-}}{S^{+}}$$

but if we attempt to derive a sentence with an unlicensed NPI, the derivation

fails because the NPI does not occur in the "+"-marked category needed to complete the derivation (indicated by "⊕" in (32):

$$\begin{array}{c|c} \text{Sue} & \frac{\text{any}}{CN^{+}} & \frac{\text{armadillos}}{CN^{+}} \\ \hline S^{+}/VP^{+} & \frac{VP^{+}}{VP^{+}} \\ \hline \end{array}$$

3.2 Negative Concord and ↓M Marking

Negative Concord is now reasonably familiar in linguistic studies, not only of non-standard English but also Romance languages such as Italian and Catalan; the data below is from Ladusaw (1991,1992)

- (33) a. Nobody said nothing to nobody. [NS English] 'Nobody said anything to anyone.'
 - b. Maria didn't say nothing to nobody. [NS English] 'Maria didn't say anything to anyone.'
 - c. Mario non ha parlato di niente con nessuno. [Italian] 'Mario hasn't spoken with anyone about anything.'
 - d. No m'ha telefonat ningú. [Catalan] 'Nobody has telephoned me.'

Ladusaw sees a dilemma for previous analyses of Negative Concord; these essentially amount to treating one of the negatively-marked elements in a multiple-negative sentence as semantically potent negation (the "real negation"), and the other negatively marked NP(s) as plain existential quantifiers. (and didn't as meaning do in (33b)). Two kinds of problems then immediately present themselves in getting the distributions of the two kinds of "negative" words to work out properly. First, there is no "logical double negation" reading for (34),

(34) Mary didn't talk to nobody.
'Mary didn't talk to anybody'

nor indeed are such logical double negation readings usually found in any NC languages at all, according to Ladusaw, so logically potent negations must somehow be limited to one per sentence. Second, the postulation of "positive" semantic sources for negative morphemes raises the problem of insuring that these never surface *except* where C-commanded by a "real" (semantically negative) negative morpheme.

Ladusaw comments on the desirability of linking the description of NC with a more general theory that encompasses the licensing of NPIs and NC in their respective types of language, and I think there is not likely to be any disagreement with him on this last point. However, he has qualms about the existing ways of carrying out this linking by making the assumption that "negative" words like nobody in NC languages have the afore-mentioned ambiguity between real negations and NPI readings. Specifically: "the theory of negative polarity items restricts the distribution of the NPIs but does nothing to restrict the distribution of the negative expressors (beyond requiring that there be one if there are any NPIs). In a language which is strictly NC, something must be added to restrict the distribution of the negative quantifiers. Otherwise every sentence which contains multiple negative phrases should have both a double negation (DN) and a NC construal" (Ladusaw 1992: 248). The conclusion drawn in that paper is that the appropriate way to relate NC to NPIs is that none of the negative phrases in a NC sentence should be regarded as "expressing the negation." Rather, all should be viewed as NPIs, with the negative meaning itself expressed "constructionally", i.e. represented as an abstract operator in logical form that is triggered by syntactic rules, not by any one morpheme. (See Ladusaw 1992 for the complete statement of this argument.)

Under the view of the importance of $\downarrow M$ contexts taken here, it is natural to suggest the following approach to NC:

(35) Negative Concord Language Hypothesis: Given that ↓M vs. jother positions are significant for natural language inference, ↓M positions are just as important (if not more so) for a language to indicate syntactically than "where the negation is" in logical form. Hence some natural languages—the NC languages—employ words which are indeterminate as to signifying negation itself (i.e. are ambiguous between negative and NPI interpretations) but these words UNIFORMLY signify ↓M positions.

(Here "negation" means a determiner meaning that is anti-additive for both arguments (no, none of, cf. Zwarts 1993) or the image of classical negation in the relevant Boolean type (not S, not VP, etc., but excludes other \downarrow M functions.)

What is interesting is that when this hypothesis is formally implemented, the result is that the two features of negative morpheme distribution mentioned by Ladusaw follow automatically, as we will see shortly. This is a rather striking argument for the relevance of $\downarrow M$ inference-marking in NC/NPI phenomena.

We could if desired soften the unpleasant appearance (for some) of a postulation of systematic lexical ambiguity, and at the same time emphasize the uniform semantic role of these words as $\downarrow M$ indicators, by introducing a new kind of category schema for them:

(36) Let $A^{(-)}/B^-$ stand for two categories which contain corresponding pairs of (lexical) items, an expression in A^+/B^- with a polarity-reversing

("negative") meaning and a homophonous expression in A^-/B^- with polarity-preserving ("positive") meaning.

To be sure, this would not eliminate the essential systematic ambiguity involved, only give us a way to describe it more methodically. An assignment of a representative set of "negative" items to categories in such a system would be as follows:

(37)

(38)

(Determiners are still assumed to be assigned to the corresponding object-position category as well, e.g. for no here, $(TV^-\backslash VP^{(-)})/CN^-$).

An illustrative derivation is in (38): NB nobody appears in the reversing category S^+/VP^- and thus has a "true" negative translation, whereas didn't and no in the respective "NPI" categories VP^-/VP^- and $(TV^-\backslash VP^-)/CN^-$ will have NPI-type translations:

$$\frac{\operatorname{catch}}{\operatorname{catch}} = \frac{\frac{\operatorname{no}}{(TV^- \backslash VP^-)/CN^-} \frac{\operatorname{armadillos}}{CN^-}}{\frac{VP^- / NP^-}{S^+ / VP^-}} \frac{\frac{\operatorname{no}}{VP^- / NP^-}}{VP^- / E} = \frac{\frac{\operatorname{no}}{(TV^- \backslash VP^-)/CN^-}}{VP^- / E} = \frac{\frac{\operatorname{no}}{VP^- / NP^-}}{VP^- / E} = \frac{\operatorname{no}}{(TV^- \backslash VP^-)/CN^-} \frac{\operatorname{armadillos}}{(TV^- \backslash VP^-)/CN^-} / E$$

If on the other hand we attempted a derivation in which all "neg" words had NPI translations, as in (39), it would fail because the one of these with widest scope would have a A^-/B^- category, which could not result in a category S^+ for the whole sentence (or else only by ignoring the proper matching of "+/-" in functor and argument):

$$\frac{\text{didn't}}{S^{-}/VP^{-}} \frac{\frac{\text{catch}}{VP^{-}/NP^{-}} \frac{\text{no}}{(TV^{-}\backslash VP^{-})/CN^{-}} \frac{\text{armadillos}}{CN^{-}}}{VP^{-}/NP^{-}} /E$$

$$\frac{\frac{\text{Nobody}}{VP^{-}/VP^{-}} \frac{VP^{-}}{VP^{-}}}{S^{\oplus}} **/E$$

Example (40) shows why an attempt at a logical double negation reading fails, and (41) shows that a derivation also fails which has both a semantically negative and NPI-interpreted negatives if the former fails to F-command the latter: in both cases, the reason is that a negative-marked word would fail to correctly signify a logically $\downarrow M$ context:

$$(40) \qquad \qquad \frac{\text{no}}{\frac{\text{catch}}{VP^{-}/NP^{-}}} \frac{\frac{\text{no}}{(TV^{-}\backslash VP^{+})/CN^{-}} \frac{\text{armadillos}}{CN^{-}}}{VP^{-}/NP^{-}} / E$$

$$\frac{\frac{\text{Sue}}{S^{+}/VP^{+}} \frac{VP^{+}/VP^{-}}{VP^{-}/NP^{-}} \frac{VP^{-}}{/E}}{\frac{VP^{+}/VP^{-}}{NP^{-}}/E}$$

$$\frac{\text{catch}}{\frac{\text{Sue}}{S^{+}/VP^{+}}} \frac{\frac{\text{catch}}{VP^{-}/NP^{-}} \frac{\frac{\text{ono}}{(TV^{-}VP^{+})/CN^{-}} \frac{\text{armadillos}}{CN^{-}}}{VP^{-}/NP^{-}} \frac{VP^{\ominus}}{VP^{\ominus}} **/E}$$

In summary, this way of treating Negative Concord (i.e. the hypothesis that concord marking uniformly indicates $\downarrow M$ context but is indifferent to its own negation status) predicts without any further apparatus that:

- In a "chain" of neg-marked words, the one with logically widest scope must have a \(\mathbb{M} \) interpretation
- All other neg-marked words in the chain must have NPI interpretations.
- (Consequently,) no logically double negations are possible.

3.3 Variations in NC Dialects:

Ladusaw (1991, 1992) observes that among negative concord dialects there is quite a bit of variation in the positions in which negative morphemes are allowed (and the readings produced by them); four variants from American dialects (which turn out to have correspondents among the Romance languages) are listed in (42) below. Here, a "*" indicates the sentence is ungrammatical, a "1" indicates the sentence is grammar on a logically "single negation" reading, and a "2" indicates a sentence which, if grammatical at all, can only be interpreted as logical double negation:

(42)			$NC-A_1$	$NC-A_2$	NC-B	NC-C
	a.	No dogs chased Felix	1	1	1	*
b. c. d. e.		No dogs didn't chase Felix	*	2	1	*
	c.	No dogs didn't chase no cats	*	2	1	*
	d.	No dogs chased no cats	1	1	1	*
	e.	Fido didn't chase no cats	1	1	1	1
	f.	Fido chased no cats	*	*	*	*

The meanings expressed by (a-b) and (c-d) in dialect NC-B are expressed instead in dialect NC-C by (f) and (g) respectively:

- (42) g. Didn't no dogs chase Felix.
 - h. Didn't no dogs chase no cats.

It turns out that all but one of these dialects can be described in the system proposed above by deleting one of the two interpretations (true negation or NPI) of one of the neg-words; this possibility depends on the fact that subject and object NPs (and therefore determiners) belong to separate categories in the categorial syntax employed here (an analytic step which does have some independent motivation in this syntactic framework, to be sure). Before it is objected too loudly that appeal to varying lexical categorizations may be ad hoc and inappropriate, the reader should be reminded that categorial grammar is after all a "radical lexicalist" theory (to use Karttunen's famous phrase), one in which all the syntax of the grammar resides in the syntactic categories and their structure (there are no independent phrase structure rules, except for slash elimination itself, and no movement rules or other independent syntactic constraints), and only by assignment of lexical items to various categories differentially is their behavior differentiated at all; thus this is in fact a rather restricted method of describing dialect differences such as those in (41).

The version of negative concord implicitly described in the rules given earlier is a kind of hypothetical, very general one, in that even the sentence-type (f), ungrammatical in all American dialects, would be generated. However, Ladusaw notes (1992:249, note 10) that Allesandro Zucchi reported that his own dialect of Italian, a NC language, allowed pattern (f) as well as the others.

To eliminate sentence pattern (f) from Dialect NC-B (and from all other NC dialects as well—except Zucchi's Italian), it suffices to delete the negative interpretations, and of course lexical entries in the syntactic categories for these, of non-subject **nobody** and **no**, leaving only their NPI interpretations. This pattern is also said to be found in Catalan.

To describe Dialect NC-A₁, we need to eliminate sentence types (b) and (c), and this is accomplished by deleting the NPI readings of neg-marked auxiliaries. The pattern is matched by Spanish and standard Italian.

As Dialect NC-A₂ is described by Ladusaw (1991), this dialect is supposed to allow double-negation readings for (b) and (c) patterns (and is otherwise like NC-A₁), but as already noted, the present system predicts that genuine double-negation readings are absent from negative-concord languages. Indeed, Ladusaw elsewhere (1992) takes the absence of such readings to be a deep generalization in need of explanation (also as noted), but if this observation of NC-A₂ is correct, then no such generalization holds. Such readings are perhaps an instance of metalinguistic negation, a possibility suggested by the examples of alleged "real" double negation in NC-dialects mentioned briefly by Labov (1992). The negative concord Romance languages definitely do allow such readings, it is now well-known, a problem to which we will return in the last section.

In Dialect NC-C, a neg-marked NP may not precede a neg-marked auxiliary: cf. (a)-(d) vs. (g), (h). To describe this, we delete negative interpretation of subject **nobody**, no (as well as the negative interpretations of auxiliaries deleted for dialect NC-A₁), and add "inverted" **didn't** of type $(S^+/(S^-/VP^-))/VP^-$ (i.e. having an interpretation that insures wide scope over the subject NP). In other words, in this dialect neg-marked NPs have only NPI interpretations and auxiliaries have only negative interpretations; with the assumption that an aux can take scope over a quantificational subject NP only if it precedes it, the hypothesis that NPI "negatives" consistently mark \downarrow M contexts, takes care of the rest.

Before concluding this section, I will mention one interesting argument relating to the choice between the present hypothesis and Ladusaw's (1992) theory of NC (that neg-marked words in NC languages are always indefinites in the sense of Discourse Representation Theory, and that the real logical negation in a negative sentence is abstract). It has often been observed for English (e.g. in the discussion of "free-choice" any vs. NPI any, cf. Carlson (1981)) that adverbs like practically, almost, virtually, nearly, absolutely can modify universal and (sufficiently large) numerical quantificational NPs, (cf. Almost everybody likes him, Practically every teenager adores him, Nearly 1000 people voted for her), including negative ones (Absolutely nobody goes there anymore), but not indefinites (cf. *Almost several people like him, *Absolutely most people like him, *Practically a teenager or two like him). The same is

apparently true in French (practiquement), according to van der Wouden and Zwarts (1993), which allows negative concord. They have observed, furthermore, that while such an adverb can modify the first or C-commanding NP in a multiply-negated sentence, it cannot modify subsequent neg-marked NPs:

- (43) a. Pratiquement personne n'a rien dit Practically nobody not has nothing said 'Practically nobody said anything'
 - b. *Personne n'a pratiquement rien dit Nobody not has practically nothing said

This is predicted immediately by analyses in which the C-commanding negation in a NC sentence is indeed a real logical negation while other neg-marked words are NPIs, but does not follow (at least, not without further explanation and/or assumptions) in a theory in which all neg-marked words in such sentences are semantically indefinites.

4 "Doubly-Licensed" NPIs: A Problem for the ↓M-Inference Hypothesis

A fact about NPI-licensing which seems to be not widely known, but is I believe very significant, is that NPIs are acceptable when in the scope of not just one but also two or more \downarrow M expressions. Examples where a conditional antecedent provides one trigger (44a, 45a) and a second may be added (44b, 45b) were cited by Hoeksema (1986). However, other kinds of cases besides conditionals can readily be constructed where a second trigger can be added (46,47), and Chris Barker has observed the actually occurring example (48):

- (44) a. If he knows anything about logic, he will know Modus Ponens.
 - b. If he doesn't know anything about logic, he will (still) know Modus Ponens. (Hoeksema, 1986)
- (45) a. If he lifts a finger, fire him.
 - b. If he doesn't lift a finger, fire him. (Hoeksema, 1986)
- (46) a. She very rarely eats anything at all for lunch.
 - b. She very rarely doesn't eat anything at all for lunch.
- (47) a. No one except John Q. Public pays any income tax (anymore).
 - b. No one except Leona Helmsley doesn't pay any income tax (at all).
- (48) Don't let her die without lifting a finger! (Sign on a video arcade game, to encourage the potential customer to rescue the heroine; observed by Chris Barker, p.c.)

This situation contrasts with that of positive polarity items, where it has been known since Baker's research in the 1970's that the combination of two $\downarrow M$ functors recreates a "positive" environment in which PPIs are again happy (Ladusaw 1979):

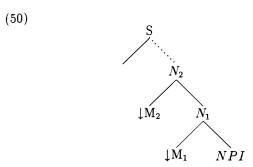
- (49) a. I would rather be in Florida right now
 - b. ??I wouldn't rather be in Florida right now
 - c. You can't convince me I wouldn't rather be in Florida now.

In other words, PPIs seem to follow the (logically justified) rule that two $\downarrow M$ functors combine to produce a $\uparrow M$ context, while NPIs do not.

The theory of Ladusaw (1979) is actually constructed so as to predict that a NPI is licensed iff it is in the scope of at least one (i.e. thus also of two or more) $\downarrow M$ expression(s) and so is consistent with such data. But insofar as two $\downarrow M$ expressions combine to yield a $\uparrow M$ logical context, that theory therefore claims that logical $\downarrow M$ context per se is not the relevant factor for NPI licensing (contrary to its usual characterization as a "purely semantic" account of NPI licensing context). The theory of this paper implies that logical $\downarrow M$ context (relative to the whole sentence in which a word occurs) should always be the determining factor, hence does not predict such data.

A simple conclusion to draw at this point would be that the condition "in the scope of (a) \downarrow M functor(s)" in Ladusaw's (1979) account is the empirically correct licensing condition, whereas the sentence-based definition of \downarrow M context derived from Sánchez's logic is not, and that is the end of the matter. However, I believe there are reasons to be less than fully satisfied with this position and to give a good deal more thought to the problem.

This data actually shows that a unique, syntactically-characterized logical form for each sentence must be appealed to in Ladusaw's account: a purely semantic definition of $\downarrow M$ context will not do, because in pure semantic terms there is no distinction between an argument in the scope of no $\downarrow M$ functions and one in the scope of exactly two $\downarrow M$ functions, any more than there is a distinction in the semantic behavior of ϕ vs. $\neg \neg \phi$ in classical logic. In order to determine whether an NPI is licensed "in terms of a semantic definition of downward monotonicity" in a sentence such as that indicated by the tree sketched in (50) for example, we must still determine for (the model-theoretic interpretation of) each node N_1 , N_2 , etc. whether the NPI is licensed with respect to that node: if we should happen to skip node N_1 and check only N_2 and higher nodes, then the NPI would fail to meet the licensing condition:



One illuminating way to see the consequences of this is to consider the Lambek-calculus, which is "structurally complete" in the sense that if a string is well-formed on one bracketing, it is well-formed on all possible bracketings and would receive the same semantic interpretation on each bracketing (baring the complications of non-concatenative operations introducing quantifier scope ambiguities, etc.). If however two $\downarrow M$ functions such as $\downarrow M_2$ and $\downarrow M_1$ in a tree like the one above were combined by functional composition (i.e. the Lambek equivalent) before being applied to their argument NPI, then their argument would not meet Ladusaw's licensing condition on this derivation. As far as I know, this would thus be the only case where alternative Lambek-derivations of the same string could have different "semantic" properties.

This realization also shows that the set of " \downarrow M contexts" relevant for NPI licensing cannot be so straightforwardly identified on empirical on semantic grounds, independently of NPI data, as is generally assumed: in fact, speakers of natural languages can only identify entailments from full sentences, not from "local" subparts of arbitrary full sentences (which is what Ladusaw's account would really need). So out of a series of sentences (i) Mary caught every unicorn (ii) Mary didn't catch every unicorn (iii) It is not the case that Mary caught every unicorn, our intuitions tell us that the CN argument of every only supports \downarrow M inferences in (i) and (iv), not (ii) and (iii). Such a superficial 'inconsistency' in empirical observation can of course be easily explained in a theory in which the property of permitting \downarrow M inference reverses when two \downarrow M functors combine. But it is that theoretical conception of \downarrow M that leads us to the divergence between NPI distribution and the theory of \downarrow M context we have observed, since NPIs would be licensed in this CN argument in all four cases.

I do not have a solution to this anomaly. It is easy to make speculations about the ultimate "explanation" of double licensing, perhaps ones deserving investigation: Natural logic might involve intuitionistic rather than classical negation⁸, which would mean in terms of the present logic that we should not assume both cancellation rules $A^+/B^- + B^- \Rightarrow A^+$ and $A^-/B^+ + B^+ \Rightarrow A^-$ but only the former, replacing the latter instead with $A^-/B^{--} + B^{--} \Rightarrow A^-$, and putting NPIs not only in B^- but also B^{--} , etc. Another possibility is

⁸This possibility was in fact suggested to me by Bill Ladusaw (p.c.), which is not to say that he advocates it seriously.

that the (well-known) sentence processing difficulty of sentences with multiple $\downarrow M$ contexts, together with the relative infrequency of occurrence of such cases (and therefore little real need to be able to deal with them "logically") led to the grammaticization of the "no licensing reversal" rule. Yet another is that "local $\downarrow M$ context"-marking serves some useful processing function, perhaps in computing an intermediate partial compositional interpretation of the sentence in advance of putting the whole sentence meaning together, a processing having no counterpart in deductive systems such as Sanchez'. But these are only guesses.

4.1 A Double-Marking System for Describing both Syntactic and Logical NPI/NC Contexts

Rather, what I will propose here is a system which simply takes this discrepancy as given. a system which distinguishes LOGICAL POLARITY from SYNTACTIC POLARITY and computes them in parallel, in the same compositional way. We will keep the "+/-" superscripts for logical polarity and add subscripts for syntactic polarity:

(51) Let the notation: A^{α}_{β} stand for a category A which has (i) LOGICAL polarity α and (ii) SYNTACTIC polarity β .

For example, A_-^-/B_+^+ would be a logical-polarity reverser that also cancels NPI licensing, whereas A_-^-/B_-^+ is a logical polarity reverser that preserves NPI licensing, i.e. allows its argument to contain NPIs even though it is no longer a logically $\downarrow M$ context.

This double-marking system is not intended as a 'theory' of NPI/NC licensing in the proper sense, but a descriptive framework within which to investigate the interaction of logical and syntactic licensing and to develop a linguistic theory. The goal of and motivation behind such as systems is this: the NPI-as-inference-marker hypothesis predicts that logical and syntactic marking should ideally coincide, so divergences should be the exception (as they apparently are). When the two DO diverge, under exactly what conditions does this happen, and, what explanations for the divergence can be found in each case? One hopes that most or all of these divergences will turn out to be explainable as consequences of other semantic or syntactic characteristics of natural languages, and the double-marking system should serve as a temporary framework for discovering these. If however, some syntactic NPI licensing contexts turn out to be "grammaticized" variants of \$\pextstyle M\$ licensing, then no superior format may every exist for describing these than one of this kind.

Note that some basic assumptions of categorial grammar are preserved: both syntactic and logical polarity properties will have to be projected "lexically", i.e. from (lexical) categories, without general constraints over kinds of syntactic structures per se, but with function/argument organization always implicit.

As a first illustration of this framework, I formulate the NPI theory of Ladusaw (1979) (and approximately that in Dowty (1993)), i.e. "irreversible" syntactic NPI-licensing:9

(52) Words in general are entered in one of two category sets:

Polarity-Preservers: Polarity-

$$\left\{ \begin{array}{l} A_{+}^{+}/B_{+}^{+} \\ A_{-}^{-}/B_{-}^{-} \\ A_{-}^{+}/B_{-}^{+} \end{array} \right\} \quad =_{def} \quad A/B \qquad \left\{ \begin{array}{l} A_{+}^{+}/B_{-}^{-} \\ A_{-}^{-}/B_{-}^{+} \\ \left[A_{-}^{+}/B_{-}^{-} \right] \end{array} \right\} \quad =_{def} \quad A/^{R}B_{-}$$

except for:

$$\begin{cases} A_{-}^{-}/B_{-}^{-} \\ A_{-}^{+}/B_{-}^{+} \end{cases} =_{def} A_{-}/B_{-} \begin{cases} A_{+}^{+}/B_{+}^{+} \\ A_{-}^{+}/B_{-}^{+} \end{cases} =_{def} A^{+}/B^{+}$$

The notation introduced for names of category sets is intended to reflect the defaults that (i) montonicity preservation is the default over monotonicity reversal; (ii) whether syntactic polarity is preserved or reserved is, by default, determined by logical polarity (presered vs. reversed), i.e. the category must otherwise be explicitly marked; (iii) an expression can occur in either positive or negative polarities (both logical and syntactic varieties) unless marked otherwise. An indirect consequence of these defaults is that syntactic polarity will typically match logical polarity except for expressions in the scope of a functor that has explicitly introduced a divergence. The category of independent sentences is, naturally, S_+^+ .

4.2 Describing Classes of Words/Environments that Affect Licensing Differentially

For another application, we turn to a kind of NPI "Island", an environment that never allows NPIs to licensing from outside—namely the "Manner-of-Speaking" verbs, illustrated below along with a category assignment that will give them this behavior:

- (53) a. Sue whispered (shouted, murmured, etc.) that armadillos were approaching.
 - b. *Sue whispered (shouted, murmured, etc.) that any armadillos were approaching.

⁹I enclose the third category on the right in brackets simply because this would have to be the category of a ↓M expression occurring inside the scope of two more ↓M expressions, and while I know of no reason to rule this possibility out, my attempts to construct sentences with three ↓M expressions and determine their embedded NPI-licensing properties are a shaky basis for claiming this category is needed.

c. *Sue didn't whisper (shout, murmur, etc.) that any armadillos were approaching.

d. Preserv.:
$$\begin{cases} A_{+}^{+}/B_{+}^{+} \\ A_{-}^{-}/B_{+}^{-} \\ A_{-}^{+}/B_{+}^{+} \end{cases} =_{def} A/B_{+} \quad Revers.: \begin{cases} A_{+}^{+}/B_{+}^{-} \\ A_{-}^{-}/B_{+}^{+} \end{cases} =_{def} A/_{R}B_{+}$$

This case is a good candidate for a potential semantic explanation of the licensing failure, to be sure. NPIs are not licensed inside \downarrow M reports of quoted speech acts (cf. Mary didn't say that John ate anything vs. *Mary didn't say, "John ate anything"), and these verbs have meanings that are connected, unmistakably, with such a report, both because their meaning involves describing the auditory properties of the speech act as well as its content, and because of other syntactic characteristics shared with quote complements (e.g. resistance to WH-extraction).

Negative Concord languages also provide examples of islands to concord licensing. Whereas English allows NPIs to be licensed inside a relative clause, most American English NC dialects do no allow concord in this position. This restriction is all the more noteworthy, because certain NC dialects exist which do, however, license concord here. A often-cited example from such an exceptional NC dialect is (54) from Labov (1972):

(54) Ain't no cat can't get in no coop.
Lit.: '(There) ain't no cat (that) can't get in no coop'
Actually means in this dialect: "There isn't a cat that can get in a coop."

As Labov notes, such examples sound bizarre or unintelligible even to speakers of English who have passive knowledge of the more prevalent NC English dialects. A category assignment for a relative pronoun **that** (assuming as in many recent categorial analyses that relative clauses and their "gaps" can be treated simply as of category S/NP is (55)

(55) that (Most NC dial.)
$$\left\{ \frac{(CN_+^+ \backslash CN_+^+)/(S/NP)_+^+}{(CN_-^- \backslash CN_-^-)/(S/NP)_+^+} \right\}$$

$$=_{def} (CN \backslash CN)/(S/NP)_+$$
that (Rare NC dial.)
$$\left\{ \frac{(CN_+^+ \backslash CN_+^+)/(S/NP)_+^+}{(CN_-^- \backslash CN_-^-)/(S/NP)_-^-} \right\}$$

$$=_{def} (CN \backslash CN)/(S/NP)$$

As mentioned earlier, Romance NC languages, unlike American English NC dialects, allow true double negation readings in many cases, so that many sentences with two negative-marked constituents are ambiguous between a NC and a double-negation reading, such as French sentences like this one:

- (56) Personne n'a rien dit
 - (i) "Nobody said anything"
 - (ii) "Nobody said nothing" (= "Everybody said something")

The double-marking system (unlike the analysis of NC presented earlier) offers a way of analyzing such cases—in fact, they would sharing the properties of 'pure' NC languages (that the 'same' morphemes serve simultaneously as $\downarrow M$ functors and also as simply indicators that their argument is in a $\downarrow M$ context, without denoting a montonicity-reversing functions themselves) with the properties of NPI-licensing in English (namely, "irreversible" syntactic licensing). This is accomplished by giving such words this sort of dual categorization:

By assigning negative words with $\downarrow M$ interpretation¹⁰ the first category set, including the category A_-^-/B_-^+ not used in NC languages for $\downarrow M$ expressions earlier, we allow it to occur embedded under another genuine negative morpheme and yield a logically double negative reading. With $\uparrow M$ interpretations, negative words occur not in a a category A_-/B_- as before, indicating a logically $\downarrow M$ context, but in category A_-/B_- , they indicate perhaps that their argument is logically $\downarrow M$ or perhaps only that it is C-commanded by a negative trigger (i.e. in double-negation sentences).

Finally, we can observe that the double-marking system turns out to allow a description of the constraint, first observed by Linebarger (1980), on NPI licensing created by the intervention of certain quantifiers between trigger and NPI (in her terms, a sub-case of the Immediate Scope Constraint): whereas (58a) allows the reading (58b) (among other readings, perhaps), the parallel example (58c) with a NPI does not. Similarly, (58d) is grammatical, but (58d) is not; in both cases it is the intervention of the determiner every that is responsible for the difference:

- (58) a. Sue didn't read every book to a student
 - b. $\neg(\forall x : book(x))(\exists y : student(y))[read(Sue, x, y)]$
 - c. Sue didn't read every book to any student
 - d. Sue didn't say Sam caught any armadillos
 - e. *Sue didn't say everybody caught any armadillos

¹⁰To be precise, with anti-additive interpretations; I do not address in this paper the question of a logic for anti-additive functions; cf. however Kas (1993)

This behavior is captured if we make this category assignment for every:¹¹

(59) **every** is assigned to:
$$\left\{ \begin{array}{l} (S_{+}^{+}/VP_{+}^{+})/CN_{-}^{-} \\ (S_{-}^{-}/VP_{+}^{-})/CN_{-}^{+} \\ \left[(S_{-}^{+}/VP_{+}^{+})/CN_{-}^{-} \right] \end{array} \right\} =_{def} (A\backslash B+)/^{R}C_{-}$$

4.3 A Syntactic Perseverance Hypothesis

One benefit of using the double-marking system to investigate divergences between logical and syntactic polarity context is the possibility of observing relationships between the two that might go unnoticed in studies where semantics and syntax and not examined side-by-side in this way. Here is one possible example. Comparison of the case of "unexpected" licensing of NPIs (double \$\frac{1}{2}M\$ context) with this last case of "unexpected" blocking of NPI licensing (Immediate Scope Constraint) reveals a a parallel:

(60) doesn't:
$$\left\{ \begin{array}{l} VP_{+}^{+}/VP_{-}^{-} \\ VP_{-}^{-}/VP_{+}^{+} \end{array} \right\} \quad \text{every:} \left\{ \begin{array}{l} (S_{+}^{+}/VP_{+}^{+})/CN_{-}^{-} \\ (S_{-}^{-}/VP_{+}^{-})/CN_{-}^{+} \end{array} \right\}$$

In both cases, the "-"-marked (result) category differs from the expected one in that the syntactic polarity feature on one (or more) argument category(ies) is, instead, the same as the syntactic feature on the corresponding "+"-marked (result) category (which is the only other choice in a binary system, to be sure): this feature is the VP_{-}^{+} for doesn't, but not only do we find "-" for the CN argument of every but also "+" for its VP argument. One might wonder, in other words, whether the quantifier intervention constraint and the irreversibility of NPI licensing exist ultimately for the same reason. Both don't and every are of course polarity reversers (for one argument). Perhaps for some reason connected with parsing complexities of $\downarrow M$ functors, the syntactic polarity marking of the $\uparrow M$ case "perseveres" in the syntactic polarity of the $\downarrow M$ case, rather than reversing as does the logical polarity:

(61) "Syntactic Perseverance" hypothesis:

In 'hard' cases,
$$\left\{ \begin{array}{c} A_+^+/B_\alpha^\alpha \\ A_-^-/B_\beta^\beta \end{array} \right\}$$
 is replaced by: $\left\{ \begin{array}{c} A_+^+/B_\alpha^\alpha \\ A_-^-/B_\beta^\beta \end{array} \right\}$.

To be sure, one would want further motivations for tying these two cases together before concluding that this is the proper kind of explanation for them—for example, finding further instances of "perseverance". Specifically, we would look for cases, besides *every*, of binary functors that are $\downarrow M$ for one argument, $\uparrow M$ for the other. Though such cases are rare, *not every* may be one. At least Barker (1991) as well as Hoeksema (1986) have discussed the monotonicity and licensing properties of *not every*; these are complicated by the question

¹¹As before, the brackets and question marks merely indicate obscure licensing possibilities whose properties are hard to verify empirically

of the proper constituent structure of a NP like not every student, and quite possibly more than one exists. However, if one possible structure has not every as a complex determiner (whether it is a lexical or complex phrasal unit is immaterial here), then it definitely follows from its semantics that it must be $\uparrow M$ for its first (CN, or Restriction) argument, $\downarrow M$ (VP, or nuclear scope) argument, just the converse of every. This would predict the reverse distribution of NPIs in the two arguments from that of every, and indeed both Barker and Hoeksema report judgments consistent with this possibility, such as:

- (62) a. *Not every child with any sense stole some candy.
 - b. Not every child ever stole any candy. (Barker, 1991)

If not every is a 'hard' case, the Syntactic Perseverance hypothesis predicts its category should be, unexpectedly, not $(S/^RVP)/CN$ or $(S/^RVP_-)/CN$ but $(S/^RVP_-/CN_+)$ (to use the default notation introduced above). That is, that when embedded under a JM functor, its category-schema instance must be $S_{-}^{-}/^{R}VP_{-}^{+})/CN_{+}^{-}$, so it should be a "barrier" to licensing of NPIs within its Restriction argument from a higher \(\psi M \) trigger (as every is a barrier for its nuclear scope) while being an irreversible licenser of NPIs within its nuclear scope. Unfortunately, this prediction is extremely hard to test. The judgments reported above by Barker and Hoeksema are already not overwhelmingly clear for some speakers, and testing the perseverance hypothesis requires embedding such sentences into a matrix clause with its own negation or other \ M trigger (for it is a quirk of not every that it must occur in subject position); even for the most natural-sounding examples I can construct, most speakers balk at trying to parse them at all even without NPIs, much less differentiate NPI licensing in different positions. Hence I omit them here. Hopefully, other ways to test the perseverance hypothesis may someday come to light.

4.4 Summary of Observed Polarity Divergences

With a system of categories that allows any functor four potentially distinct polarity markings for each functor categories (cf. $A^{\alpha}_{\beta}/B^{\delta}_{\gamma}$) there are 2^4 (=16) distinct polarity possibilities, and with assignments of items to sets of polarity-marked subcategories, the number of polarity possibilities grows to the power set of this set—seemingly an embarrassingly rich set of options. Of the various kinds of polarity mismatches discussed (and speculated about) here, however, only the following eight types of category are needed, when the defaults introduced above are employed:

(63)

$$\begin{array}{lll} \textit{Default Name:} & \textit{Cat. Set:} & \textit{Use:} \\ (i) & \textit{A}/\textit{B} & = \{A^{\circ}_{\beta}/A^{\circ}_{\beta}\} & \uparrow \text{ M functors} \\ (iia) & \textit{A}^{R}/\textit{B} & = \{A^{\circ}_{\beta}/B^{-\alpha}_{-\alpha}\} & \downarrow \text{ M funct's (in NC langs.)} \\ (iib) & \textit{A}^{R}/\textit{B}_{-} & = \{A^{\circ}_{\alpha}/B^{-\alpha}_{-\alpha}\} & \downarrow \text{ M funct's (in NPI langs.)} \\ (iii) & \textit{A}_{-}/\textit{B}_{-} & = \{A^{\circ}_{\alpha}/B^{\circ}_{\alpha}\} & \text{NPIs} \\ (iv) & \textit{A}_{-}/\textit{B}_{-} & = \{A^{\circ}_{\alpha}/B^{\circ}_{\alpha}\} & \text{NPI "Islands"} \\ (v) & \textit{A}/\textit{B}_{+} & = \{A^{\circ}_{\beta}/B^{\circ}_{+}\} & \text{NPI "Islands"} \\ (vi) & (\textit{A}/\textit{B}_{+})/^{R}\textit{C}_{-} & = \{(A^{\circ}_{\beta}/B^{\circ}_{+})/\textit{C}_{-}^{-\alpha}\} & \text{every} \\ \\ (vii) & \left\{ \begin{array}{c} \textit{A}/^{R}\textit{B} & \downarrow \textit{M} \\ \textit{A}_{-}/\textit{B}_{-}^{-} & \uparrow \textit{M} \end{array} \right\} & \text{"N-words", Engl. NC, nongoing the context of the context o$$

(Here, I use α and β as variables over $\{+, -\}$, with $-\alpha$ indicating the opposite value of α .) Though I have listed **every** separately, it is of course the combination of a NPI polarity island (v) with a NPI-lang.-type reversing functor (iib). "Pure" NC languages, like American NC English dialects, would not require the separate syntactic polarity marking at all except for their islands to embedding polarity licensing such as relative clauses.

References

Barker, Chris (1991) Possessive Descriptions, University of California Santa Cruz dissertation.

Benthem, Johan van (1991) Logic in Action: Categories Lambdas, and Dynamic Logic, North-Holland.

Carlson, Greg N. (1981) "Distribution of Free-Choice Any", Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society, ed. R.A. Hendrick et al, Chicago Linguistic Society, Chicago, 8-23.

Dowty, David (1993) "Deductive versus Semantic Accounts of Reasoning: the Relevance of Negative Polarity and Negative Concord Marking", Papers from the 29th Regional Meeting of the Chicago Linguistic Society, Chicago Linguistic Society, Chicago.

Dowty, David (in prep.) "Montonicity-Based Logic and Why Natural Languages have Negative Polarity and Negative Concord Marking."

Hoeksema, Jack (1986) "Monotonicity Phenomena in Natural Language", Linguistic Analysis 16.1-2, 235-250.

- Kas, Mark (1993) Essays on Boolean Functions and Negative Polarity, University of Groningen Dissertation, Groningen.
- Labov, William (1972) "Negative Attraction and Negative Concord." Language 48:773-818.
- Ladusaw, William (1979) Polarity Sensitivity as Scope Relation (University of Texas dissertation).
- Ladusaw, William (1991) "Interpreting Negative Concord Structures", Paper presented at the 1991 Annual Meeting of the Linguistic Society of America, January 1991.
- Ladusaw. William (1992) "Expressing Negation", Proceedings of SALT 2, ed. C. Barker and D. Dowty, Ohio State University Linguistics Department, Columbus, 237–260.
- Lakoff, George (1972) "Linguistics and Natural Logic", in Semantics of Natural Language, ed. G. Harmon and D. Davidson, Reidel, Dordrecht, 545-665.
- Linebarger, Marcia (1980) The Grammar of Negative Polarity. MIT dissertation.
- Moortgat, Michael (1990) Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus. (GRASS 9) Mouton-De Gruyter, Berlin.
- Roberts, Craige (1989) Modal Subordination, Anaphora, and Distributivity (University of Massachusetts Diss.), Garland Publications.
- Sánchez Valencia, Víctor (1991) Studies on Natural Logic and Categorial Grammar, University of Amsterdam dissertation.
- van der Wouden, Ton, and Franz Zwarts (1993) "A Semantic Analysis of Negative Concord", SALT III, ed. U., Lahiri and Adam Wyner, Cornell University Linguistics, Department, Ithaca.
- Zwarts, Franz (1993) "Three Kinds of Negative Polarity", ms., Groningen, to appear in *Plural Quantification*, ed. F. Hamm and E. Hinrichs.