Reasoning *with* Categorial Grammar Logic

Raffaella Bernardi

Universiteit Utrecht – NL
Università 'G. D'Annunzio' – Chieti, Italy
bernardi@let.uu.nl

Abstract

The article presents the first results we have obtained studying natural reasoning from a proof-theoretic perspective. In particular we focus our attention on monotonic reasoning. Our system consists of two parts: (i) A Formal Grammar – a multimodal version of classical Categorial Grammar – which while syntactically analysing linguistic expressions given as input, computes semantic information (In particular information about the monotonicity properties of the components of the input string are displayed.); (ii) A simple Natural Logic which derives (monotonicity) inferences using as vehicle the parsed output. The monotonicity markers assigned in the lexicon are propagated through the proofs via a combination of the structural and the logical rules for the unary operators of Multimodal Categorial Grammar (MMCG) [Moo97]. We have chosen to work with an expressive 'grammar logic', in order to avoid the use of extra-logical marking devices and extra-logical structural reasoning. Having MMCG as parser, our system is able to make the derivations simply within the logic. This new approach makes the implementation of the theory an easier task. We have implemented the theoretical results, so far obtained, using Grail, a theorem prover for Categorial Grammar Logics [Moo98].

1 Introduction: Natural Logic

The task of accounting for the role of language in drawing inferences is commonly considered to belong to the domain of formal semantics, i.e., the study of the *meaning* of natural language expressions through *formal logics*. Most of the literature in natural reasoning assumes a *model-theoretic* perspective using a formal language as an intermediate step in which natural language expressions are translated.

In the paper we assume a new perspective: instead of using logic forms as vehicles of inference, *natural* language expressions can be used directly, and instead of giving the derivation via an interpretation into a model, an inference is read off the derivation. The system so obtained is a Natural Logic.

A similar approach has been taken in [SV91], where linguistic expressions are analyzed by a non-directional Categorial Grammar (i.e., Lambek calculus with Permutation, LP). Since LP cannot account for the monotonicity marking on its own it is enriched with a marking algorithm. As we will see in the next sections, Multimodal Categorial Grammar (MMCG), instead, is a 'grammar logic' with the right expressiveness. It is able to carry out the required tasks, simply via the logical and the structural rules, without using any extra-logical devices.

Moreover, due to the fact that in natural reasoning the felicity of the inferences is determined both by the syntactic and by the semantic properties of the linguistic expressions, the well-known property of MMCG of tightly integrating these two levels is fundamental.

The paper gives a brief introduction to Categorial Grammar (CG) explaining the needs for the multimodal system on which our Natural Logic is based. Then a fragment of Natural Logic is discussed.

2 Multimodal Categorial Grammar

2.1 Classical Categorial Grammar

In order to make the article self-contained, before going into the details of MMCG, we briefly introduce

CG, refering the reader to [OBW88] for a more detailed overview and to [AC96] for more recent results.

Classical CG is a language recognition device first described in Ajdukiewicz [Ajd35]. It is based on a finite set of primitive categories and on complex categories recursively built. In its original presentation the latter were denoted by the fraction notation $\frac{A}{B}$; the combination of the category $\frac{A}{B}$ and B, was marked by $\frac{A}{B}$ x B or B x $\frac{A}{B}$, equal to A.

Bar-Hillel [BH64] split the $\frac{A}{B}$ into $B\backslash A$ and A/B, in order to discriminate between complex expressions which will produce an expression of category A when composed with an arbitrary expression of category B to the left, and to the right, respectively.

The notation \, /, derives from [Lam58]. In this paper important steps has been taken to achieve a logical calculus for natural language: (i) a completely formalized logical language and (ii) a proof-system of/for natural language are given.

The logical language (cf. i) is enriched with the product (composition) operator $\bullet$, which makes it possible to explicitly refer to the combination of categories.

As concerning the proof-system (cf. ii), the new step towards a 'Grammar Logic' consists on the introduction of the *hypothetical reasoning* using Gentzen-style natural deduction format. For the sake of simplicity in this subsection we present the Lambek calculus (L) informally. The formal format will be used describing MMCG.

[1a] If A $\in \alpha/\beta$, and B $\in \beta$,, then AB $\in \alpha$;	[1b] If AB $\in \alpha$ and B $\in \beta$, then A $\in \alpha/\beta$;

i.e., $X \bullet Y \to Z$ iff $X \to Z/Y$

[2a] If A $\in \alpha$, and B $\in \alpha\backslash\beta$, then AB $\in \beta$;	[2b] If AB $\in \beta$, and A $\in \alpha$, then B $\in \alpha\backslash\beta$;

i.e., $X \bullet Y \to Z$ iff $Y \to X\backslash Z$

The first rule says that if an expression B of category β and an expression A of category α/β are given, then the composition of A and B is of category α.

This logic characterizes some fragments of natural language. The system proves the grammaticality of sentences starting from the lexicon (where categories are assigned to single words), and applying the logical rules. Here follows a simple example. Let n (for 'common noun'), s (for 'sentence') and np (for 'noun phrase'), be the set of primitive categories,

Example 2.1 *Sara reads the book*

Proof

(1) book $\in n$	lexical entry
(2) the $\in np/n$	lexical entry
(3) reads $\in (np\backslash s)/np$	lexical entry
(4) Sara $\in np$	lexical entry
(5) the book $\in np$	(2), (1); [1a]
(6) reads the book $\in np\backslash s$	(3), (5); [1a]
(7) Sara reads the book $\in s$	(6), (4); [2a]

The proof starts with the lexical entries assigned to the words present in the given string, then applies the schema [1a] instantiating the premise $A \in \beta$ with 'book $\in n$' and the premise $B \in \alpha/\beta$ with 'the $\in np/n$'. Thus the proof goes on in this way till the input is proved to be of category s.

As we have said, an important logical aspect introduced by Lambek is the hypothetical reasoning. An example will clarify the meaning and the use of it.

Example 2.2 *The book that Sara reads*

Proof. Let (1), (4) as before,

(5) that $\in (n\backslash n)/(s/np)$	lexical entry
(6) x $\in np$	hypothesis
(7) reads x $\in (np\backslash s)$	(3), (6); [1a]
(8) Sara reads x $\in s$	(7), (4); [2a]
(9) Sara reads $\in s/np$	(8), (6); [1b]
(10) that Sara reads $\in n\backslash n$	(5), (9); [1a]
(11) book that Sara reads $\in n$	(10), (1); [1a]
(12) The book that Sara reads $\in np$	(2), (11); [1a]

In order to prove that the given string is of category np, the introduction of the hypothesis 'x' and the elimination of it (via [1b]) are needed (i.e. the hypotethical reasoning). Doing so we manipulate the functional application's order. If we make this explicit via brackets, we have:

(8) Sara (reads x)
(8') (Sara reads) x
(9) Sara reads

The right-associativity property, [P1], is required. However, allowing this structural property *globally*, will lead to both undergeneration and overgeneration problems. A relative clause such as 'the book that Sara gives to Noradin', would still be underivable: [P1] makes accessible only the right-peripheral position. As we will better explain in the last section a *controlled* form of [P1] will solve the problems.

2.2 Multimodal Categorial Grammar

In the last decade new versions of the formalism we have described, has been developed. In this paper we work with Multimodal Categorial Grammar (MMCG) [Moo97] which overcomes some expressive limitations of the Lambek calculus.

Following the insight that languages are different from each other basically in the way they structurally realize the form/meaning correspondence, MMCG consists of two independent parts: (i) a base logic, in which no structural rules are available; and (ii) packages of structural rules that can be lexically controlled via unary operators, $\diamond$ and $\Box^{\downarrow}$, in addition to the classical binary ones. The base logic is meant to capture invariants of grammatical compositions shared by all languages. It is a variation of the Lambek calculus described in the previous subsection, enriched with the logical rules for the unary operators and dispossessed of the structural rules. However, these rules are reintroduced in the structural packages which, instead, are language-specific. Moreover, the control on the structural reasoning is made possible by means of *modes* which share the base logic, but (can) differ in their structural properties.

The logical rules of MMCG follow. To make the comparison with their use in the coming examples easier, we present the rules in Sequent Natural Deduction form.

Logical Rules

Logical Rules for the binary operators

$$\frac{\Gamma \vdash A/_iB \quad \Delta \vdash B}{(\Gamma \circ_i \Delta) \vdash A}\,[/_iE] \qquad \frac{(\Gamma \circ_i B) \vdash A}{\Gamma \vdash A/_iB}\,[/_iI]$$

$$\frac{\Gamma \vdash B \quad \Delta \vdash B\backslash_i A}{(\Gamma \circ_i \Delta) \vdash A}\,[\backslash_iE] \qquad \frac{(B \circ_i \Gamma) \vdash A}{\Gamma \vdash B\backslash_i A}\,[\backslash_iI]$$

$$\frac{\Delta \vdash A \bullet_i B \quad \Gamma[(A \circ_i B)] \vdash C}{\Gamma[\Delta] \vdash C}\,\bullet_iE \qquad \frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma \circ_i \Delta \vdash A \bullet_i B}\,\bullet_iI$$

These rules are equivalent to the ones given in the previous subsection: the Elimination rules correspond to the (a)s, the Introduction rules to the (b)s.

A deduction of the category A from the combination of the structures Γ and Δ is denoted by the sequent $(\Gamma \circ_i \Delta) \vdash A$. Structures are built up from formulas with the binary operation $(..\circ_i..)$, structural counterpart of $\bullet_i$. The 'i' indicates the mode of the combination. Finally, $\Gamma[B]$ means that the structure Γ contains a distinguished occurrence of the formula of category B.

Having these reading keys, the $/_iE$ says that if from the structure Γ derives the category $A/_iB$ and from the structure Δ the category B, then the composition of the two structures is A.

Besides the logical rules for the classical operators, MMCG consists of Introduction and Elimination rules for the unary ones:

Logical Rules for the unary operators

$$\frac{\Delta \vdash \diamond_i A \quad \Gamma[\langle A\rangle^i] \vdash B}{\Gamma[\Delta] \vdash B}\,[\diamond_iE] \qquad \frac{\Gamma \vdash A}{\langle\Gamma\rangle^i \vdash \diamond_i A}\,[\diamond_iI]$$

$$\frac{\Gamma \vdash \Box_i^{\downarrow} A}{\langle\Gamma\rangle^i \vdash A}\,[\Box_i^{\downarrow}E] \qquad \frac{\langle\Gamma\rangle^i \vdash A}{\Gamma \vdash \Box_i^{\downarrow} A}\,[\Box_i^{\downarrow}I]$$

At this point the reading of these rules might be clear, the only thing that should be noticed is that $\langle\cdot\rangle^i$ is the structural operator corresponding to $\diamond_i$.

To complete the picture we give an example of structural rules:

Structural Rules

$$\frac{[\Delta_1 \circ_i (\Delta_2 \circ_j \Delta_3)] \vdash C}{[(\Delta_1 \circ_i \Delta_2) \circ_j \Delta_3] \vdash C}\,[P1]$$

$$\frac{[(\Delta_1 \circ_i \Delta_2) \circ_j \Delta_3] \vdash C}{[\Delta_1 \circ_i (\Delta_2 \circ_j \Delta_3)] \vdash C}\,[P2]$$

As is evident from the rules, the structural inferences effect only the structure of the sentence, leaving the category C unchanged. Moreover, they are allowed only for the interaction of specific modes – i, j.

In this section we have seen how the Formal Grammar, on which the Natural Logic is based, syntactically analyses linguistic expressions. In the next we describe how it is able to carry the semantic information involved in deriving monotonicity inference in parallel with this parsing.

3 A Fragment of Natural Logic

3.1 Monotonicity Reasoning

Natural language inference turns out to cover many forms of reasoning of quite different complexities

[Ben87]. In this article we focus our attention on monotonicity reasoning, which is a widespread phenomenon in natural reasoning.

The idea on which monotonicity reasoning is based is the following definition:

Definition 3.1 *Let $f : A \to B$ be a function and $\leq_A$, $\leq_B$ be two partial orders on A and B respectively, then*

a. f is an increasing monotonic function *iff $\forall x, y \in A$, if $x \leq_A y$ then $f(x) \leq_B f(y)$;*

b. f is a decreasing monotonic function *iff $\forall x, y \in A$, if $x \leq_A y$ then $f(y) \leq_B f(x)$;*

In other words, an increasing (resp. decreasing) function preserves (resp. inverts) the partial order of its arguments.

The application of the above definition to natural language can be understood considering some examples:

1. Mary reads an *interesting newspaper* ⇒ Mary reads a *newspaper*.
2. No boy bought a *newspaper* ⇒ No boy bought an *interesting newspaper*.
3. Mary *carefully listens* to the news ⇒ Mary *listens* to the news.
4. No boy *listens* to the news ⇒ No boy *carefully listens* to the news.
5. Mary doesn't *carefully listen* to the news ⇒ Mary doesn't *eagerly and carefully listen* to the news.

The inferences all involve substituting an expression by an expression the denotation of which is a superset (see (1) and (3)) or a subset (see (2), (4) and (5)) of the denotation of the original expression. We will denote this inclusion relation by [[P]] ≤ [[Q]], i.e., P denotes a subset of the set denoted by Q, and we summarize the described behavior with the following inference schema:

If [[P]] ≤ [[Q]] and N is a monotonic context, then

$$\frac{N[Q]}{N[P]}\ (a) \qquad \text{or} \qquad \frac{N[P]}{N[Q]}\ (b)$$

The examples (1) and (3) instantiate (*b*), while (2), (4) and (5) exemplify (*a*).

As we have anticipated in the introduction, building a fragment of Natural Logic (restricted to monotonicity reasoning), means to have a logic which derives (monotonicity) inferences using natural language expressions. The idea of using a formal grammar as basis for such a system has been inspired by V. Sanchez [SV91], who has worked with a non directional CG as parser. The choice of having this formal grammar as basis of his Natural Logic, does not give the system any logical control on the sentence structure and the word order. Natural language, however, is sensible to these two aspects, which also have effect on (monotonicity) inferences. Furthermore, his choice has required extra-logical devices for the monotonicity markers propagation, namely a three-step algorithm.

Dowty [Dow94] proposes a solution to this last problem, using a simple labeled grammar. See [Ber99] for a comparison of these two approaches with the proposal here described. In the paper we show that having MMCG as parser makes it possible to account for both the monotonicity and the structural reasoning simply within the logic.

As we have seen (MM)CG assigns to each word either a primitive (*complete*) category or a complex (*incomplete*) category. The distinction, *complete* vs. *incomplete* categories, has an important role in our framework: the latter are seen as functions $\alpha\backslash\beta$ and β/α which take argument of category α and yield values of category β. This makes the link with monotonicity reasoning possible. The property of CG of assigning a functor-argument structure to arbitrary sentences, allows the application of Definition 3.1 to natural language. In [Zwa86] detailed tests are given for establishing these properties of single words.

Moreover, using semantic information we can assume a certain partial order among linguistic expressions (e.g., [[big animal]] ≤ [[animal]]). Therefore, we have all the information required to apply the above definition and derive the right inference.

However, we still miss an appropriate formalization displaying the monotonic positions, both on the lexical entries and on the parsed output. Definition 3.1 entails that the argument of an increasing monotonic function is in an increasing monotonic position whereas the argument of a decreasing monotonic function is in a decreasing monotonic position. We will use this fact to assign the markers in the lexicon. What we want to achieve is to have a parser which automatically gives an output with this information marked so that the expression in an increasing (resp. decreasing) position can be replaced with denotationally bigger (resp. smaller) expression. In the next subsection is shown that MMCG is a suitable tool for this task.

3.2 Monotonicity Reasoning with MMCG

As we have just mentioned, in order to produce a monotonicity marking in the course of the derivation, first of all, we need to display this information at the lexical entry level. Furthermore, we have to propagate the markers from the lexical entries of our "grammar logic" to the final output of the parsing.

We use '+' and the '–' as monotonicity markers, for the increasing and the decreasing monotonic position, respectively. Reflecting the fact that words in functor position are always in an increasing monotonic position, functors are decorated with $\Box^{\downarrow}_{+}$. This will keep track, via the $[\Box^{\downarrow}_{+}E]$ rule of the lexical entries which figure as functors in the proof. The propagation of monotonicity properties from a functor to the words that function as its arguments is given by the logical rules for the $\Diamond$ operator.

One of the advantages of assuming a direct proof-theoretic perspective in the study of natural reasoning is that the Natural Logic thus obtained can be implemented more easily. The system described in this paper has been implemented in Grail, an automated theorem prover for CG Logics developed by Moot [Moo98]. We will use its proof format in the presentation of the examples below.

When attempting to prove (the grammaticality of) a sentence, Grail starts from the lexical entries stored in a database and applies the logical and structural rules of the system.

Example 3.1 *John walks*

Lexical entries:

$\texttt{John} \vdash np$
$\texttt{walks} \vdash \Box^{\downarrow}_{+}(\Diamond_{+}np\backslash s)$

Derivation

$$\dfrac{\dfrac{\texttt{John} \vdash np}{\langle\texttt{John}\rangle^{+} \vdash \Diamond_{+}np}\,[\Diamond_{+}I] \quad \dfrac{\texttt{walks} \vdash \Box^{\downarrow}_{+}(\Diamond_{+}np\backslash s)}{\langle\texttt{walks}\rangle^{+} \vdash \Diamond_{+}np\backslash s}\,[\Box^{\downarrow}_{+}E]}{\langle\texttt{John}\rangle^{+} \circ \langle\texttt{walks}\rangle^{+} \vdash s}\,[\backslash E]$$

As we can see from the proof, rule $[\Diamond_{+}I]$ is applied because of the $\Diamond_{+}$ which heads the argument np of the functor 'walks' and it marks the argument, 'John', with the required polarity. As sketched above, rule $[\Box^{\downarrow}_{+}E]$, instead, keeps track of the functional application, marking the lexical entry 'walks', which is the functor, with the structural operator $\langle\cdot\rangle^{+}$.

Based on this marked and structurally analyzed sentence which is the output of our "grammar logic" and making use of the fact that $[\![\text{walks}]\!] \leq [\![\text{moves}]\!]$, the Natural Logic can draw the following inference using the schema given above:

$$\dfrac{\langle\texttt{John}\rangle^{+} \circ \langle\texttt{walks}\rangle^{+} \vdash s}{\langle\texttt{John}\rangle^{+} \circ \langle\texttt{moves}\rangle^{+} \vdash s}$$

However, if we give more complex sentences as input information on the markers distribution are required.

A comparison of the examples (3) "Mary carefully listens to the news" and (5) "Mary doesn't carefully listen to the news", can give the intuition of the important role played by the monotonic function composition: the presence of "doesn't" in (5) changes the polarity of the verb phrase, allowing different substitutions. From monotonicity calculus we know that the composition of monotonic function works simply as an algebraic calculus:

Monotonic Function Composition

(i)	$\uparrow$Mon $\circ$ $\uparrow$Mon =	$\downarrow$Mon $\circ$ $\downarrow$Mon	= $\uparrow$Mon;
(ii)	$\downarrow$Mon $\circ$ $\uparrow$Mon =	$\uparrow$Mon $\circ$ $\downarrow$Mon	= $\downarrow$Mon.

This should be reflected in the proofs given by the formal grammar. To explicate this task we encode the above table of monotonic function composition by means of structural rules which derive $\Gamma[\Delta'] \vdash C$ from $\Gamma[\Delta] \vdash C$, where Δ and Δ' are as specified in the table below:

	Δ	Δ'
$[Mon++]$	$\langle\langle\Sigma\rangle^{+}\rangle^{+}$	$\langle\Sigma\rangle^{+}$
$[Mon--]$	$\langle\langle\Sigma\rangle^{-}\rangle^{-}$	$\langle\Sigma\rangle^{+}$
$[Mon-+]$	$\langle\langle\Sigma\rangle^{-}\rangle^{+}$	$\langle\Sigma\rangle^{-}$
$[Mon+-]$	$\langle\langle\Sigma\rangle^{+}\rangle^{-}$	$\langle\Sigma\rangle^{-}$

which simply says that $++ = -- = +$ and that $+- = -+ = -$.

The monotonicity distribution of a marker over all the elements within its scope is accounted for by the following rule. Let $j \in \{+,-\}$, then

$$\dfrac{\Gamma[\langle\Delta_1 \circ_i \Delta_2\rangle^{j}] \vdash C}{\Gamma[\langle\Delta_1\rangle^{j} \circ_i \langle\Delta_2\rangle^{j}] \vdash C}\,[Monj]$$

In order to see how these postulates together with the logical rules succeed in determining the polarity markers in the process of a proof, we consider the example of the object wide scope reading of the sentence *No boy reads a novel*.

Example 3.2 *No boy reads a novel*

Lexical entries:

$\texttt{no} \vdash \Box^{\downarrow}_{+}((s/\Diamond_{-}(\Diamond_{j} np\backslash s))/\Diamond_{-} n) \qquad \texttt{boy} \vdash n$

$\texttt{a} \vdash \Box^{\downarrow}_{+}((\Diamond_{+}(s/\Diamond_{j} np)\backslash s)/\Diamond_{+} n) \qquad \texttt{novel} \vdash n$

$\texttt{reads} \vdash \Box^{\downarrow}_{+}((\Diamond_{+} np\backslash s)/\Diamond_{+} np)$

The category assigned to 'no' expresses that it is a decreasing monotonic function in both its arguments $\Diamond_{-} n$ and $\Diamond_{-}(\Diamond_{j} np\backslash s)$. This is captured by the fact that both arguments are headed by $\Diamond_{-}$, while the second argument of 'no' can be an increasing or a decreasing function – since its np argument is headed by $\Diamond_{j}$, which ranges over $\Diamond_{+}, \Diamond_{-}$. Applying these entries and applying the logical and the structural rules to the above entries, the proof in figure 1 is derived.

Adopting a top to bottom perspective, we see that the proof is given by hypothetical reasoning, assuming the arguments taken by 'reads' – r_0 and r_2 – and substituting them with 'no boy' and 'a novel', respectively. In particular, the second substitution which proceeds via $[/I]^4$, is made possible by applying the structural rules $[Mon-]$ $[Mon+-]$, which distribute the monotonicity markers in accordance with the monotonic function composition. Thus in the application of $[/I]^4$ the fact that the np is abstracted from a negative context is marked by the $\Diamond_{-}$. After that, the proof goes on as usual. A clear explanation of the use of hypothetical reasoning in MMCG can be found in [Moo97].

Once the formal grammar has syntactically parsed the string given as input and assigned the monotonicity markers to each component, the following inference can be derived:

If $[\![(\text{a} \circ \text{novel})]\!] \leq [\![(\text{a} \circ \text{book})]\!]$, then

$$\frac{((\texttt{no} \circ \texttt{boy}) \circ \texttt{reads}) \circ \langle \texttt{a} \circ \texttt{novel}\rangle^{+} \vdash s}{((\texttt{no} \circ \texttt{boy}) \circ \texttt{reads}) \circ \langle \texttt{a} \circ \texttt{book}\rangle^{+} \vdash s}$$

3.2.1 Structural Rules: A closer look

In Example 3.2 we have used a general form of associativity. However, now that the reader is more familiar with MMCG logic, we can come back to the problem the introduction of a so general property entails.

As we have said the unary operators are meant to lexically control the structural rules. Therefore, in order to have structural rules only *locally* available, we can restrict them to specific modally decorated formula. The lexical entries for such specific formula will be modified in the lexicon, and the postulates refined in the way given below. As an example we derive the relative clause we have said to be underivable in L. We use the right-associativity $[P2]$ and a mixed associativity.

Example 3.3 *that Sara gives to Noradin*

Structural Rules

$$\frac{[(\Delta_1 \circ \langle\Delta_3\rangle) \circ \Delta_2] \vdash C}{[(\Delta_1 \circ \Delta_2) \circ \langle\Delta_3\rangle] \vdash C}\ [P3]$$

$$\frac{[\Delta_1 \circ (\Delta_2 \circ \langle\Delta_3\rangle)] \vdash C}{[(\Delta_1 \circ \Delta_2) \circ \langle\Delta_3\rangle] \vdash C}\ [P2]$$

Lexical entries:

$\texttt{that} \vdash \Box^{\downarrow}_{+}((\Diamond_{+} n\backslash n)/\Diamond_{+}(s/\Diamond\Box^{\downarrow}\Diamond_{j} np))$

$\texttt{gives} \vdash \Box^{\downarrow}(((\Diamond_{+} np\backslash s)/\Diamond_{+} pp)/\Diamond_{+} np)$

$\texttt{to} \vdash \Box^{\downarrow}_{+}(pp/\Diamond_{+} np)$

Thanks to these changes the derivation of 'that Sara gives to Noradin' in Figure 2 can be obtained. As the reader can see, the lexical entry assigned to 'that' requires the introduction of an hypothesis of category $\Box^{\downarrow}\Diamond_{+} np$ which substitutes the $\Diamond_{+} np$. The $\Box^{\downarrow}$ which heads this hypothesis plays a fundamental role on the structural side: By means of the $[\Box^{\downarrow}E]$ 'p_1' is marked by the structural operator $\langle\cdot\rangle$, which allows the application of [P3] and [P2]. Once the structure has been properly modified the hypothesis $\Box^{\downarrow}\Diamond_{+} np$ is eliminated via the $[\Diamond E]^2$ rule and the introduction of the hypothesis $[r_0 \vdash \Diamond\Box^{\downarrow}\Diamond_{+} np]^1$ which will be withdrawn by $[/I]^1$. Then the proof goes on as in figure 1. We abbreviate the category assigned to 'that' with rel.

Figure 1

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\vdots\\ \langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^- \vdash s/\Diamond_-(\Diamond_+ np\backslash s)\end{array} \quad \dfrac{\dfrac{\dfrac{[\mathrm{r}_0 \vdash \Diamond_+ np]^1 \quad \begin{array}{c}\vdots\\ \langle\mathrm{p}_1\rangle^+ \circ (\langle\mathtt{reads}\rangle^+ \circ \langle\mathrm{r}_2\rangle^+) \vdash s\end{array}}{\mathrm{r}_0 \circ (\langle\mathtt{reads}\rangle^+ \circ \langle\mathrm{r}_2\rangle^+) \vdash s}[\Diamond E]}{\langle\mathtt{reads}\rangle^+ \circ \langle\mathrm{r}_2\rangle^+ \vdash \Diamond_+ np\backslash s}[\backslash I]^1}{\langle\langle\mathtt{reads}\rangle^+ \circ \langle\mathrm{r}_2\rangle^+\rangle^- \vdash \Diamond_-(\Diamond_+ np\backslash s)}[\Diamond_- I]}{(\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\langle\mathtt{reads}\rangle^+ \circ \langle\mathrm{r}_2\rangle^+\rangle^- \vdash s}[/E]}{(\langle\mathtt{no}\rangle^+ \circ \langle\langle\mathtt{boy}\rangle^+\rangle^-) \circ (\langle\mathtt{reads}\rangle^- \circ \langle\langle\mathrm{r}_2\rangle^+\rangle^-) \vdash s}[Mon-]}{(\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ (\langle\mathtt{reads}\rangle^- \circ \langle\mathrm{r}_2\rangle^-) \vdash s}[Mon+-]}{((\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^-) \circ \langle\mathrm{r}_2\rangle^- \vdash s}[P1] \quad [\mathrm{p}_2 \vdash \Diamond_- np]^4}{((\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^-) \circ \mathrm{p}_2 \vdash s}[\Diamond_- E]^3}{(\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^- \vdash s/\Diamond_- np}[/I]^4}{\langle(\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^-\rangle^+ \vdash \Diamond_+(s/\Diamond_- np)}[\Diamond_+ I] \quad \langle\mathtt{a}\rangle^+ \circ \langle\mathtt{novel}\rangle^+ \vdash \Diamond_+(s/\Diamond_- np)\backslash s}{\dfrac{\langle(\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^-\rangle^+ \circ (\langle\mathtt{a}\rangle^+ \circ \langle\mathtt{novel}\rangle^+) \vdash s}{((\langle\mathtt{no}\rangle^+ \circ \langle\mathtt{boy}\rangle^-) \circ \langle\mathtt{reads}\rangle^-) \circ (\langle\mathtt{a}\rangle^+ \circ \langle\mathtt{novel}\rangle^+) \vdash s}[Mon+]}[\backslash E]
$$

Figure 2

$$
\dfrac{\begin{array}{c}\vdots\\ \langle\mathrm{that}\rangle^+ \vdash rel\end{array} \quad \dfrac{\dfrac{\dfrac{[\mathrm{r}_0 \vdash \Diamond\Box^{\downarrow}\Diamond_+ np]^1 \quad \dfrac{\dfrac{\dfrac{\dfrac{[\mathrm{p}_1 \vdash \Box^{\downarrow}\Diamond_+ np]^2}{\langle\mathrm{p}_1\rangle \vdash \Diamond_+ np}[\Box^{\downarrow}E] \quad \begin{array}{c}[\mathrm{q} \vdash np]^3\\ \vdots\\ \langle\mathrm{sara}\rangle^+ \circ (((\langle\mathrm{gives}\rangle^+ \circ \langle\mathrm{q}_1\rangle^+) \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+) \vdash s\end{array}}{\langle\mathrm{sara}\rangle^+ \circ (((\langle\mathrm{gives}\rangle^+ \circ \langle\mathrm{p}_1\rangle) \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+) \vdash s}[\Diamond E]^3}{\langle\mathrm{sara}\rangle^+ \circ (((\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+) \circ \langle\mathrm{p}_1\rangle) \vdash s}[P3]}{(\langle\mathrm{sara}\rangle^+ \circ (\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+)) \circ \langle\mathrm{p}_1\rangle \vdash s}[P2]}{(\langle\mathrm{sara}\rangle^+ \circ (\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+)) \circ \mathrm{r}_0 \vdash s}[\Diamond E]^2}{\langle\mathrm{sara}\rangle^+ \circ (\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+) \vdash s/\Diamond\Box^{\downarrow}\Diamond_+ np}[/I]^1}{\langle\langle\mathrm{sara}\rangle^+ \circ (\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+)\rangle^+ \vdash \Diamond_+(s/\Diamond\Box^{\downarrow}\Diamond_+ np)}[\Diamond I]}{\langle\mathrm{that}\rangle^+ \circ \langle\langle\mathrm{sara}\rangle^+ \circ (\langle\mathrm{gives}\rangle^+ \circ \langle\langle\mathrm{to}\rangle^+ \circ \langle\mathrm{noradin}\rangle^+\rangle^+)\rangle^+ \vdash \Diamond_+ n\backslash n}[/E]
$$

Before concluding we want to make the reader aware of some aspects of scope ambiguity phenomena underlined by Natural Logic. As it is well-known sentences for which more than one proof are derivable have more than one meaning, and consequently different inferences could be derived from them.

For example, the derivation of the narrow scope object reading of the sentence considered in Example 3.2, *Nobody reads a novel*, gives rise to a marked output which justifies the following monotonicity inference, which, as expected differs from the one above derived.

If $[\![(\mathrm{a} \circ (\mathrm{nice} \circ \mathrm{novel}))]\!] \leq [\![(\mathrm{a} \circ \mathrm{novel})]\!]$, then

$$\frac{(\texttt{no} \circ \texttt{boy}) \circ (\texttt{reads} \circ \langle \texttt{a} \circ \texttt{novel} \rangle^{-}) \vdash s}{(\texttt{no} \circ \texttt{boy}) \circ (\texttt{reads} \circ \langle \texttt{a} \circ (\texttt{nice} \circ \texttt{novel}) \rangle^{-}) \vdash s}$$

However, the Natural Logic described so far derives also inferences which are *logically* correct but *intuitively* wrong. Let's take a look at an example.

Example 3.4 *Mary reads no book*
If $[\![(\mathrm{every} \circ \mathrm{woman})]\!] \leq [\![\mathrm{mary}]\!] \leq [\![(\mathrm{a} \circ \mathrm{woman})]\!]$, then

– the wide-scope object reading entails:

$$\frac{(\langle \texttt{mary} \rangle^{-} \circ \texttt{reads}) \circ (\texttt{no} \circ \texttt{book}) \vdash s}{(\langle \texttt{every} \circ \texttt{woman} \rangle^{-} \circ \texttt{reads}) \circ (\texttt{no} \circ \texttt{book}) \vdash s}$$

– and the narrow-scope object reading entails:

$$\frac{\langle \texttt{mary} \rangle^{+} \circ (\texttt{reads} \circ (\texttt{no} \circ \texttt{book})) \vdash s}{\langle \texttt{a} \circ \texttt{woman} \rangle^{+} \circ (\texttt{reads} \circ (\texttt{no} \circ \texttt{book})) \vdash s}$$

These two inferences are both logically correct, however, only the second could intuitively be accepted[1].

Sanchez [SV91] discussing these phenomena derives an interesting philosophical implication. He observes that this could imply that proper names are scope-less *semantically*, but they are not so *inferentially*. Our intent is to investigate these phenomena looking for a general explanation.

These examples make us think that the distribution of the monotonicity markers involves semantic explanations, as well as syntactic ones. Our aim is to account for the semantic properties involved in the monotonicity reasoning simply via the logic of MMCG. Due to the presence of the structural operators and the structural postulates MMCG has so far shown to be powerful enough to achieve this goal.

[1] If, for example, we translate the sentence *Mary reads no book* into Predicate Logic we clearly see that the inference is logically correct: if $\forall x(B(x) \rightarrow \neg R(m, x))$ and $W(m)$ holds, then $\forall x(B(x) \rightarrow \neg \forall y(W(y) \rightarrow R(y, x)))$ also holds, although nobody will consider the corresponding natural language inference valid.

4 Further research

The perspective we have assumed and the choice of working with MMCG presents several advantages with respect to (more traditional) approaches and opens interesting problems for further investigations.

In particular, the possibility of controlling the structural variation among languages could shed new light on the investigation of natural reasoning from a cross-linguistic perspective, analysing the way in which differences in the structure of the sentence might influence the derivation.

References

[AC96] M. Abrusci and C. Casadio, editors. *Proofs and Linguistic Categories*, Proceedings Roma Workshop. CLUEB, 1996.

[Ajd35] Ajdukiewicz. Die Syntaktische Konnexität. *Studia Philosophica*, (1:1-27), 1935.

[Ben87] J. van Benthem. Meaning: Interpretation and Inference. *Synthese*, (73):451–470, 1987.

[Ber99] R. Bernardi. Monotonic Reasoning from a Proof-Theoretic Perspective. In G. Kruijff and R. Oehrle, editors, *Proceedings of Formal Grammar 1999*, pages 13–24, 1999.

[BH64] Y. Bar-Hillel. *Language and Information*. Addison-Wesley, New York, 1964.

[Dow94] D. Dowty. The Role of Negative Polarity and Concord Marking in Natural Language Reasoning. In *Semantics and Linguistic Theory*, volume IV, pages 114–144. Cornell University, 1994.

[Lam58] J. Lambek. The Mathematics of Sentence Structure. *American Mathematical Monthly*, 65:154–170, 1958. trad. 'La matematica della struttura di frase'. In *Linguaggio e Sistemi Formali*, De Palma A. (a cura di), Torino, Einaudi, pp. 165-186.

[Moo97] M. Moortgat. Categorial Type Logics. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 93–178. The MIT Press, Cambridge, Massachusetts, Cambridge, 1997.

[Moo98] R. Moot. Grail: An Automated Proof Assistant for Categorial Grammar Logics. In R. Backhouse, editor, *Proceedings of the 1998 User Interfaces for Theorem Provers Conference*, pages 120–129, 1998.

[OBW88] R. Oehrle, E. Bach, and D. Wheeler, editors. *Categorial Grammars, and Natural Langauge Structures*. Reidel, Dordrecht, 1988.

[SV91] V. Sanchez Valencia. *Studies on Natural Logic and Categorial Grammar*. PhD thesis, University of Amsterdam, 1991.

[Zwa86] F. Zwarts. *Categoriale Grammatica en Algebraische Semantiek*. PhD thesis, University of Groningen, 1986.