

Adjacency, dependency and order

Michael Moortgat (OTS) and Richard Oehrle (University of Arizona)

To appear in Dekker & Stokhof (eds.)
Proceedings 9th Amsterdam Colloquium

Abstract

This paper is a contribution to the study of linguistic inference in the context of a logic of structured resources. We investigate a structural parameter of *adjacency* and its interaction with the parameters of linear precedence, dominance and dependency. The theoretical framework is a *multimodal* logic of categorial type inference — a logic where families of residuated multiplicatives live together. The communication between multiplicatives is implemented semantically via frame conditions linking the distinct accessibility relations interpreting the multiplicative connectives. Proof-theoretically, the frame conditions find expression in terms of Inclusion Axioms and Interaction Axioms. The multimodal logic of is applied to linguistic phenomena of verb-raising, adverb placement, and extraposition — non-adjacent modes of combination which are analysed here in terms of head attraction and head adjunction.

1 Multimodal grammatical inference

Classical systems of grammatical deduction — familiar as the non-associative Lambek calculus NL, its associative counterpart L, and the associative, commutative Lambek-Van Benthem system LP — are based on a binary structure-building product operator (represented by the sequent comma in Gentzen style presentations), with a matching product connective ‘o’ in the type language together with its left and right residuals ‘/’ and ‘\’. These classical systems are differentiated by the presence or absence of structural rules characterizing the product operator as associative and/or commutative.

From the classical perspective, structural rules are a global option, either always available or never. But natural languages display a characteristic control of structural resources beyond the reach of this global strategy. Systems of *multimodal* grammatical inference provide a setting in which some of the subtleties of linguistic composition can be addressed from a logical perspective.

Multimodal systems contain a multiplicity of co-existing structure-building product operators — a possibility anticipated in the logical literature by Dunn [10], by Belnap's 'Display Logic' [3], and recently exploited in Girard's [11] 'Logique Universelle'. Corresponding to each structural product connective is a logical product connective (type-constructor) with its left and right residuals. In addition to this uniform approach to the individual products and their associated implications, there may be non-trivial properties relating the properties of two or more products. In Kripke frame semantics for these systems, such relations take the form of linking principles (or correlation postulates) involving the ternary accessibility relations interpreting the different products in question. In postulate presentations, the correlation principles are realized as axiom schemata relating the type-constructors corresponding to the underlying product operations. In sequent presentations, these axiom schemata are recast as structural rules so as to allow cut-free proof search.

In this paper we consider how the multimodal perspective applies to the structural parameter of *adjacency*, and to the interaction of this parameter with the parameters of linear precedence, dominance, and dependency. Our aim is to address some linguistic questions first treated in the categorial literature in the pioneering work of Emmon Bach [1, 2] — questions which in the GPSG/HPSG tradition are dealt with by means of ID/LP statements, in the GB tradition by extraposition and head adjunction rules. Our focus is on structures involving semantic composition between configurations A and B , with A obligatorily preceding B , where in some cases we find that A must immediately precede B — they form an adjacent configuration — while in other cases precedence is required, but adjacency is not. After defining a multimodal framework appropriate to the investigation of these properties, we apply the system to a variety of linguistic problems beyond the reach of classical systems.

1.1 Substructural communication: model theory, proof theory

The category formulae for multimodal type logic are defined inductively on the basis of a set of category atoms \mathcal{A} and a set of indices I as shown below. We'll refer to the $i \in I$ as *resource management modes*, or modes for short.

$$\mathcal{F} ::= \mathcal{A} \mid \mathcal{F} /_I \mathcal{F} \mid \mathcal{F} \circ_I \mathcal{F} \mid \mathcal{F} \setminus_I \mathcal{F}$$

The semantics for this type language is a straightforward generalisation of *ternary frame semantics* for the standard systems (cf. [8] where completeness and decidability results are discussed). The ingredients of the multimodal approach can be summarized as follows.

(i) Rather than interpret multiplicative type-constructors in terms of *one* privileged notion of linguistic composition, we throw different forms of linguistic com-

position together and interpret in multimodal ternary frames $\langle W, (R_i)_{i \in I} \rangle$, where the ‘worlds’ W are to be seen as the linguistic resources, and where the accessibility relations R_i model the relevant notions of linguistic composition. (Think of R_ixyz as ‘putting together resources x and y in mode i yields z ’.) A *valuation* on a ternary frame $\langle W, (R_i)_{i \in I} \rangle$ consists of a map $\mathcal{F} \rightarrow \mathcal{P}(W)$ which respects the structure of the complex types in the following way, for each of the modes $i \in I$:

$$\begin{aligned} v(A \circ_i B) &= \{z \mid \exists x \exists y [R_i xyz \ \& \ x \in v(A) \ \& \ y \in v(B)]\} \\ v(C /_i B) &= \{x \mid \forall y \forall z [(R_i xyz \ \& \ y \in v(B)) \Rightarrow z \in v(C)]\} \\ v(A \setminus_i C) &= \{y \mid \forall x \forall z [(R_i xyz \ \& \ x \in v(A)) \Rightarrow z \in v(C)]\} \end{aligned}$$

(ii) The resource management properties of the individual R_i are obtained by imposing frame conditions (‘constraints’) which restrict the possible interpretations of the R_i . For example, let \circ_c be a commutative product, then in the semantics the R_c interpreting this connective will be constrained to satisfy $(\forall x, y, z \in W) R_c xyz \Rightarrow R_c yxz$.

(iii) Communication between different modes of composition \circ_i and \circ_j is obtained by further frame conditions linking R_i and R_j . Following the logical tradition, we call these *correlation postulates*. Two forms of correlation will be prominent in this paper:

– ordering of the accessibility relations R_i in terms of the information they provide about the structure of the linguistic resources. In terms of frame conditions (for R_i more informative than R_j) this means $(\forall x, y, z \in W) R_i xyz \Rightarrow R_j xyz$. (For example, since the non-commutative product \bullet is more informative than the commutative product \otimes , we’d have the constraint: $(\forall x, y, z) R_\bullet xyz \Rightarrow R_\otimes xyz$.) Corresponding to this form of correlation, there will be Inclusion Axioms in the logic.

– correlation postulates ‘mixing’ distinct R_i (for example: mixed associativity or commutativity laws). This form of correlation is expressed in the logic in terms of Interaction Axioms. They form our main concern in §3 and §4.

We can present the type logic in a variety of ways. In an axiomatic presentation, each product \circ_i is systematically connected to corresponding implication operators by residuation laws of the (schematic) form:

$$A \longrightarrow C /_i B \quad \text{iff} \quad A \circ_i B \longrightarrow C \quad \text{iff} \quad B \longrightarrow A \setminus_i C$$

In sequent presentation, we follow Belnap [3] in making a distinction between *logical* connectives (i.e. the type constructors, taking types into types) and *structural* connectives (taking structures into structures). Each residuated family of multiplicatives $\{/_i, \circ_i, \setminus_i\}$ has a matching structural connective, or sequent punctuation $(\cdot, \cdot)^i$. Antecedent terms are inductively defined as $\mathcal{T} ::= \mathcal{F} \mid (\mathcal{T}, \mathcal{T})^I$. A sequent is a pair $(\mathcal{T}, \mathcal{F})$, which we write as $\mathcal{T} \Rightarrow \mathcal{F}$. Axiom, Cut and the logical

rules are uniform. Logical rules insist that use and proof of connectives respect the resource management modes.

$$\begin{array}{c}
[Ax] \frac{}{A \Rightarrow A} \quad \frac{\Gamma \Rightarrow A \quad \Delta[A] \Rightarrow C}{\Delta[\Gamma] \Rightarrow C} [Cut] \\
[R/i] \frac{(\Gamma, B)^i \Rightarrow A}{\Gamma \Rightarrow A/i, B} \quad \frac{\Gamma \Rightarrow B \quad \Delta[A] \Rightarrow C}{\Delta[(A/i, B, \Gamma)^i] \Rightarrow C} [L/i] \\
[R\backslash i] \frac{(B, \Gamma)^i \Rightarrow A}{\Gamma \Rightarrow B\backslash i, A} \quad \frac{\Gamma \Rightarrow B \quad \Delta[A] \Rightarrow C}{\Delta[(\Gamma, B\backslash i, A)^i] \Rightarrow C} [L\backslash i] \\
[Lo_i] \frac{\Gamma[(A, B)^i] \Rightarrow C}{\Gamma[A \circ_i B] \Rightarrow C} \quad \frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{(\Gamma, \Delta)^i \Rightarrow A \circ_i B} [Ro_i]
\end{array}$$

In addition to the residuation inferences which are shared by all resource management modes, there are mode-specific structural options. In axiomatic style, they take the form of *structural axioms*; in sequent presentation, there is corresponding *structural rules*. As an illustration, see the structural axioms/rules for the commutative mode c for which we gave the corresponding frame condition above:

$$A \circ_c B \longleftrightarrow B \circ_c A \quad \frac{\Gamma[(\Delta_2, \Delta_1)^c] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^c] \Rightarrow A} [Comm]$$

Multimodal communication follows the same general agenda model-theoretically and proof-theoretically: for a correlation postulate linking R_i, R_j in the semantics, we find, in the deductive presentation, an axiom schema capturing the restriction on the models, and in the Gentzen presentation a corresponding structural rule. As an illustration, suppose again we have products \circ_n, \circ_a (non-associative vs associative composition) interpreted by the accessibility relations R_n, R_a , with R_n more informative than R_a , i.e. $R_n \subseteq R_a$. The correlation is expressed by the Inclusion Axiom (deductive presentation) or Rule (sequent style) below.

$$\frac{A \circ_n B \longrightarrow A \circ_a B}{\text{logically}} \quad \frac{\Gamma[(\Delta, \Delta')^a] \Rightarrow A}{\Gamma[(\Delta, \Delta')^n] \Rightarrow A} \quad \text{Double arrow}$$

The classical systems LN, LNP, L, and LP already provide examples of distinct, but compatible, families of structural products and their corresponding type-constructors. In the multimodal setting, we can put these different systems together, and reason about mixed configurations. The linking principles regulating the communication take the form of *loss* of information, moving from the non-associative to the associative mode, or from the non-commutative to the commutative mode. For what is to follow, it is important to realize that the classical options do not exhaust the range of possible products in the multimodal setting. We rely below on ‘dependency types’, introduced and developed by Moortgat

& Morrill [18], which are based on the left-headed structural product $(\cdot, \cdot)^l$ and the right-headed product $(\cdot, \cdot)^r$. The dependency dimension introduces an asymmetry (the opposition between heads and dependents) which is orthogonal to the function/argument asymmetry. The product and residuation properties of the corresponding asymmetrical type-constructors—the families $\{/r, \circ_r, \backslash_r\}$ and $\{/l, \circ_l, \backslash_l\}$ —follow the general principles laid out above. One obvious linking principle ‘forgets’ the head/non-head distinction.¹

We shall shortly turn to a more detailed examination of another structural parameter: adjacency. Before taking up this question in detail, however, we extend the basic type system in another direction—sensitivity to prosodic sorting.

1.2 Refinement: sortal decoration

So far we have been discussing types with respect to their resource management properties. The resource management component is just one dimension of the linguistic resources. In general, we want to be able to associate types with information of different kinds — for example, information relating to morphophonological or prosodic properties of the linguistic resources. In order to make the type language expressive enough to take these extra dimensions into account we move to a *labelled* type system. In the extreme case, one might imagine incorporating a complete description of the prosodic or morphophonological structure of an expression or its arguments. This suggests the possibility that type-driven composition could be sensitive to arbitrarily fine prosodic or morphophonological detail. It is worthwhile to investigate the expressive power that can be achieved on the basis of a simple system of labels.

Labelled types are obtained from the old unlabelled types by decorating every subtype with a label, which we will write as a subscript. The labelled type formulae are then defined on the basis of our old set of type atoms A and a set of labels S as follows.

$$\mathcal{FL} ::= A_S \mid (\mathcal{FL} /_I \mathcal{FL})_S \mid (\mathcal{FL} \circ_I \mathcal{FL})_S \mid (\mathcal{FL} \backslash_I \mathcal{FL})_S$$

The labeling system used in the present paper will be extremely simple. We think of the labels as providing information about the *prosodic sort* of the linguistic resources, so that we can make a distinction between, say, resources of ‘lexical’ sort and resources of ‘phrasal’ sort. The prosodic sort labeling makes it possible to express morphophonological constraints on the operands and the result of the

¹Interactions with associativity and commutativity parameters are more complex and pose non-trivial choices. To illustrate, a commutative product can be regarded as ‘headed’ in two distinct ways: the ‘head’ might be associated with a specific element of the pair; or it might be associated with a specific property of the pair, e.g. the rightmost of the two. These two modes of linking offer new type-theoretic models of empirical linguistic phenomena. A more complete exploration of the space of structural products can be found in Moortgat & Oehrle [19].

multiplicative operations, so that one can distinguish, for example, composition at the lexical level ('word formation') from composition at the phrasal level ('syntax proper'). In line with the work on typed feature structures within the framework of HPSG, we assume the prosodic sorts to be structured as an inheritance hierarchy $\langle S, \sqsubseteq \rangle$. We follow the practice of [6] and write $a \sqsubseteq b$ for $a, b \in S$ if b inherits from a , i.e. a is more general than (subsumes) b , or b is more specific than a .

The move to a labeled type system requires some adjustments in the model theory and proof rules. In the semantics, we replace our monodimensional models where formulae are evaluated at a resource point $w \in W$, by two-dimensional models with evaluation at pairs (w, s) where s is a prosodic sort.² The semantics is set up in such a way that in a deductive presentation of the labeled type logic, the logical axiom schema and the residuation inferences take the following form.³

$$A_a \longrightarrow A_{a'} \quad \text{if} \quad a' \sqsubseteq a$$

$$A_a \longrightarrow (C_c/B_b)_a \quad \text{iff} \quad (A_a \circ B_b)_c \longrightarrow C_c \quad \text{iff} \quad B_b \longrightarrow (A_a \setminus C_c)_b$$

Note that on this account, the relations among different sorts depends on properties of the model: the type system itself, for example, does not determine that the sort of the product of two sorted types of sorts a_1 and a_2 subsume both a_1 and a_2 .

Likewise, in the sequent presentation we move to labelled sequents (Δ, A) , also written as $\Delta \Rightarrow A$, with $\Delta \in \mathcal{T}$ and $A \in \mathcal{FL}$, where labelled antecedent terms are defined as usual: $\mathcal{T} ::= \mathcal{FL} \mid (\mathcal{T}, \mathcal{T})_s^I$. In accordance with the sort inheritance hierarchy, we allow replacement of antecedent subterms by less informative subterms, in the following sense:

$$\frac{\Gamma[(\Delta_1, \Delta_2)_{a'}] \Rightarrow B_b}{\Gamma[(\Delta_1, \Delta_2)_a] \Rightarrow B_b} \quad a' \sqsubseteq a$$

The labelled version of the logical rules is straightforward. We give the product rules, and leave the residual implications to the reader.

$$\frac{\Gamma[(A_a, B_b)_c] \Rightarrow C_d}{\Gamma[(A_a \circ B_b)_c] \Rightarrow C_d} \text{L}(\cdot \circ \cdot)_c \quad \frac{\Gamma \Rightarrow A_a \quad \Delta \Rightarrow B_b}{(\Gamma, \Delta)_c \Rightarrow (A_a \circ B_b)_c} \text{R}(\cdot \circ \cdot)_c$$

We close this section with a very simple illustration of how one can keep word-formation and syntax apart by making the type assignment sort-sensitive. Compare the adjectival modifier *very* and the prefix *un-*. In the dimension of syntax/semantics, both qualify as functors taking adjectives into adjectives, i.e. we

²We cannot fully work out the details of such two-dimensional interpretation in the present paper. The interested reader is referred to [15].

³We omit the resource management mode indices, so that the reader can concentrate on the flow of information wrt the sort specifications.

can assign them the type a/a . Such an assignment is not fine-grained enough to capture the distinction between the grammatical phrase *very unhappy* and the ungrammatical *un-very happy* where phrasal and lexical composition do not properly interact. To take into account the morphophonological dimension, we move to the sorted typing shown below. The sort labeling characterizes *very* as a syntactic word, taking phrasal expressions into phrasal expressions, whereas *un-* is typed as an affix, taking lexical expressions into lexical expressions.

$$\text{un} := (a_w/a_w)\text{affix} \quad \text{very} := (a_{ph}/a_{ph})w$$

We also assume $ph \sqsubseteq w$ in the sort order, expressing the fact that a syntactic word can do duty as a phrasal expression, but not vice versa. Given the sort-decorated types, the expression *very unhappy* is derivable. The attempt to derive the ungrammatical *un-very happy* fails because of a violation of the sort constraints. As soon as *very* has combined with *happy*, the result is of phrasal sort, which is in conflict with the lexical sort required by the affix *un*. Below is a (natural deduction style) derivation of the well-formed case.

$$\frac{\frac{\text{very}}{(a_{ph}/a_{ph})w} \quad \frac{\frac{\text{un}}{(a_w/a_w)\text{affix}} \quad \frac{\text{happy}}{a_w}}{a_w} \text{E}}{a_{ph}} \text{E}}{a_{ph}} \text{E}$$

2 The adjacency parameter

The connectives discussed in the previous section have a common characteristic: linguistic composition is defined only for adjacent configurations. In Combinatory Categorical Grammar [23, 25], adjacency is taken to be a defining characteristic of categorial composition — the so-called ‘Adjacency Principle’. From the multimodal perspective, however, it is more natural to regard adjacency as a *parameter* of resource structuring. We adopt this perspective here and investigate variants of the multiplicatives where adjacency is not a necessary condition for composition. Schematically, we will then face one of the following situations:

$$(\star) \quad (\Gamma[\Delta_1], \Delta_2) \rightsquigarrow \Gamma[(\Delta_1, \Delta_2)] \quad (\star\star) \quad \Gamma[(\Delta_1, \Delta_2)] \rightsquigarrow (\Gamma[\Delta_1], \Delta_2)$$

On the lefthand side of ‘ \rightsquigarrow ’ we represent the actual structures we encounter, on the righthand side the structural alternative that is required for the logical rules to be applicable. In the case of (\star) the configuration Δ_2 is put together with Γ whereas in fact it should enter into composition with a subconfiguration Δ_1 of Γ . The $(\star\star)$ case is symmetric: Γ here contains a subconfiguration (Δ_1, Δ_2) whereas Δ_2 in fact requires linguistic composition with $\Gamma[\Delta_1]$. In the remainder

of the paper we investigate a number of situations of non-adjacent composition which are sensitive to the *dependency* structure of the linguistic resources. In that setting, (\star) could represent the case of a dependent Δ_2 *attracted* by the head Δ_1 , whereas ($\star\star$) would be the case of a configuration Δ_2 *adjoining* itself to the head Δ_1 of the configuration over which it has scope.

The core of our analysis of non-adjacent composition will be a set of interaction principles regulating the communication between the adjacent and the non-adjacent modes. In an axiomatic presentation of the logic, the communication will be captured by Interaction Axioms of Mixed Associativity and Mixed Commutativity.

$$\begin{aligned} Ax1 : & (A \circ_i B) \circ_j C \longleftrightarrow (A \circ_j C) \circ_i B & : Ax1' \\ Ax2 : & (A \circ_i B) \circ_j C \longleftrightarrow A \circ_i (B \circ_j C) & : Ax2' \end{aligned}$$

The j -mode here generically represents the non-adjacent mode of composition, the i -mode adjacent dependency composition (e.g. left or right headed structuring). Specific cases will be obtained by instantiating the generic i and j modes by concrete type constructors.

The axiom schemata $Ax1/Ax2$, $Ax1'/Ax2'$ correspond to frame conditions restricting the allowable models, cf. $F1/F2$, $F1'/F2'$ below ($\forall xyzuv \in W$):

$$\begin{aligned} F1 : & (R_i xyz \ \& \ R_j zuv) \Rightarrow \exists t (R_j xut \ \& \ R_i tyv) \\ F2 : & (R_i xyz \ \& \ R_j zuv) \Rightarrow \exists t (R_j yut \ \& \ R_i xtv) \\ F1' : & (R_j xyz \ \& \ R_i zuv) \Rightarrow \exists t (R_i xut \ \& \ R_j tyv) \\ F2' : & (R_j xyz \ \& \ R_i zuv) \Rightarrow \exists t (R_i xut \ \& \ R_j tyv) \end{aligned}$$

In sequent presentation, we recast the axiom schemata in the usual structural rule format. The structural rules below represent the situation where a dependent configuration Δ_3 is attracted from the left by a head.

$$\frac{\Gamma[((\Delta_1, \Delta_3)^j, \Delta_2)^i] \Rightarrow A}{\Gamma[((\Delta_1, \Delta_2)^i, \Delta_3)^j] \Rightarrow A} R1 \quad \frac{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^j)^i] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^i, \Delta_3)^j] \Rightarrow A} R2$$

Depending on whether the target head is a subconfiguration of Δ_1 or Δ_2 , Mixed Commutativity (R1) or Mixed Associativity (R2) will be applicable. Symmetric cases for attraction from the right can be added. The situation for head adjunction is dual: structural rules are obtained by switching premise and conclusion around.

$$\frac{\Gamma[(\Delta_1, \Delta_2)^i, \Delta_3)^j] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_3)^j, \Delta_2)^i] \Rightarrow A} R1' \quad \frac{\Gamma[(\Delta_1, \Delta_2)^i, \Delta_3)^j] \Rightarrow A}{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^j)^i] \Rightarrow A} R2'$$

Linguistic applications. The linguistic applications we will consider in §3 and §4 below are (i) verb raising in Dutch and German and its interaction with adverb placement, and (ii) some cases of extraposition in English.

Under (i) we consider cases such as

- (dat je) soms je boeken thuis kan laten
that one sometimes one's books at-home may leave
- (dat je) je boeken soms thuis kan laten
- *(dat je) je boeken thuis soms kan laten

We treat both adverb placement and verb-cluster formation as cases of head adjunction. Following the lead of Dowty [9], we will distinguish distinct degrees of coherence of structural units within the verb phrase, using the sortal machinery developed in the previous section: the combination *thuis laten* is penetrable for the verb raising trigger *kan*, but it is not for the adverb *soms*. The head adjunction approach naturally accounts for the scoping relations between adverbials and verb-raising triggers, improving on earlier proposals to treat verb raising in terms of 'lexical rules'.

Under (ii) we offer an account for cases such as

- a book appeared about linear logic
- an easy book to read

The first of these cases, we regard as head attraction, by the head *book*. We exploit somewhat different characteristics of the multimodal perspective to account for the discontinuity of the adjective *easy* and its complement phrase, motivated not just by this particular case, but by the typology of English noun modification.

3 Head wrapping

We start our analysis of non-adjacent forms of linguistic composition with an account of head-wrapping — the ($\star\star$) case. We first draw the map of the 'logical space' for head-wrapping operations in general terms. Then we illustrate head-wrapping with verb-raising in Dutch and German, showing how the general framework naturally accomodates the parametric variation between these languages.

The account of head wrapping to be presented here explores the options for structuring the linguistic resources we already saw in the dependency system: the head wrapping operations indeed turn out to differ from the dependency family only in the structural parameter of *adjacency*. Just as we distinguished a left and a right headed form of concatenation ($A \circ_l B$ versus $A \circ_r B$), with four matching implications, we will now distinguish left and right dominant forms of head wrapping. Whereas the dependency products consist of the concatenation of an *A* resource (the prefix) and a *B* resource (the suffix), the head-wrapping products are combinations of an *infix* and a *circumfix* (or host): the infix syntactically adjoins itself to the head of the host circumfix. The structural parameters of precedence

and dependency offer four ways of realising the head adjunction: the infix can go before or after the head of the host, and it can adjoin either as a head (in which case the infix determines the head of the resulting combination) or as a dependent (in which case the host determines the head of the resulting combination). Together these options give rise to *four* head wrapping products, each determining two residual implications. We will use mnemonic labels \circ_{lh} , \circ_{ld} , \circ_{rh} , \circ_{rd} , with l (r) indicating that the left (right) subtype is the infix, and h (d) that the infix (circumfix) determines the head of the combination. See the schema below (with A the infix subtype, B the circumfix):

$A \circ_{lh} B$	$A \circ_{ld} B$
$B \circ_{rd} A$	$B \circ_{rh} A$

In the communication between the dependency and the head-wrapping system, we can distinguish base cases and recursive cases. The **BASE CASES** tell us under what conditions head-wrapping reduces to simple dependency adjunction. We assume that the dependency adjunctions represent ‘concrete’ configurations (structure which is phonetically interpretable), whereas the wrapping configurations represent ‘abstract’ steps in the process of grammatical inference. On the level of resource management, the base cases are handled in terms of inclusion axioms/frame conditions relating the family $\{\circ_{lh}, \circ_{ld}, \circ_{rh}, \circ_{rd}\}$ to $\{\circ_l, \circ_r\}$. On the prosodic level, the applicability of the base cases is dependent on sortal information, with different sortal constraints characterizing distinct types of head adjunction. The **RECURSIVE CASES** establish the communication between head-wrapping and dependency operators in terms of mixed Associativity and Commutativity laws allowing the head-wrapping infix to ‘travel’ along the head-projection until it reaches the landing site characterized by the base cases.

Head wrapping: postulates, structural rules. Let us see what has to be added to the residuation inferences of our multimodal type logic to handle head wrapping. We first concentrate on resource management. In our discussion of specific linguistic phenomena we introduce the sortal finestructure to appropriately characterize the base cases of the recursion.

In a deductive presentation of the type logic, head wrapping requires the Inclusion Axioms $A1/A1'$ and $A2/A2'$, and Interaction Axioms $A3$, $A4$, $A5$, $A6$. The Inclusion Axioms will be constrained to specific prosodic sort information. The Interaction Axioms preserve whatever sort information there is on the subtypes involved. In the Interaction Axioms, lw (rw) stands for either lh or ld (rh or rd).

$$\begin{array}{ll}
 A1 : & A \circ_l B \longrightarrow A \circ_{lh} B \\
 A2' : & B \circ_l A \longrightarrow B \circ_{rd} A \\
 A1' : & A \circ_r B \longrightarrow A \circ_{ld} B \\
 A2 : & B \circ_r A \longrightarrow B \circ_{rh} A
 \end{array}$$

$$\begin{aligned}
A3: & (A \circ_{rw} B) \circ_l C \longrightarrow (A \circ_l C) \circ_{rw} B \\
A4: & A \circ_r (B \circ_{rw} C) \longrightarrow (A \circ_r B) \circ_{rw} C \\
A5: & A \circ_r (B \circ_{lw} C) \longrightarrow B \circ_{lw} (A \circ_r C) \\
A6: & (A \circ_{lw} B) \circ_l C \longrightarrow A \circ_{lw} (B \circ_l C)
\end{aligned}$$

Recasting of the axiom schemata above in structural rules for sequent presentation follows the pattern that will be familiar by now.

$$\frac{\Gamma[(\Delta, \Delta')^{lh}] \Rightarrow C}{\Gamma[(\Delta, \Delta')^l] \Rightarrow C} R1 \quad \frac{\Gamma[(\Delta, \Delta')^{ld}] \Rightarrow C}{\Gamma[(\Delta, \Delta')^r] \Rightarrow C} R1'$$

$$\frac{\Gamma[(\Delta, \Delta')^{rd}] \Rightarrow C}{\Gamma[(\Delta, \Delta')^l] \Rightarrow C} R2' \quad \frac{\Gamma[(\Delta, \Delta')^{rh}] \Rightarrow C}{\Gamma[(\Delta, \Delta')^r] \Rightarrow C} R2$$

$$\frac{\Gamma[(\Delta_2, (\Delta_1, \Delta_3)^r)^{lw}] \Rightarrow A}{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^{lw})^r] \Rightarrow A} \text{Comm R/LW} \quad \frac{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^l)^{lw}] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^{lw}, \Delta_3]^l] \Rightarrow A} \text{Ass L/LW}$$

$$\frac{\Gamma[((\Delta_1, \Delta_3)^l, \Delta_2)^{rw}] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^{rw}, \Delta_3]^l] \Rightarrow A} \text{Comm L/RW} \quad \frac{\Gamma[(\Delta_1, \Delta_2)^r, \Delta_3)^{rw}] \Rightarrow A}{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^{rw})^r] \Rightarrow A} \text{Ass R/RW}$$

In Figure 1 we give a graphical representation of the Interaction Axioms/Rules.

Illustration: verb-raising, Dutch versus German

We can turn now to a concrete illustration, and compare the core cases of Dutch and German verb raising clusters⁴:

(omdat hij) boeken wil kunnen lezen
 (because he) books wants may read
(weil er) Bücher lesen dürfen will
 (because he) books read may wants

On the descriptive level, the difference between the two languages can be characterized as follows. Main verbs are typed as left-oriented head functors, in German and Dutch. In both German and Dutch, the verb raising triggers are head functors (i.e. the head-adjunction infix becomes the head of the configuration it adjoins into); in Dutch the triggers are prefix functors, leading to the crossed dependencies between verbs and their arguments; in German, the triggers are suffixes to the host head.

As remarked above, the Inclusion Axioms which mark the end-point of the recursive descent of a verb-raising trigger into the phrase over which it takes scope

⁴What follows is intended as an illustration of how the multimodal communication principles work, not as an in-depth analysis of the many complexities of verb-raising phenomena!

Comm R/LW		Ass L/LW	
Conclusion	Premise	Conclusion	Premise
$ \begin{array}{c} o_r \\ / \quad \backslash \\ \Delta_1 \quad o_{lw} \\ \quad / \quad \backslash \\ \Delta_2 \quad \Delta_3 \end{array} $	$ \begin{array}{c} o_{lw} \\ / \quad \backslash \\ \Delta_2 \quad o_r \\ \quad / \quad \backslash \\ \Delta_1 \quad \Delta_3 \end{array} $	$ \begin{array}{c} o_l \\ / \quad \backslash \\ o_{lw} \quad \Delta_3 \\ / \quad \backslash \\ \Delta_1 \quad \Delta_2 \end{array} $	$ \begin{array}{c} o_{lw} \\ / \quad \backslash \\ \Delta_1 \quad o_l \\ \quad / \quad \backslash \\ \Delta_2 \quad \Delta_3 \end{array} $
$ \begin{array}{c} o_r \\ / \quad \backslash \\ \Delta_1 \quad o_{rw} \\ \quad / \quad \backslash \\ \Delta_2 \quad \Delta_3 \end{array} $	$ \begin{array}{c} o_{rw} \\ / \quad \backslash \\ o_r \quad \Delta_3 \\ / \quad \backslash \\ \Delta_1 \quad \Delta_2 \end{array} $	$ \begin{array}{c} o_l \\ / \quad \backslash \\ o_{rw} \quad \Delta_3 \\ / \quad \backslash \\ \Delta_1 \quad \Delta_2 \end{array} $	$ \begin{array}{c} o_{rw} \\ / \quad \backslash \\ o_l \quad \Delta_2 \\ / \quad \backslash \\ \Delta_1 \quad \Delta_3 \end{array} $
Conclusion Ass R/RW		Conclusion Comm L/RW	

Figure 1: Head-wrapping: Interaction Axioms/Rules

semantically are sensitive to the prosodic status of the elements combined. Consider the following sort-decorated refinement of the Inclusion Axioms we saw above.

$$\begin{aligned}
 A1: & (A_i o_l B_v)_v \longrightarrow (A_i o_{lh} B_v)_{ph} \\
 A2': & (A_v o_r B_i)_v \longrightarrow (A_v o_{rh} B_i)_{ph}
 \end{aligned}$$

In the prosodic sort hierarchy, we take sort v (for ‘verb cluster’) to be a supralexical but subphrasal sort, i.e. $ph \sqsubseteq v \sqsubseteq w$. This place in the hierarchy makes it possible for the verb cluster to contain multiple lexical verbs, as will be the case when verb-raising is applied recursively. On the other hand, full phrasal complements are excluded (at least in standard Dutch). With sort i (for ‘infix’) we label the verb-raising triggers. We can now paraphrase the sort-decorated version of Inclusion Axioms A1, A2’ as follows: a verb cluster is a headed structure recur-

sively built from a head-wrapping infix and a verb cluster (ultimately, the lexical head); if we find a configuration meeting these sortal characteristics, it qualifies as a phrase-level head-wrapping configuration which has access to non-adjacent modes of composition through the Interaction Axioms A3-A6.

The lexical type assignments below implement what has been said so far.

boeken := np_{ph}	Bücher := np_{ph}
wil := $(vp_{ph}/lh\ inf_{ph})_i$	will := $(inf_{ph}\backslash_r\ vp_{ph})_i$
kunnen := $(inf_{ph}/lh\ inf_{ph})_i$	dürfen := $(inf_{ph}\backslash_r\ inf_{ph})_i$
lezen := $(np_{ph}\backslash_r\ inf_{ph})_w$	lesen := $(np_{ph}\backslash_r\ inf_{ph})_w$

We present the derivations schematically, leaving out sortal decoration, and omitting the last step, where wrapping reduces to simple dependency adjunction, under the appropriate sortal conditions.

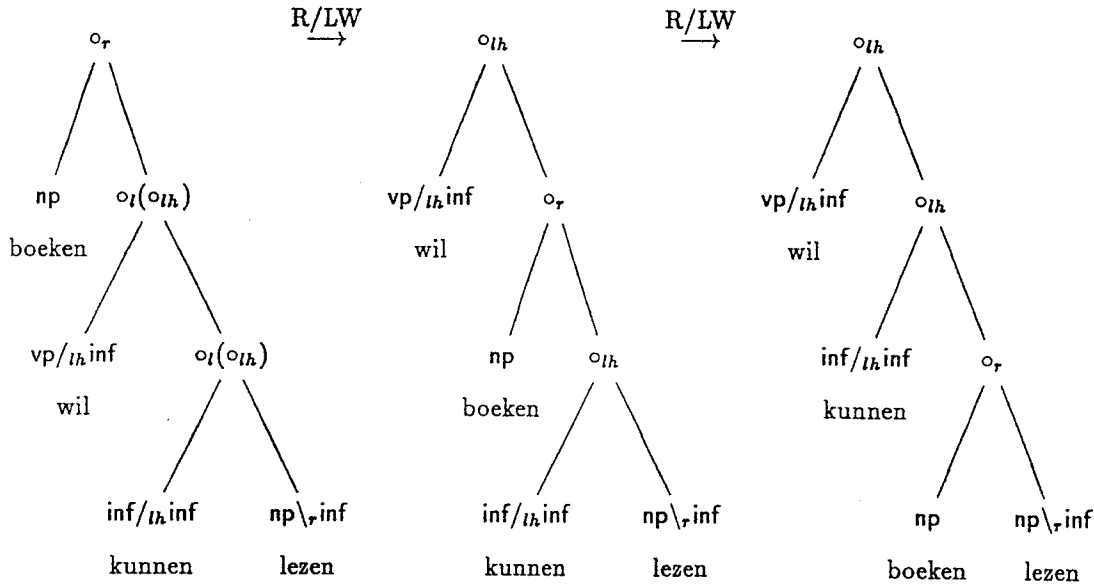


Figure 2: Verb-raising: Dutch

Our analysis of head wrapping can be compared with the initial categorial attempts to deal with the crossed dependencies of Dutch in terms of so-called *mixed composition* rule schemata: variants of functional composition where the directionality of the functors does not harmonize (cf. [24]):

$$Bx : A/B \bullet C \backslash B \longrightarrow C \backslash A \quad Bx' : B/C \bullet B \backslash A \longrightarrow A/C$$

The combinator Bx is not valid as a theorem about the concatenation operators $'/'$, $'\backslash'$. It is valid in LP, a type system with a global structural rule of Permutation, where the distinction between $'/'$ and $'\backslash'$ can no longer be made. But in

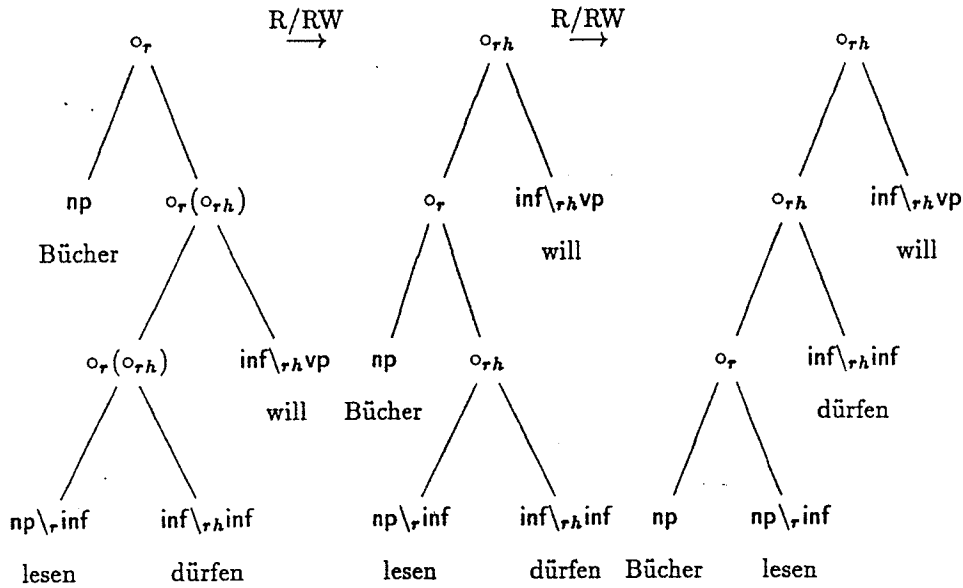


Figure 3: Verb-raising: German

that system, any directional variant of Bx which preserves the polarities of the subtypes is valid. In order to single out just Bx and its symmetrical variant Bx', Steedman postulates two principles (Directional Consistency and Directional Inheritance) which together ensure that in cases of 'partial combination' (functional composition) the directionality requirements of the principal functor and of the transmitted argument(s) are respected. In the account of head-wrapping presented here, the mixed composition laws are derivable as theorems involving the head-wrapping and dependency implications. Also, the principles of Directional Consistency and Directional Inheritance follow from the logic of the types — they do not have to be stipulated as postulates. Below is a derivation for mixed composition as it occurs in the Dutch crossed dependency construction.

$$\begin{array}{l}
 \frac{np \Rightarrow np \quad inf \Rightarrow inf}{(np, np \backslash_r inf \Rightarrow inf)} L \backslash_r \quad \frac{vp \Rightarrow vp}{(vp /_{lh} inf, (np, np \backslash_r inf)^r)^{lh} \Rightarrow vp} L /_{lh} \\
 \frac{(vp /_{lh} inf, (np, np \backslash_r inf)^r)^{lh} \Rightarrow vp}{(np, (vp /_{lh} inf, np \backslash_r inf)^{lh})^r \Rightarrow vp} R / LW \\
 \frac{(np, (vp /_{lh} inf, np \backslash_r inf)^{lh})^r \Rightarrow vp}{(vp /_{lh} inf, np \backslash_r inf)^{lh} \Rightarrow np \backslash_r vp} R \backslash_r
 \end{array}$$

Schematically, in 'Geach' version, we have the following valid type transition (and the symmetric case):

$$A /_{lw} B \Rightarrow (C \backslash_r A) /_{lw} (C \backslash_r B)$$

We close this section with some remarks about the interaction between verb raising and adverb placement. As we saw above, verb phrase adverbials have a

certain flexibility as to where they occur within the phrase they have in their scope semantically. We repeat some typical examples.

(dat je) snel je armen in de lucht moet gooien
 that one quickly one's arms in the air must throw
 (dat je) je armen snel in de lucht moet gooien
 *(dat je) je armen in de lucht snel moet gooien

We propose to account for this flexibility by treating adverbs not only as prefixes to the VP, but also as head-wrapping infixes. They differ in two respects from the verb-raising triggers we discussed above. First, on the level of resource management, whereas the verb-raising triggers are *heads* with respect to the host they infix into, adverbials will be typed as *dependents* with respect to the verb phrase they modify. This is the difference between the implications /*ih* and /*ld*. Secondly, on the prosodic level, we key the base cases for the Inclusion Axiom for verb-raising triggers and adverbial infixes to different prosodic sorts:

(VR triggers) : $(A_i \circ_l B_v)_v \rightarrow (A_i \circ_{lh} B_v)_{ph}$
 (VP adverbs) : $(A_{adv} \circ_r B_v)_{v'} \rightarrow (A_{adv} \circ_{ld} B_v)_{ph}$

Verb-raising triggers of sort *i* are adjoined to the verb-cluster level *v*. Adverbials (sort *adv*) are adjoined at the sort level *v'*, which is the sort we'll assign to 'impenetrable' phrasal verbal composites such as the combination *in de lucht gooien*. Below we give the lexical type declarations, and the abbreviations used in the schematic derivation.

moet : (vp/*ih*iv)_i (= *A_i*)
snel : (iv/*ld*iv)_{adv} (= *B_{adv}*)
je armen : np_{ph} (= *C_{ph}*)
in de lucht : pp_{ph} (= *D_{ph}*)
gooien : (pp_{ph}_rtv_v)_v (= *E_v*)

The end-sequent in the following partial derivation corresponds to the surface rendering *(dat je) je armen snel in de lucht moet gooien*. We trace the Interaction and Inclusion rules up to the point where the conditions for elimination of the

logical connectives are met.

$$\begin{array}{c}
(A_i, (B_{adv}, (C_{ph}, (D_{ph}, Ev)_v^r)_v^r)_v^r)_v^r)^{ld})_{ph}^{lh} \Rightarrow vp \\
\hline
(A_i, (C_{ph}, (B_{adv}, (D_{ph}, Ev)_v^r)_v^r)_v^r)_v^r)^{ld})_{ph}^{lh} \Rightarrow vp \\
\hline
(A_i, (C_{ph}, (B_{adv}, (D_{ph}, Ev)_v^r)_v^r)_v^r)_v^r)^{lh} \Rightarrow vp \\
\hline
(C_{ph}, (A_i, (B_{adv}, (D_{ph}, Ev)_v^r)_v^r)_v^r)_v^r)^{lh} \Rightarrow vp \\
\hline
(C_{ph}, (B_{adv}, (A_i, (D_{ph}, Ev)_v^r)_v^r)_v^r)_v^r)^{lh} \Rightarrow vp \\
\hline
(C_{ph}, (B_{adv}, (D_{ph}, (A_i, Ev)_v^r)_v^r)_v^r)_v^r)^{lh} \Rightarrow vp \\
\hline
(C_{ph}, (B_{adv}, (D_{ph}, (A_i, Ev)_v^r)_v^r)_v^r)_v^r)^{lh} \Rightarrow vp
\end{array}$$

It may be interesting to compare the head wrapping analysis with earlier proposals to treat verb-raising in terms of a *lexical* rule of Disharmonic Composition/Division. The lexical rule analysis has been proposed as an alternative to syntactic rule schemata Bx, Bx'. The addition of the latter to a calculus which is order sensitive has the effect of destroying the order-sensitivity of the system: schemata such as Bx, Bx' — invalid with respect to the interpretation of '/', '\ — make the logic degenerate into the permutation-closed type system LP (cf. [16]). The collapse is the result of the free interaction of the invalid Bx, Bx' with standard valid type transitions. This interaction rests on the fact that the derivability relation is transitive. Pushing the effect of Bx, Bx' back into the lexicon would block free interaction with the otherwise valid type transitions in the 'syntactic' module.⁵

$$\frac{A/B \rightarrow C/D}{A/B \rightarrow (E \setminus C) / (E \setminus D)} \quad \frac{B \setminus A \rightarrow D \setminus C}{B \setminus A \rightarrow (D/E) \setminus (C/E)}$$

The recursive schemata above induce a form of lexical polymorphism: on the basis of an initial type assignment (say, vp/inf, for a verb-raising trigger) an infinite number of derived lexical assignments can be obtained via the disharmonic division schemata. Derivation of a verb-raising construction, on this view, would involve a 'lexical' stage of guessing the appropriate type-assignment for the verb-raising triggers, followed by a 'syntactic' stage, where type inference obeys the usual laws of the order-sensitive mode of inference, such as by L or one of its relatives.

The major problem of the 'lexical' analysis is that it would lead one to expect that verb-raising does not interact with syntax proper. But such interaction does

⁵The associated semantics is that of the Geach rule (functional composition):

$$\frac{t \rightarrow u}{t \rightarrow \lambda x \lambda y. u(x(y))}$$

take place: the interplay of verb-raising and adverbial modification is a clear case in point. Maintaining the ‘lexical’ analysis in the light of these facts would force one to assume that adverbial modification in itself is a ‘lexical’ process: adverbs are added to the lexical type assignment via a recursive type-lifting rule schema. This consequence has indeed been accepted by defenders of the ‘lexical’ analysis. We think it leads to an unattractive demarcation between lexicon and syntax, and an unrestrictive notion of ‘possible lexical rule’.⁶ The head-wrapping analysis proposed in this paper formulates a *logic* for the wrapping operators so that they can freely participate in the process of grammatical inference.

4 Head attraction / complement postponement

We move now to a discussion of the (\star) variety of non-adjacent composition: the case where a configuration Δ is in construction with $\Gamma[\Delta']$ whereas in fact Δ forms a unit with the subconfiguration Δ' . English displays a variety of constructions involving communication between a head and a modifier or complement to its right. We focus here on two cases: the interaction of nouns with complex modifiers, as in *easy book to read* and *different person than Smith*; and extraposition from np, as in *a paper appeared about the adjacency parameter*. These discontinuous constructions in English are sometimes optional, sometimes obligatory, under conditions whose precise formulation poses a challenge to grammatical analysis. In what follows, we address aspects of this challenge. And we argue that the multimodal framework provides some explanatory insights into the basic properties of these constructions.

4.1 Attributive adjectives and adjective phrases

We begin our discussion with attributive adjectives and adjective phrases, which exhibit the behavior displayed below:

- angry rabbit / *rabbit angry / *angry rabbit at the world / rabbit angry at the world
- eager rabbit / *rabbit eager / *eager to run rabbit / *eager rabbit to run / rabbit eager to run
- different rabbit / *rabbit different / *different than Bugs rabbit / different rabbit than Bugs / rabbit different than Bugs

⁶Carpenter [5] shows that recursive lexical rule schemata of the type discussed here make an otherwise restrictive system degenerate into a type-0 formalism.

- happier rabbit / *rabbit happier / *happier than Bugs rabbit / happier rabbit than Bugs / rabbit happier than Bugs
- easy technique / *technique easy / easy to use technique / easy technique to use / technique easy to use

All of the adjectives here — *angry, eager, angry, different, happier, easy* — occur preminally when they are uncomplemented. On these grounds, we provisionally assign them the type n/n . With complements, however, we see a richer array of patterns.

The general pattern with *pp*-complements is illustrated by *angry*, which must occur post-nominally when it is complemented. What is the nature of the adjective + *pp* combination? In predicative position, *angry* need not be adjacent to its *pp*-complement. Along with *he was angry at the world all day* (an instance of connected complementation), we have *he was angry all day at the world* (where the complement is extraposed). Similarly, we find extraposition of the complement in *wh*-constructions such as *how angry was he at the world*. On these grounds, we regard *angry at the world* as an instance of a configuration $(\cdot, \cdot)^{la}$ — ‘*la*’ for ‘left attraction’ — that does not demand string adjacency. This configuration communicates with left- and right-headed structures according to the postulates below. In these postulates, we use $L\cdot$ to denote the binary, left-headed product which is *right associative*: that is, where \circ_h is any headed product, it satisfies

$$(A \circ_l (B \circ_h C)) \longrightarrow ((A \circ_l B) \circ_l C)$$

The distinction between $(\cdot, \cdot)^l$ and non-associative $(\cdot, \cdot)^l$ plays an important role in our account of extraposition: the lack of a counterpart for the non-associative product to the rule *Ass L·/LA* blocks recursion down the right branch of any non-associative structure and thus yields a variety of empirical consequences, including the Right Roof constraint on extraposition from *np*.

$$\begin{aligned} \text{Comm } L\cdot/LA &: ((A \circ_l B) \circ_{la} C) \longrightarrow ((A \circ_{la} C) \circ_l B) \\ \text{Comm } R/LA &: ((A \circ_r B) \circ_{la} C) \longrightarrow ((A \circ_{la} C) \circ_r B) \\ \text{Ass } L\cdot/LA &: ((A \circ_l B) \circ_{la} C) \longrightarrow (A \circ_l (B \circ_{la} C)) \end{aligned}$$

In the sequent presentation, these take the forms:

$$\frac{\Gamma[(\Delta_1, \Delta_3)^{la}, \Delta_2]^l \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^l, \Delta_3]^{la} \Rightarrow A} \text{Comm } L\cdot/LA \quad \frac{\Gamma[(\Delta_1, \Delta_3)^{la}, \Delta_2]^r \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^r, \Delta_3]^{la} \Rightarrow A} \text{Comm } R\cdot/LA$$

$$\frac{\Gamma[(\Delta_1, (\Delta_2, \Delta_3)^{la})^l] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^l, \Delta_3]^{la} \Rightarrow A} \text{Ass } L\cdot/LA$$

These allow the schematic derivation steps:

$$\begin{array}{l} \vdots \\ \frac{(\text{how, (angry, (at the world)}^{la})^l, (\text{was he}))^l \Rightarrow q}{(((\text{how, angry})^l, (\text{at the world}))^{la}, (\text{was he}))^l \Rightarrow q} \text{ Ass } L \cdot / LA \\ \frac{(((\text{how, angry})^l, (\text{was, he}))^l, (\text{at the world}))^{la} \Rightarrow q}{(((\text{how, angry})^l, (\text{was, he}))^l, (\text{at the world}))^{la} \Rightarrow q} \text{ Comm } L \cdot / LA \end{array}$$

On the other hand, the mode of combination between noun and post-nominal *angry at the world* requires connectivity, since extraposition of post-nominal *ap*'s is impossible (though care must be exercised to distinguish subject-bound predicative constructions such as *Smith/A man came in tired* from extraposed structures such as *A man/*Smith answered from Toulouse*, which require a subject *np* compatible with the extraposed complement; in the latter structures, extraposed *ap*'s are impossible, as exemplified by *A man took the floor from Toulouse/*angry at the Whig policies*). On these grounds, then, we may assign to *angry at np* the type $n \setminus_l n$, and to *angry* itself the type $(n \setminus_l n) /_{la} at - p$, where $/_{la}$ is the residual of the product $(\cdot, \cdot)^{la}$: left-headed, non-adjacency-demanding.

The multimodal setting makes it possible to tie the ordering conditions of the residual slashes of prenominal *angry* and post-nominal *angry at np* to prosodic sorts in an interesting way, a way which leads to a simplification of the lexical specification of attributive adjectives and phrases built up from them. Although in general there are many natural linking laws between binary products, such natural links do not always exist, or do not exist uniquely. For example, the following links between the right-headed non-associative structure $(\cdot, \cdot)^r$ and the non-headed, non-associative structure $(\cdot, \cdot)^*$, and between the latter and the (non-associative) commutative structure $(\cdot, \cdot)^\otimes$ are natural, since they arise purely by 'forgetting'—first headedness, then order:

$$(\Delta, \Delta')^r \longrightarrow (\Delta, \Delta')^* \longrightarrow (\Delta, \Delta')^\otimes$$

Similarly, there is a natural link between a 'pointed' commutative product $(\cdot^*, \cdot)^\otimes$ (equivalently, $(\cdot, \cdot^*)^\otimes$) and the simple commutative product $(\cdot, \cdot)^\otimes$, according to the simple rule: forget the pointing. But there is no natural link between right- or left-headed structures and pointed commutative products: this is because while one may *forget* the head, one must then *add* the point. Nevertheless, just because there is no natural link between these structures, there may still be linking principles.

Consider the following principles, which govern relaxation of right- and left-headed structures into pointed commutative structures:

$$\text{Relax } R / \odot \frac{\Gamma[(\Delta_w, \Theta^*)^\otimes] \Rightarrow A}{\Gamma[(\Delta_w, \Theta)^r] \Rightarrow A} \quad \text{Relax } L / \odot \frac{\Gamma[(\Theta^*, \Delta_{ph})^\otimes] \Rightarrow A}{\Gamma[(\Theta, \Delta_{ph})^l] \Rightarrow A}$$

Now instead of assuming that *angry* by itself is of type $n /, n$ and that *angry at np* is of type $n \setminus_l n$, we may take them to be of the same type, distinguished only

by prosodic sorts: *angry* of type $(n/\circ n^*)_w$, and *angry at np* of type $(n/\circ n^*)_{ph}$. If we then make the reasonable assumption that commutative products are only induced on expressions as a result of natural and specified communication principles — reasonable because the expressions in question are in fact linearly ordered in time — this prosodic distinction between *angry* and *angry at np* is sufficient to yield the paradigm: *angry politician*, **angry at the Whigs politician*, **angry politician at the Whigs*, **politician angry*, *politician angry at the Whigs*. We leave the analysis of the individual cases to the reader.

To achieve this result on the basis of a single lexical specification for the word *angry*, we need an account of optional arguments.⁷ Morrill e.a. [20] have shown how to do this using modalities in the style of linear logic. The basic idea is easily adaptable to the multimodal setting. Three steps are necessary. First, for each structural product α , we add a new structural product (α) . Second, we need to ensure that any product α relaxes into (α) . Third, we need to ensure that any residual type — say, $A/(\alpha)B$ — relaxes into the an appropriate (α) configuration — say, $(A/(\alpha)B, B)^{(\alpha)}$. The second two steps point the way to the obvious frame conditions and corresponding logical principles, exemplified by the structural rules below.

$$\text{Relax } L/() \frac{\Gamma[(\Delta_1, \Delta_2)^{(l)}] \Rightarrow A}{\Gamma[(\Delta_1, \Delta_2)^l] \Rightarrow A} \quad \text{Opt } A/(\iota)B \frac{\Gamma[(A/(\iota)B, B_\top)^{(l)}] \Rightarrow C}{\Gamma[A/(\iota)B] \Rightarrow C}$$

Relative to these assumptions, we now associate *angry* with the type:

$$((n/\circ n^*)_x \sqcup w /(\iota_a) pp_x)w.$$

Together with our communication postulates, this assignment leads to the following behavior:

$$\begin{aligned} (angry, politician)^r &\Rightarrow n && \text{Opt, Relax } R/\circ \\ (politician, (angry, at the Whigs)^{\iota_a})^l &\Rightarrow n && \text{Relax } \iota_a/(), \text{ Relax } L/\circ \\ politician \text{ angry} &\not\Rightarrow n \\ angry \text{ politician at the Whigs} &\not\Rightarrow n \\ angry \text{ at the Whigs politician} &\not\Rightarrow n \end{aligned}$$

Eager. The case of uncomplemented *eager* and complemented *eager to please* is exactly analogous. The uncomplemented adjective appears obligatorily pre-nominally; the complemented phrase appears obligatorily post-nominally, under obligatorily adjacency; and the infinitive complement is extraposable in structures such as *how eager is he to please?*. We treat it in exactly the same way as *angry*. Thus, *eager* is associated with the type:

⁷In some cases, it is reasonable not to distinguish between ‘optional arguments’ and ‘modifiers’. But it need not be reasonable to conflate these two concepts in general. The topics discussed here are treated more smoothly from the point of view of optional arguments.

$$((n/\circ n^*)_x \sqcup w /_{(la)} \text{inf}_x)w.$$

Happier/Different. Consider now the cases of *happier (than np)* and *different (than np)*. Abstractly, we think of these adjectives as combining with a *than*-phrase (in the *la*-mode, to the right) and an *n* to form an *n*. If we allow the adjective to combine with these arguments in either order—that is, first with the *than*-phrase and then with the *n*, or first with the *n* and then with the *than*-phrase—and if we require that the expression that combines with the *n* satisfy the same prosodic constraints imposed for *angry (at np)* above, then we allow exactly *different / happier rabbit, rabbit different/happier than Bugs*, and *different/happier rabbit than Bugs*. On this account, then, the *different/happier* class of adjectives differs from the *angry* class in that the latter imposes a strict order on how it combines with its arguments—first, the *pp*-complement, and then the noun—with the consequence that the relative of order of *angry (at np)* and *n* is specifiable only after a decision has been made concerning the occurrence of the complement; in the former case, however, the adjective may combine with the noun first (in which case, it must precede the noun) and then combine (optionally) with a *than*-phrase, or it may combine first with the *than*-phrase and then combine with the noun (in which case, as a phrase, it must follow the noun, exactly as in the case of *angry at np*). We denote the type of *different* in the following way:

$$n \begin{array}{l} /_{\circ}; n^* \\ /_{(la)} \text{than-p} \end{array}$$

The associated frame conditions require that expressions of this type belong to the intersection of the sets of expressions of types $(n /_{\circ} n^*) /_{(la)} \text{than-p}$ and $(n /_{(la)} \text{than-p}) /_{\circ} n^*$. Prosodic sorts are to be interpreted accordingly. We suppress the complete statement of the associated structural properties.

It is possible, of course, to have cases involving both a *pp*-complement and a comparative phrase complement. One possible extension of the above system to such cases would allow the comparative adjective to combine with (optional) *pp*-complement, (optional) *than*-phrase complement, and (obligatory) *n* in any order, giving:

happier man with these results than anyone else
happier man than anyone else with these results
happier man than anyone else
happier man with these results
man happier with these result (than anyone else)
man happier than anyone else (with these results)

Easy. The *easy* class of adjectives exhibits all the behavior observed with the *different/happier* class. The members of this class can be assimilated to the current system by allowing them the additional privilege of combining with an

expression of type $\text{inf} /_1 \text{np}$ and prosodic sort w to form an expression of type $n /_0$ n^* of prosodic sort w . This assumption allows these cases to occur preminimally. The reader may find it amusing to show how our assumptions provide an account of the following interesting data:

- easy to understand proof of this theorem*
- proof of this theorem easy to understand*
- * *proof easy to understand of this theorem*
- * *proof easy of this theorem to understand*

4.2 Extraposition from noun phrases

We now turn to extraposition from np , starting with extraposition from subject position, as in the alternants *A book on that subject appeared recently* \sim *A book appeared recently on that subject*. We call the sub-expression *a book* the 'host' and the sub-expression *on that subject* the 'guest'. The extraposed variant of the alternation is subject to independent constraints involving properties of the host and the guest. We mention two of these briefly, then turn to the question of how communication between host and guest governs the possibility of the extraposition alternation.

There is an important pragmatic restriction on extraposition, noted by Guéron [12]: extraposition is impossible if the referent of the host is referentially determinate. While *A man object from Toulouse* is possible, interposing a deictic anacoluthon between host and guest fixing the referent of the host is impossible: thus, **A man—that guy over there to Amy's left—objected from Toulouse*.

There are also interesting differences associated with predicates: individual-level predicates in the sense of Carlson [4] are impossible, as demonstrated by the contrast (see [22]) between *No proof will ever be found*/**sound of this theorem*.

Not every post-nominal constituent is extraposable. There are categorial restrictions on what constituents are extraposable, illustrated in the table below:

category	examples
pp	<i>a man from Toulouse objected</i> \sim <i>a man objected from Toulouse</i>
ap	<i>a man present objected</i> \sim * <i>a man objected present</i>
rel	<i>a man who/that didn't like the proposal objected</i> \sim <i>a man objected who/that didn't like the proposal</i>
rel	<i>a man who/that/∅ I think you know objected</i> \sim <i>a man objected who/that/*∅ I think you know</i>
vp-ing	<i>a man walking by came in</i> \sim * <i>a man came in walking by</i>
past-prt	<i>a man suspected of perjury objected.</i> \sim * <i>a man objected suspected of perjury.</i>

The distinction between extraposable and non-extraposable constituents is expressed directly in the type system adopted here: extraposable constituents depend on products in the *la*-mode; non-extraposable constituents depend on products in the *l*-mode.

Type communication in extraposition. Consider now an allowable host for extraposition—such as *a person*—and a compatible post-nominal complement. If the complement is not an extraposable, then we regard the complement as being of type $n \backslash_l n$. Adjacency is enforced. On the other hand, if the complement is of the extraposable kind, then we regard it as being of type $n \backslash_{la} n$. Given the communication rules *Comm R/LA* and *Ass L/LA*, we have the following proof for *A man objected from Toulouse*:

$$\begin{array}{l} \vdots \\ \frac{((np \backslash_l n, (n, n \backslash_{la} n)^{la})^l, np \backslash_r s)^r \Rightarrow s \quad \text{antecedent residual introductions}}{(((np \backslash_l n, n)^l, n \backslash_{la} n)^{la}, np \backslash_r s)^r \Rightarrow s \quad \text{Ass } L \cdot /LA} \\ \hline \frac{(((np \backslash_l n, n)^l, np \backslash_r s)^r, n \backslash_{la} n)^{la} \Rightarrow s \quad \text{Comm } R/LA}{} \end{array}$$

This system also allows complements of *pre-nominal* adjectives to be extraposed, as in *A different man objected than Ilya*. The proof is the same as above, modulo the replacement of *man*, of type *n*, by *different man* of type $n /_{(la)} \text{than-p}$ and of the phrase *from Toulouse* of type $n \backslash_{la} n$ by the *than-p than Ilya*.

The same method yields double extraposition cases as well, such as *A book appeared on that subject by a well-known figure*.

This system does not, however, allow extraposition of the *than*-phrase in **A man different objected from Ilya* nor of the *about*-phrase in *A picture of a book appeared about French cooking*. In order for postnominal *different* to combine with *man*, it must first combine with the extraposed constituent to form a phrase which matches the conclusion of the inference rule *Relax L/⊙*. Thus, we must have the configuration:

$$\dots \left(n, n \begin{array}{l} /_{\odot} n^* \\ /_{(la)} \text{than-p} \end{array} \right)^l \dots$$

The postulates *Comm R/LA* and *Ass L/LA* allow communication between the extraposed *than*-phrase and the structure corresponding to $((man, different)^l, \text{than Ilya})^{la}$. But no further communication is possible, and the proof fails at this point. Similarly, with respect to extraposition from within a *pp*-complement or modifier, the preposition itself combines with its object in the mode $(pp \backslash_r np, np)^r$ —which does not communicate with the *la*-mode in a way comparable to the communicative postulate *Ass L/LA*.

There are other syntactic aspects of extraposition worthy of scrutiny. For example, the extraposed phrase must appear to the right of obligatory arguments of the predicate (if there are any). To see this, compare the variants below:

A rumor of the battle at Tuzla quickly reached the troops
A rumor quickly reached the troops of the battle at Tuzla.
A rumor quickly reached of the battle at Tuzla the troops (under Chichikov).

The last of these is emphatically impossible.

In traditional forms of generative grammar, explanation of issues of this kind is typically sought in the study of comparative tree geometry. Our approach is somewhat different: we simply impose a particular type structure which controls the dynamic interaction of the resources of the sentence. Suppose, then, that the predicate *reached* is assigned the type $(np \setminus_r s)/_l np$. On this type assignment, we have:

$$(np, ((np \setminus_r s)/_l np, np)^l)^r \Rightarrow s$$

The extraposition case is based on the validity of:

$$(((np/_l n, n)^l, ((np \setminus_r s)/_l np, np)^l)^r, n \setminus_{la} n)^{la} \Rightarrow s$$

This much is straightforward. But the second case is straightforward as well: there simply is no structure jointly compatible with the demands of the individual resources. To observe the source of these incompatible demands, note that the extraposed constituent must combine in the *la*-mode, so that the substructure corresponding to *reached of the battle at Tuzla* must be $((np \setminus_r s)/_l np, n \setminus_l n)^{la}$. Adding the other information in any mode yields one of two possible bracketings:

$$\begin{aligned} & (((np/_l n, n)^l, ((np \setminus_r s)/_l np, n \setminus_{la} n)^{la})^r, np)^l \\ & ((np/_l n, n)^l, (((np \setminus_r s)/_l np, n \setminus_{la} n)^{la}, np)^r)^l \end{aligned}$$

But neither of these is resolvable: in fact, in neither case is either of the accessible type constructors removable.

An interesting contrast, which follows directly from assumptions adopted here, is the fact that *pp*'s extraposed from within a verbal argument need not appear following all verbal arguments, as shown below:

The NATO command sent a bulletin out about the events to all parties involved in the siege of Sarajevo.

The possibility of this kind of 'vp-internal' extraposition is easily treated. A sketch of the relevant proof substructure:

$$\frac{((np \setminus_r s)/_l ((np \otimes out) \cdot to-p), (((np/_l n, (n, n \setminus_{la} n)^{la}), out)^r, to-p)^r)^l \Rightarrow np \setminus_r s}{((np \setminus_r s)/_l ((np \otimes out) \cdot to-p), (((np/_l n, n)^l, n \setminus_{la} n)^{la}), out)^r, to-p)^l \Rightarrow np \setminus_r s} \\ ((np \setminus_r s)/_l ((np \otimes out) \cdot to-p), (((np/_l n, n)^l, out)^r, n \setminus_{la} n)^{la}, to-p)^l \Rightarrow np \setminus_r s$$

5 Conclusion

We have presented a multimodal architecture for *combining* logics of categorial type inference. The central research theme for the mixed architecture is *communication*: what principles govern the process of reasoning when different modes of grammatical composition are combined? We have investigated two such principles. Inclusion principles impose a natural order on the distinct modes of composition in terms of the constraints they impose on the structure of the linguistic resources. Interaction principles regulate the communication in mixed configurations. We have shown how *non-adjacent* modes of composition arise in the multimodal setting through inclusion and interaction principles. We have illustrated the potential of the approach with an analysis of head adjunction and head attraction phenomena — phenomena that have resisted categorial treatment in the individual classical type systems.

References

- [1] Bach, E. (1981) 'Discontinuous constituents in generalized categorial grammar'. *NELS* 11, 1–12.
- [2] Bach, E. (1981) 'Some generalizations of categorial grammars'. In Landman & Veltman (eds.) *Varieties of Formal Semantics*, Foris, Dordrecht, 1–23.
- [3] Belnap, N.D. (1982) 'Display Logic'. *Journal of Philosophical Logic*, 11, 375–417.
- [4] Carlson, G. (1977) *Reference to kinds in English*. Ph.D. dissertation. University of Massachusetts at Amherst.
- [5] Carpenter, B. (1991) 'The generative power of categorial grammars and head-driven phrase structure grammars with lexical rules', *Computational Linguistics*, 17(3), 301–313.
- [6] Carpenter, B. (1992) *The Logic of Typed Feature Structures*. Cambridge University Press.
- [7] Došen, K. (1988,1989) 'Sequent systems and groupoid models'. *Studia Logica* 47, 353–385, 48, 41–65.
- [8] Došen, K. (1992) 'A brief survey of frames for the Lambek calculus'. *Zeitschr. f. math. Logik und Grundlagen d. Mathematik* 38, 179–187.
- [9] Dowty, D. (1991) 'Towards a minimalist theory of syntactic structure'. In Sijtsma and Van Horck (eds.) *Discontinuous constituency*. Mouton-De Gruyter (to appear).

- [10] Dunn, J.M. (1973) 'A "Gentzen system" for positive relevant implication'. *Journal of Symbolic Logic*, 38, 356–357.
- [11] Girard, J.-Y. (1993) 'On the unity of logic'. *Annals of Pure and Applied Logic* 59, 201–217.
- [12] Guéron, J. (1980) 'On the syntax and semantics of PP extraposition'. *Linguistic Inquiry*, 11, 637–678.
- [13] Hinrichs, E. and T. Nakazawa (1989) 'Flipped out: AUX in German'. *CLS Papers* 25, 187–202.
- [14] Hoeksema, J. (1991) *A categorial theory of reanalysis phenomena*. ms University of Groningen.
- [15] Kurtonina, N. & M. Moortgat (in prep) 'Sorts and types'. To appear in Abrusci and Moortgat (eds.) *Proceedings First Rome Workshop on Linear Logic and Lambek Calculus*.
- [16] Moortgat, M. (1988) *Categorial Investigations*. Foris, Dordrecht.
- [17] Moortgat, M. (1992) 'Labelled deductive systems for categorial theorem proving'. In *Proceedings Eighth A'dam Colloquium*, ILLC, Amsterdam.
- [18] Moortgat, M. and G. Morrill (1991) 'Heads and phrases. Type calculus for dependency and constituency'. Ms.OTS Utrecht (to appear in JoLLI).
- [19] Moortgat, M. and R. Oehrle (1993) *Logical parameters and linguistic variation. Lecture notes on categorial grammar*. 5th European Summer School in Logic, Language and Information. Lisbon.
- [20] Morrill, G., N. Leslie, M. Hepple, and G. Barry (1990) 'Categorial deductions and structural operations'. In G. Barry and G. Morrill, eds. *Edinburgh working papers in cognitive science 5: Studies in categorial grammar*. 1-21. Centre for Cognitive Science. University of Edinburgh.
- [21] Noord, G.J. van and G. Bouma (1993) 'Recursion in the lexicon'. To appear in *Proceedings CLIN93*.
- [22] Oehrle, R.T. (1977) 'Comments on the paper by Selkirk'. In A. Akmajian, P. Culicover, and T. Wasow, eds. *Formal Syntax*. 317–327. Academic Press, New York.
- [23] Steedman, M. (1988) 'Combinators and grammars'. In Oehrle, Bach and Wheeler (eds) *Categorial Grammars and Natural Language Structures*. Reidel, Dordrecht, 417–442.

- [24] Steedman, M. (1985) 'Dependency and coördination in the grammar of Dutch and English'. *Language* 61, 523-568.
- [25] Steedman, M. (1993) 'Categorial grammar. Tutorial overview'. *Lingua* 90,221-258.

