Incremental Processing and Acceptability

Glyn Morrill*
Universitat Politècnica de Catalunya

We present a left to right incremental algorithm for the processing of Lambek categorial grammar by proof net construction. A simple metric of complexity, the profile in time of the number of unresolved valencies, correctly predicts a wide variety of performance phenomena: garden pathing, left to right quantifier scope preference, centre embedding unacceptability, preference for lower attachment, and heavy noun phrase shift.

Introduction

Contemporary linguistics rests on abstractions and idealizations which, however fruitful, must eventually be integrated in a mature cognitive science with human computational performance in language use. In this paper we consider the case of language processing on the basis of Lambek categorial grammar (Lambek 1958). We argue that an incremental algorithm of proof net construction affords an account of various processing phenomena, including:

(1)
- garden pathing
- left to right quantifier scope preference
- centre embedding unacceptability

Garden pathing (Bever 1970) is illustrated by the following contrasts:

(2) a. The horse raced past the barn.
b. ?The horse raced past the barn fell.

(3) a. The boat floated down the river.
b. ?The boat floated down the river sank.

(4) a. The dog that knew the cat disappeared.
b. ?The dog that knew the cat disappeared was rescued.

Typically, although the 'b' sentences are perfectly well-formed they are perceived of as being ungrammatical due to a strong tendency to interpret their initial segments as in the 'a' sentences.

Left to right quantifier scope preference is illustrated by:

(5) a. Someone loves everyone.
b. Everyone is loved by someone.

Both sentences exhibit both quantifier scopings:

* Departament de Llenguatges i Sistemes Informàtics, Mòdul C 5 - Campus Nord, Jordi Girona Salgado 1-3, E-08034 Barcelona. E-mail: morrill@lsi.upc.es; http://www-lsi.upc.es/~glyn/.

(6) a. $\exists x \forall y(love\ y\ x)$
b. $\forall y \exists x(love\ y\ x)$

However, while the dominant reading of (5a) is (6a), that of (5b) is (6b), i.e. the preference is for the first quantifier to have wider scope. Note that the same effect is observed when the nature of the quantifications is swapped:

(7) a. Everyone loves someone.
b. Someone is loved by everyone.

While both sentences in (7) have both quantifier scopings, the preferred readings give the first quantifier wide scope.

Centre embedding unacceptability is illustrated by the fact that while the nested subject relativizations of (8) exhibit little variation in acceptability, the nested object relativizations (9) exhibit a sever deterioration in acceptability (Chomsky 1965, ch. 1).

(8) a. The dog that chased the cat barked.
b. The dog that chased the cat that saw the rat barked.
c. The dog that chased the cat that saw the rat that ate the cheese barked.

(9) a. The cheese that the rat ate stank.
b. ?The cheese that the rat that the cat saw ate stank.
c. ??The cheese that the rat that the cat that the dog chased saw ate stank.

We argue that a single simple metric of categorial processing complexity explains these and other performance phenomena.

1 Lambek calculus

We shall assume some familiarity with Lambek categorial grammar as presented in e.g. Moortgat (1988), Morrill (1994), Moortgat (1997) or Carpenter (1998), and limit ourselves here to reviewing some technical and computational aspects.

The types, or (category) formulas, of Lambek calculus are freely generated from a set of primitives by binary infix connectives / ("over"), \ ("under") (directional divisions) and $\cdot$ ("product"). With respect to a semigroup algebra $(L, +)$ (i.e. a set L closed under an associative binary operation $+$ of adjunction) each formula A is interpreted as a subset $[\![A]\!]$ of L by residuation as follows.

(10)
$$\begin{array}{rcl} [\![A{\cdot}B]\!] & = & \{s_1{+}s_2 | s_1 \in [\![A]\!] \wedge s_2 \in [\![B]\!]\} \\ [\![A\backslash B]\!] & = & \{s | \forall s' \in [\![A]\!], s'{+}s \in [\![B]\!]\} \\ [\![B/A]\!] & = & \{s | \forall s' \in [\![A]\!], s{+}s' \in [\![B]\!]\} \end{array}$$

A sequent, $\Gamma \Rightarrow A$, comprises a succedent formula A and an antecedent configuration Γ which is a a finite sequence of formulas.[1] A sequent is valid if and only if in all interpretations the ordered adjunction of elements inhabiting the antecedent formulas always yields an element inhabiting the succedent formula. The following Gentzen-style sequent presentation is sound and complete for this interpretation (Buszkowski 1986, Došen 1992), and indeed for free semigroups (Pentus 1994): hence the Lambek calculus

1 Officially the antecedent is non-empty, a detail we gloss over.

can make an impressive claim to be *the* logic of concatenation; the parenthetical notation $\Gamma(\Delta)$ represents a configuration containing a distinguished subconfiguration Δ.

(11)

a. $A \Rightarrow A$ id $\quad \dfrac{\Gamma \Rightarrow A \quad \Delta(A) \Rightarrow B}{\Delta(\Gamma) \Rightarrow B}\text{Cut}$

b. $\dfrac{\Gamma \Rightarrow A \quad \Delta(B) \Rightarrow C}{\Delta(\Gamma, A\backslash B) \Rightarrow C}\backslash\text{L} \qquad \dfrac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A\backslash B}\backslash\text{R}$

c. $\dfrac{\Gamma \Rightarrow A \quad \Delta(B) \Rightarrow C}{\Delta(B/A, \Gamma) \Rightarrow C}/\text{L} \qquad \dfrac{\Gamma, A \Rightarrow B}{\Gamma \Rightarrow B/A}/\text{R}$

d. $\dfrac{\Gamma(A, B) \Rightarrow C}{\Gamma(A{\cdot}B) \Rightarrow C}{\cdot}\text{L} \qquad \dfrac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A{\cdot}B}{\cdot}\text{R}$

By way of example, "lifting" A $\Rightarrow$ B/(A\B) is generated as follows:

(12)

$$\dfrac{\dfrac{\mathrm{A} \Rightarrow \mathrm{A} \quad \mathrm{B} \Rightarrow \mathrm{B}}{\mathrm{A}, \mathrm{A}\backslash\mathrm{B} \Rightarrow \mathrm{B}}\backslash\mathrm{L}}{\mathrm{A} \Rightarrow \mathrm{B}/(\mathrm{A}\backslash\mathrm{B})}/\mathrm{R}$$

And "composition" A\B, B\C $\Rightarrow$ A\C is thus:

(13)

$$\dfrac{\dfrac{\mathrm{A} \Rightarrow \mathrm{A} \qquad \dfrac{\mathrm{B} \Rightarrow \mathrm{B} \quad \mathrm{C} \Rightarrow \mathrm{C}}{\mathrm{B}, \mathrm{B}\backslash\mathrm{C} \Rightarrow \mathrm{C}}\backslash\mathrm{L}}{\mathrm{A}, \mathrm{A}\backslash\mathrm{B}, \mathrm{B}\backslash\mathrm{C} \Rightarrow \mathrm{C}}\backslash\mathrm{L}}{\mathrm{A}\backslash\mathrm{B}, \mathrm{B}\backslash\mathrm{C} \Rightarrow \mathrm{A}\backslash\mathrm{C}}\backslash\mathrm{R}$$

Every rule with the exception of Cut, where the Cut formula A does not appear in the conclusion, has exactly one connective occurrence less in its premisses than in its conclusion; Lambek (1958) proved Cut elimination —that every proof has a Cut-free counterpart— hence a decision procedure for theoremhood is given by backward-chaining proof search in the Cut-free calculus. The non-atomic instances of the id axiom are derivable from atomic instances by the rules for the connectives. But even in the Cut-free atomic-id calculus there is spurious ambiguity: equivalent derivations differing only in irrelevant rule ordering.

One approach to this problem consists in defining normal form derivations in which the succession of rule application is regulated (König 1989, Hepple 1990; see e.g. Hendriks 1993 ch. 3). Each sequent has a distinguished category formula (underlined) on which rule applications are keyed:

(14)

a. $\underline{P} \Rightarrow P$ id

b. $\dfrac{\Gamma \Rightarrow \underline{A} \quad \Delta(\underline{B}) \Rightarrow C}{\Delta(\Gamma, \underline{A\backslash B}) \Rightarrow C}\backslash\text{L} \qquad \dfrac{A, \Gamma \Rightarrow \underline{B}}{\Gamma \Rightarrow \underline{A\backslash B}}\backslash\text{R}$

c. $$\frac{\Gamma \Rightarrow \underline{A} \qquad \Delta(\underline{B}) \Rightarrow C}{\Delta(\underline{B/A}, \Gamma) \Rightarrow C}/L \qquad \frac{\Gamma, A \Rightarrow \underline{B}}{\Gamma \Rightarrow \underline{B/A}}/R$$

d. $$\frac{\Gamma \Rightarrow \underline{A} \qquad \Delta \Rightarrow \underline{B}}{\Gamma, \Delta \Rightarrow \underline{A{\cdot}B}}{\cdot}R$$

e. $$\frac{\Gamma(\underline{A}) \Rightarrow B}{\Gamma(A) \Rightarrow \underline{B}}$$

In the regulated calculus there is no spurious ambiguity, and provided there is no explicit or implicit antecedent product, i.e. provided $\cdot$L is not needed, $\Gamma \Rightarrow A$ is a theorem of the Lambek calculus iff $\Gamma \Rightarrow \underline{A}$ is a theorem of the regulated calculus. However, apart from the issue regarding $\cdot$L, there is a general cause for dissatisfaction with this approach: it assumes the initial presence of the entire sequent to be proved, i.e. it is in principle non-incremental; on the other hand, allowing incrementality on the basis of Cut would reinstate with avengance the problem of spurious ambiguity: for then what are to be the Cut formulas? Consequently, the sequent approach is ill-equipped to address the basic asymmetry of language, the asymmetry of its processing in time, and has never been forwarded in a model of the kind of processing phenomena cited in the introduction.

An alternative, combinatory, presentation would comprise combinatory schemata such as the following (together with a Cut rule, feeding one rule application into another):

(15)

a. $A, A\backslash B \Rightarrow B \qquad B/A, A \Rightarrow B$

b. $A \Rightarrow (B/A)\backslash B \qquad A \Rightarrow B/(A\backslash B)$

c. $A\backslash B, B\backslash C \Rightarrow A\backslash C \qquad C/B, B/A \Rightarrow C/A$

By a result of Zielonka (1981) the Lambek calculus is not axiomatizable by any finite set of combinatory schemata, so no such combinatory presentation can constitute *the* logic of concatenation in the sense of Lambek calculus. Combinatory categorial grammar (Steedman 1997) does not concern itself with the capture of all (or only) the concatenatively valid combinatory schemata, but from its inception (Ades and Steedman 1981) it has emphasized the capacity of schemata such as those of (15) to produce left-branching, and therefore incrementally processable, analyses, e.g. on a shift-reduce design. An approach, also based on regulation of the succession of rule application, to the associated problem of spurious ambiguity is given in Hepple and Morrill (1989) but again, to our knowledge, there is no predictive relation between incremental combinatory processing and the kind of processing phenomena cited in the introduction.

2 Proof nets

Lambek categorial derivations are usually presented in the style of natural deduction or sequent calculus. Here we concern ourselves with categorial proof nets (Roorda 1991) as the fundamental structures of proof in categorial logic, in the same sense that linear proof nets were originally introduced by Girard (1987) as the fundamental structures of proof in linear logic. (Cut-free) proof nets exhibit no spurious ambiguity and play the role in categorial grammar that parse trees play in phrase structure grammar.

A *polar category formula* is a Lambek categorial type labelled with *input* ($^\bullet$) or *output* ($^\circ$) polarity. A *polar category formula tree* is a binary ordered tree in which the leaves are labelled with polar atoms (*literals*) and each local tree is one of the following (*logical*) *links*:

(16)

a. $$\frac{A^\circ \quad B^\bullet}{A\backslash B^\bullet}\text{ii} \qquad \frac{B^\circ \quad A^\bullet}{A\backslash B^\circ}\text{i}$$

b. $$\frac{B^\bullet \quad A^\circ}{B/A^\bullet}\text{ii} \qquad \frac{A^\bullet \quad B^\circ}{B/A^\circ}\text{i}$$

c. $$\frac{A^\bullet \quad B^\bullet}{A{\cdot}B^\bullet}\text{i} \qquad \frac{B^\circ \quad A^\circ}{A{\cdot}B^\circ}\text{ii}$$

Without polarities, a formula tree is a kind of construction tree of the formula at its root: daughters are labelled with the immediate subformulas of their mothers. The polarities indicate sequent sidedness, input for antecedent and output for succedent; the polarity propagation follows the sidedness of subformulas in the sequent rules: in the antecedent (input) rule for $A\backslash B$ the subformula A goes in a succedent (output) and the subformula B goes in an antecedent (input), in the succedent (output) rule for $A\backslash B$ the subformula A goes in an antecedent (input) and the subformula B goes in a succedent (output), etc. The labels i and ii indicate whether the corresponding sequent rule is unary or binary. Note that in the output links the order of the subformulas is switched; this corresponds to a cyclic reading of sequents: the succedent type is adjacent to the first antecedent type.

A *proof frame* is a finite sequence of polar category formula trees exactly one of which has a root of output polarity (corresponding to the unique succedent of sequents).

An *axiom linking* on a set of literal labelled leaves is a partitioning of the set into pairs of complementary leaves which is planar in their ordering, i.e. there are no two pairs $\{L_1, L_3\}, \{L_2, L_4\}$ such that $L_1 < L_2 < L_3 < L_4$. Geometrically, planarity means that where the leaves are ordered on a line, paired leaves can be connected in the half plane without crossing. Axiom links correspond to id instances in a sequent proof.

A *proof structure* is a proof frame together with an axiom linking on its leaves. A *proof net* is a proof structure in which every elementary (i.e. visiting vertices at most once) cycle crosses the edges of some i-link.[2] Geometrically, an elementary cycle is the perimeter of a face or cluster of faces in a planar proof structure. There is a proof net with roots $A^\circ, A_1{}^\bullet, \ldots, A_n{}^\bullet$ iff $A_1, \ldots, A_n \Rightarrow A$ is a valid sequent.

2 This criterion, adapted from that of Lecomte and Retoré (1995), derives from Girard's (1987) 'long trip condition', which is a highly involved mathematical result. Danos and Regnier (1989) express it in terms of acyclicity and connectivity of certain subgraphs. Intuitively, acyclicity assures that the subformulas of ii-links (binary rules) occur in *different* subproofs, whereas connectivity assures that the subformulas of i-links (unary rules) occur in the *same* subproofs (attributed to Jean Gallier by Philippe de Groote, p.c.). However the single-succedent (intuitionistic) nature of Cut-free categorial proofs in fact renders the connectivity requirement redundant, hence we have just an acyclicity test.

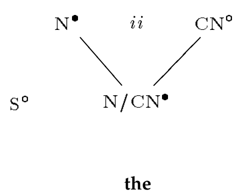
Let us assume the following lexical assignments:

(17)

	barn	–	*barn*
		:=	CN
	horse	–	*horse*
		:=	CN
	past	–	$\lambda x \lambda y \lambda z(past\ x\ (y\ z))$
		:=	$((N\backslash St)\backslash(N\backslash St))/N$
	raced	–	*race*
		:=	$N\backslash S+$
	the	–	*the*
		:=	N/CN

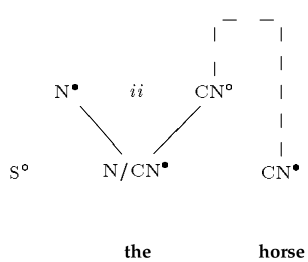
The feature + on S marks the projection of a tensed verb form; a verb phrase modified by "past" need not be tensed. Let us consider the incremental processing of (2a) as proof net construction. We assume initially that an S is expected; after perception of the word "the" there is the following partial proof net (for simplicity we omit features, included in lexical entries, from proof nets):

(18)



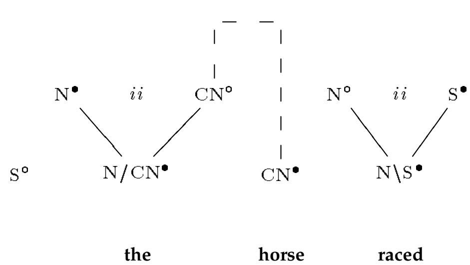
Here there are three unmatched valencies/unresolved dependencies; no axiom links can yet be placed, but after "horse" we can build:

(19)



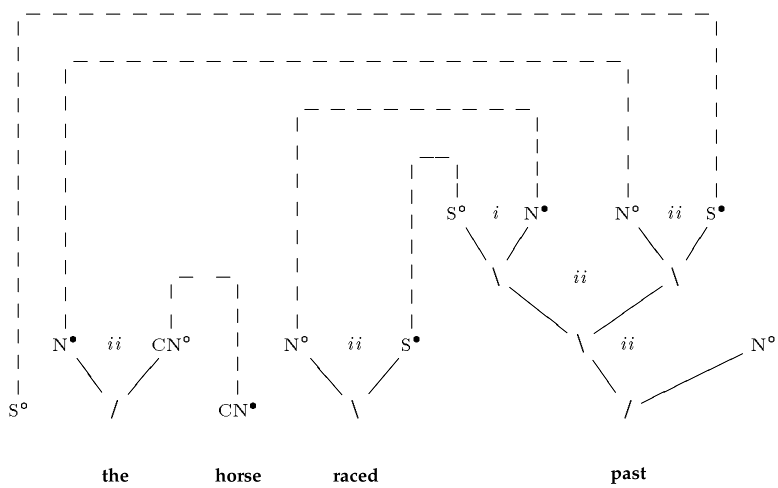
Now there are only two unmatched valencies. After "raced" we have, on the correct analysis, the following:

(20)



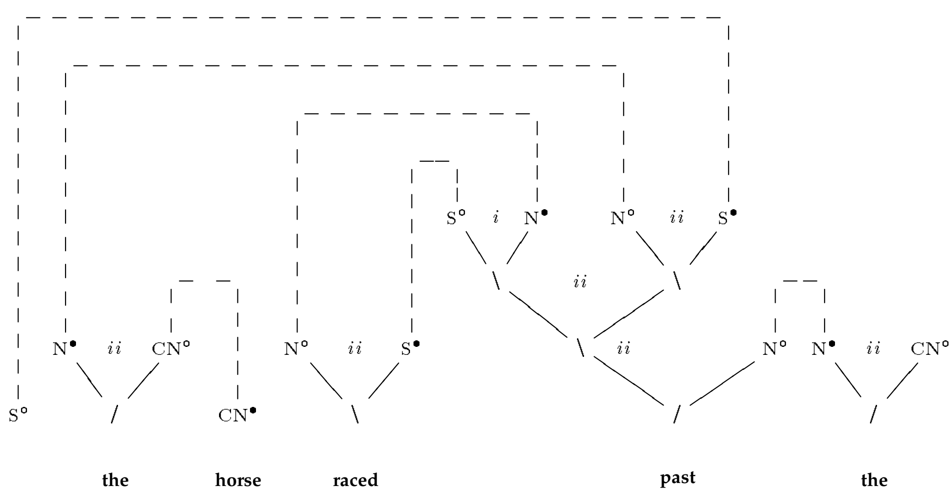
Note that linking the Ns is possible, but we are interested in the history of the *correct* analysis, and in that the verb valencies are matched by the adverb that follows (henceforth we indicate only the principal connective of a mother node):

(21)



Observe that a cycle is created, but as required it crosses the edges of a i-link. At the penultimate step we have:

(22)



The final proof net analysis is given in figure 1.

The semantics associated with a categorial proof net, i.e. the proof as a lambda term (intuitionistic natural deduction proof, under the Curry-Howard correspondence) is extracted by associating a distinct index with each output division node and travelling as

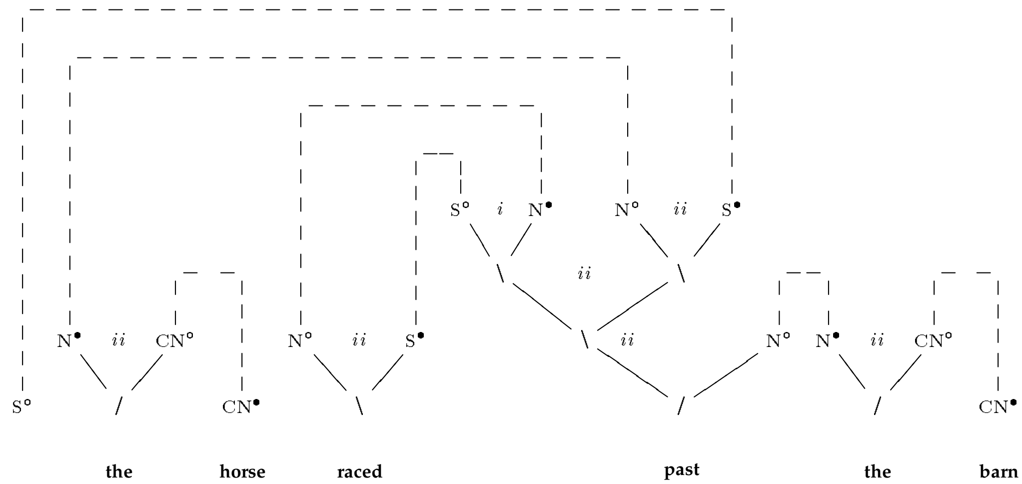


Figure 1
"the horse raced past the barn"

follows, starting by going up at the unique output root (de Groote and Retoré 1996):

(23)

- travelling up at the mother of an output division link, perform the lambda abstraction with respect to the associated index of the result of travelling up at the daughter of output polarity;
- travelling up at the mother of an output product link, form the ordered pair of the result of travelling up at the right daughter (first component) and the left daughter (second component);
- travelling up at one end of an axiom link, continue down at the other end;
- travelling down at an (input) daughter of an input division link, perform the functional application of the result of travelling down at the mother to the result of travelling up at the other (output) daughter;
- travelling down at the left (resp. right) daughter of an input product link, take the first (resp. second) projection of the result of travelling down at the mother;
- travelling down at the (input) daughter of an output division link, return the associated index;
- travelling down at a root, return the associated lexical semantics.

Thus for our example we obtain (24a), which is logically equivalent to (24b).

(24) a. $(\lambda x \lambda y \lambda z(\textit{past}\ x\ (y\ z))\ (\textit{the barn})\ \lambda 1(\textit{race}\ 1)\ (\textit{the horse}))$
b. $(\textit{past}\ (\textit{the barn})\ (\textit{race}\ (\textit{the horse})))$

The analysis of (2b) is less straightforward. Whereas in (2a) "raced" expresses a one-place predication ("go quickly"), in (2b) it expresses a two-place predication (there was some agent racing the horse); "horse" is modified by an agentless passive participle, but the adverbial "past the barn" is modifying "race". Within the confines of the Lambek calculus the characterization we offer assumes the lexical assignment to the passive participle given in the following.[3]

(25)
fell – *fall*
:= N\S+
raced – $(\lambda x \lambda y \lambda z[(y\ z) \wedge \exists w(x\ z\ w)], \textit{race2})$
:= ((CN\CN)/(N\(N\S−)))·(N\(N\S−))

Here "raced" is classified as the product of an untensed transitive verbal type, which can be modified by the adverbial "past the barn" by composition, and an adnominalizer of this transitive verbal type. According to this, (2b) has the proof net analysis given in figure 2. The semantics extracted is (26a), equivalent to (26b)

(26) a. $(\textit{fall}\ (\textit{the}\ (\pi_1(\lambda x \lambda y \lambda z[(y\ z) \wedge \exists w(x\ z\ w)], \textit{race2})\ \lambda 29 \lambda 30$
$(\lambda u \lambda v \lambda w(\textit{past}\ u\ (v\ w))\ (\textit{the barn})\ \lambda 41(($
$(\pi_2(\lambda p \lambda s \lambda t[(s\ t) \wedge \exists q(p\ t\ q)], \textit{race2})\ 29\ 41)\ 30)\ \textit{horse})))$
b. $(\textit{fall}\ (\textit{the}\ \lambda 8[(\textit{horse}\ 8) \wedge \exists 7(\textit{past}\ (\textit{the barn})\ (\textit{race2}\ 8\ 7))]))$

Let us assign to each proof net analysis a complexity profile which indicates, before and after each word, the number of unmatched literals, i.e. unresolved valencies or dependencies, following the processing up to that point. This is a measure of the course of memory load in optimal incremental processing. We are not concerned here with resolution of lexical ambiguity or serial backtracking: we are supposing sufficient resources that the non-determinism of selection of lexical entries and their parallel consideration is not the critical burden. Rather, the question is: which among parallel competing analyses places the least load on memory?

The complexity profile is easily read off a completed proof net: the complexity inbetween two words is the number of axiom links bridging rightwards (forwards in time) at that point. Thus for (2a) and (2b) analysed in figures 1 and 2 the complexity profiles are as follows:

(27)

6				b				
5								
4			b	a				
3		ab			b	b		
2			a				b	
1	ab				a	a		
0							a	b
a.	the	horse	raced	past	the	barn		
b.	the	horse	raced	past	the	barn	fell	

We see that after the first words the complexity of the correct analysis of (2b) is con-

3 In general grammar requires the expressivity of more powerful categorial logics than just Lambek calculus; however, so far as we are aware, the characterizations we offer within the Lambek calculus bear the same properties with regard to our processing considerations as their more sophisticated categorial logic refinements, because the latter concern principally generalisations of word order, whereas the semantic dependencies on which our complexity metric depends remain the same.

sistently higher than that of its 'garden path' (2a), just as would be expected on the assumption that in (2b) the less costly but incorrect analysis is assailant.

Let us consider now quantifier scope preference. A rudimentary account of sentence-peripheral quantifier phrase scoping is obtained in Lambek calculus by means of lexical assignments such as the following:

(28)

everyone	–	$\lambda x \forall y(x\ y)$
	:=	St/(N\S$t$)
everyone	–	$\lambda x \forall y(x\ y)$
	:=	(St/N)\S$t$
someone	–	$\lambda x \exists y(x\ y)$
	:=	St/(N\S$t$)
someone	–	$\lambda x \exists y(x\ y)$
	:=	(St/N)\S$t$

Then one analysis of (5a) is that given in figure 3. This is the subject wide scope analysis: its extracted and simplified semantics is as in (29).

(29) a. $(\lambda x \exists y(x\ y)\ \lambda 1(\lambda x \forall y(x\ y)\ \lambda 2(\mathit{love}\ 2\ 1)))$
b. $\exists x \forall y(\mathit{love}\ y\ x)$

A second analysis is that given in figure 4. This is the object wide scope analysis: its extracted and simplified semantics is as in (30).

(30) a. $(\lambda x \forall y(x\ y)\ \lambda 2(\lambda x \exists y(x\ y)\ \lambda 1(\mathit{love}\ 2\ 1)))$
b. $\forall y \exists x(\mathit{love}\ y\ x)$

Let us compare the complexity profiles for the two readings:

(31)

4		b		
3			ab	
2		a		
1	ab			
0				ab
a. b.	someone	loves	everyone	∃∀ (subject wide scope, figure 3) ∀∃ (object wide scope, figure 4)

At the only point of difference the subject wide scope reading, the preferred reading, has the lower complexity.

For the passive (5b) let there be assignments as in (32). The preposition "by" projects an agentive adverbial phrase; "is" is a functor over (post-)nominal modifiers ("the man outside", "John is outside", etc.) and passive "loved" is treated exactly like passive "raced" in (25).

(32)

by	–	$\lambda x \lambda y \lambda z[[z = x] \wedge (y\ z)]$
	:=	((N\S−)\(N\S−))/N
is	–	$\lambda x \lambda y(x\ \lambda z[z = y]\ y)$
	:=	(N\S+)/(CN\CN)
loved	–	$(\lambda x \lambda y \lambda z[(y\ z) \wedge \exists w(x\ z\ w)], \mathit{love})$
	:=	((CN\CN)/(N\(N\S−)))·(N\(N\S−))

A ∀∃ analysis of (5b) is given in figure 5. This has semantics, after some simplification, (33), which is equivalent to (30).

(33) $\forall 16 \exists 9 \exists 7[[9 = 7] \wedge (\mathit{love}\ 16\ 9)]$

An $\exists\forall$ analysis of (5b) is given in figure 6. This has semantics, after some simplification, (34), which is equivalent to (29).

(34) $\exists 16 \forall 14 \exists 7[[7 = 16] \wedge (\mathit{love}\ 14\ 7)]$

Again, the preferred reading has the lower complexity profile:

(35)

6				b		
5						
4		b	b	a		
3					ab	
2		a	a			
1	ab					
0						ab
a. b.	everyone	is	loved	by	someone	$\forall\exists$ (figure 5) $\exists\forall$ (figure 6)

Turning now to the final performance phenomenon listed in the introduction, for subject and object relativisation we assume the relative pronoun lexical assignments (36).

(36) **that** – $\lambda x \lambda y \lambda z[(y\ z) \wedge (x\ z)]$
:= (CN\CN)/(N\S+)
that – $\lambda x \lambda y \lambda z[(y\ z) \wedge (x\ z)]$
:= (CN\CN)/(S+/N)

Sentence (8b) is analysed in figure 7. Sentence (9b) is analysed in figure 8. Let us compare the complexities:

(37)

9								b				
8									b			
7						b	b					
6					b							
5										b		
4			ab	ab			a	a				
3		ab			a	a			a	a		
2											ab	
1	ab											
0												ab
a.	the	dog	that	chased	the	cat	that	saw	the	rat	barked	
b.	the	cheese	that	the	rat	that	the	cat	saw	ate	stank	

Again, the profile of (9b) is higher; indeed it rises above 7–8, thus reaching what are usually taken to be the limits of short term memory.

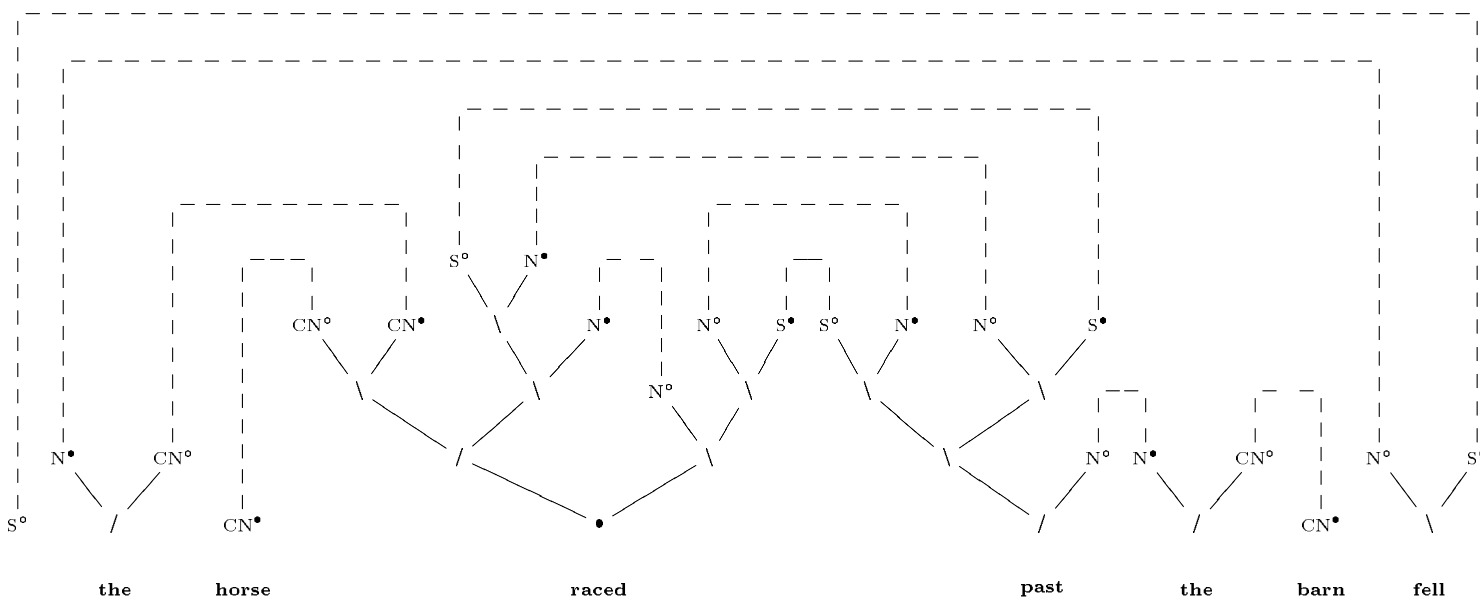


Figure 2
"the horse raced past the barn fell"

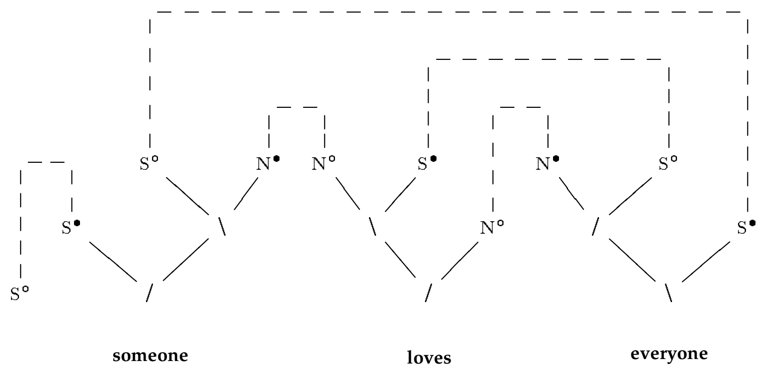


Figure 3
"someone loves everyone" (∃∀)

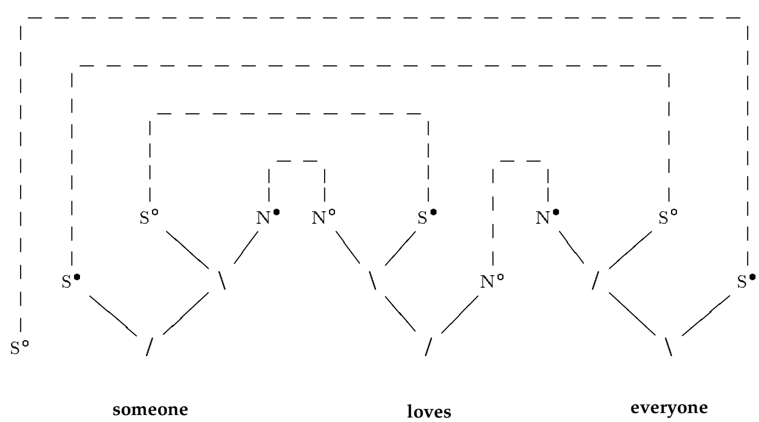


Figure 4
"someone loves everyone" (∀∃)

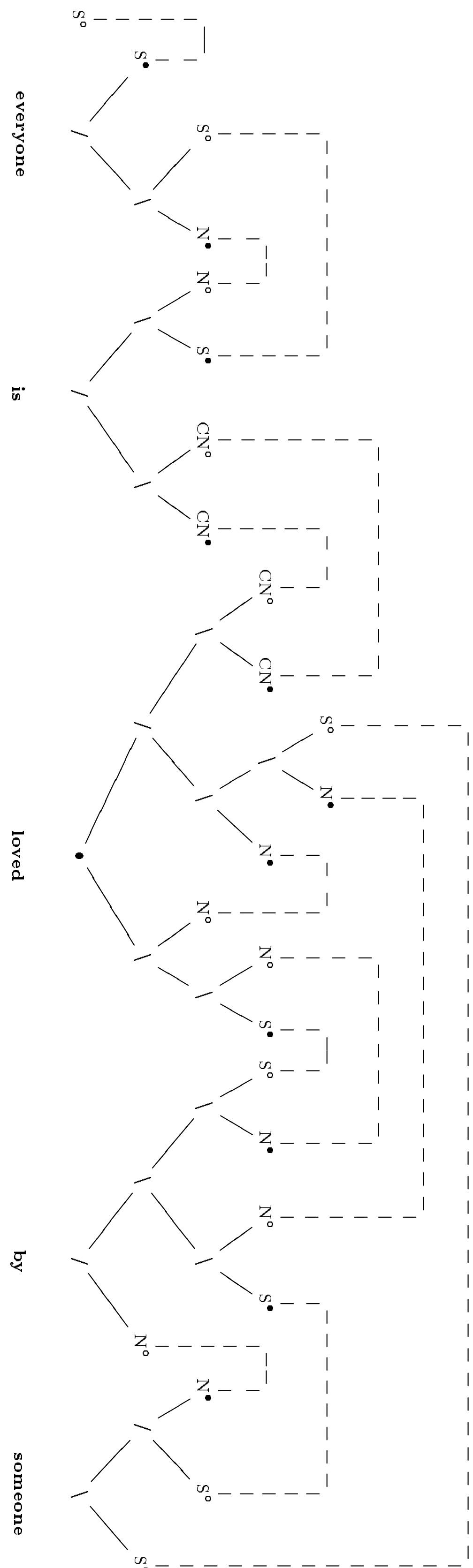


Figure 5
"everyone is loved by someone" (∀∃)

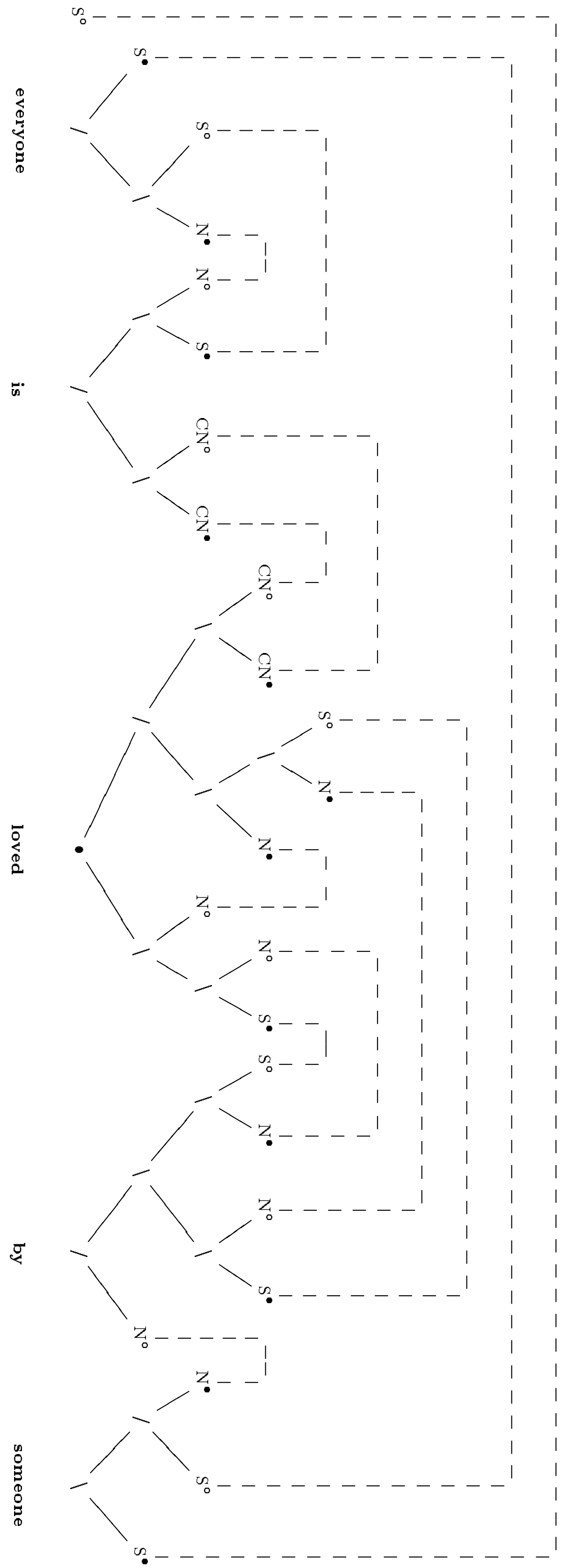


Figure 6
"everyone is loved by someone" (∃∀)

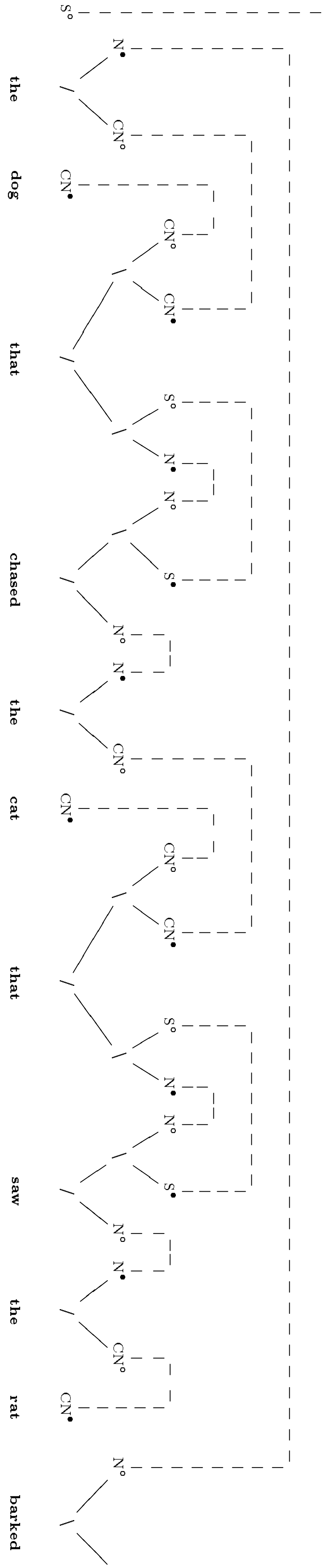


Figure 7
"the dog that chased the cat that saw the rat barked"

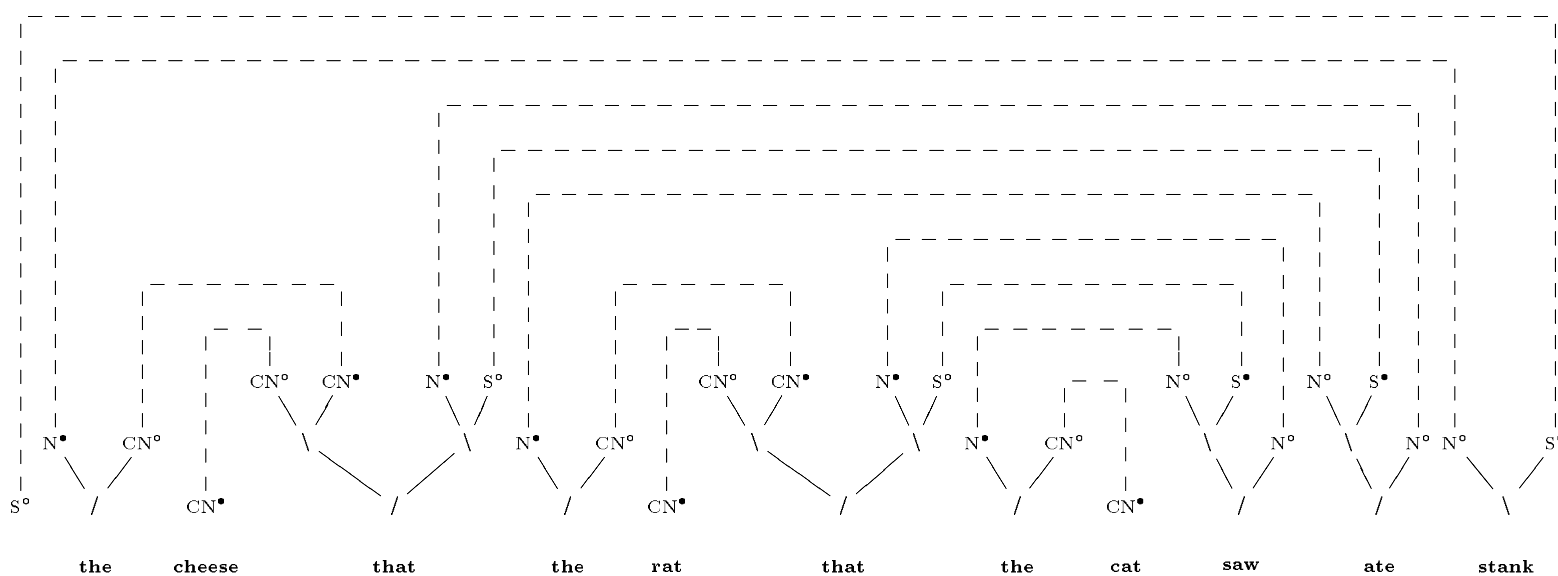


Figure 8
"the cheese that the rat that the cat saw ate stank"

3 Further cases

Another dramatic example of unacceptability is provided by the following:

(38)
a. That two plus two equals four surprised Jack.
b. ?That that two plus two equals four surprised Jack astonished Ingrid.
c. ??That that that two plus two equals four surprised Jack astonished Ingrid bothered Frank.

The passive parafrases, however, seem more or less equally acceptable:

(39)
a. Jack was surprised that two plus two equals four.
b. Ingrid was astonished that Jack was surprised that two plus two equals four.
c. Frank was bothered that Ingrid was astonished that Jack was surprised that two plus two equals four.

In figure 9 we give the analysis of (38b) and in figure 10 that of (39b). We now abbreviate proof nets by flattening formula trees into their linear representations; since this conceals the order switching of output links the notation belies the underlying planarity. It is very interesting to observe that the complexity profile of the latter is in general lower even though the analysis has more than twice the total number of links.

(40)

7					a									
6				a		a								
5			a				a							
4				b				ab						
3		a							a		b			
2		b	b			b	b			ab		b		
1	ab				b				b		a		b	
0												a		b
a.	that	that	two	plus	two	equals	four	surp.	jack	ast.	Ingrid			
b.	Ingrid	was	ast.	that	Jack	was	surp.	that	two	plus	two	equals	four	

By way of another example, Kimball (1973, p.27) observes that in a sentence such as (41), three ways ambiguous according to the attachment of the adverb, the lower the attachment is, the higher the preference (what he terms ‘Right Association’).

(41) Joe said that Martha believed that Ingrid fell today.

In figure 11 we give the analyses for the highest, the middle, and the lowest attachments. Accordingly, the complexity profiles are:

(42)

6					c			bc		
5			c	c		bc	bc			
4									abc	
3										
2		abc			ab			a		
1	abc		ab	ab		a	a			
0										abc
a.										lowest attachment
b.	Joe	said	that	Martha	believed	that	Ingrid	fell	today	middle attachment
c.										highest attachment

The same effect occurs strongly in (43), where the preferred reading is the one given by the lowest attachment, even though that one is the nonsensical reading.

(43) the book that shocked Mary’s title

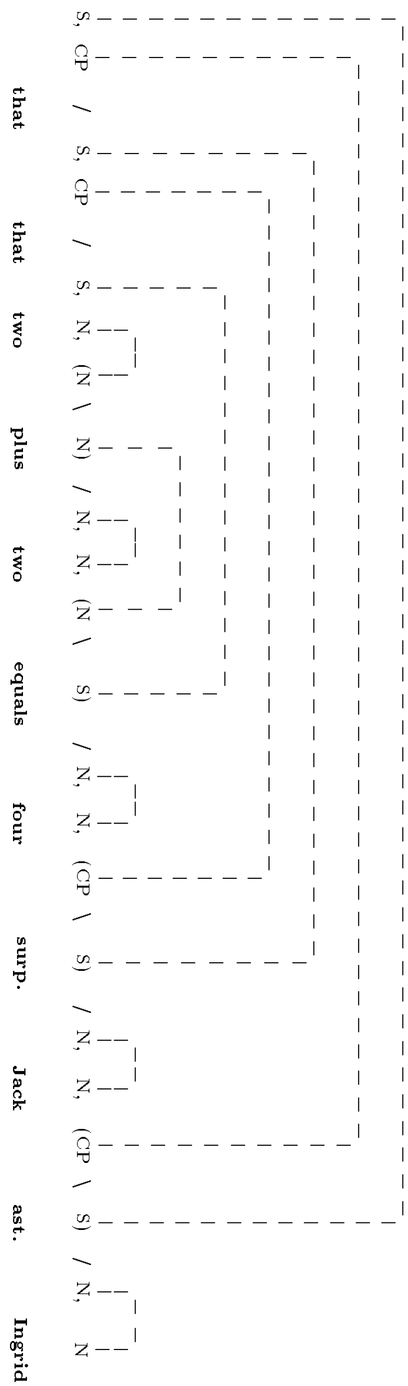


Figure 9
"that that two plus two equals four surprised Jack astonished Ingrid"

The analyses are given in figure 12. The complexities are thus:

(44)

4			a	a				
3		a			a			
2			b	b		ab		
1	ab	b			b		ab	
0								ab
a. b.	the	book	that	shocked	Mary	’s	title	sensical nonsensical

Finally, our account appears to explain the preference for heavy noun phrases to appear at the end of the verb phrase (heavy noun phrase shift). Of the following the second is more acceptable:

(45) a. ?John gave the painting that Mary hated to Bill.
b. John gave Bill the painting that Mary hated.

The analyses are given in figure 13. The complexities are thus:

(46)

4							a			
3					a	a		b		
2		ab	ab	a		b	b			
1	ab			b	b			a	a	
0									b	a
a.	John	gave	the	painting	that	Mary	hated	to	Bill	
b.	John	gave	Bill	the	painting	that	Mary	hated		

References

Ades, Anthony E. and Mark J. Steedman: 1982, ‘On the Order of Words’, *Linguistics and Philosophy* **4**, 517–558.

Bever, T.: 1970, ‘The cognitive basis for linguistic structures’, in J. R. Hayes (ed.) *Cognition and the Growth of Language*, Wiley, New York.

Buszkowski, Wojciech: 1986, ‘Completeness results for Lambek syntactic calculus’, *Zeitschrift für mathematische Logik und Grundlage der Mathematik* **32**, 13–28.

Carpenter, Bob: 1998, *Type-Logical Semantics*, MIT Press, Cambridge, Massachusetts.

Chomsky, Noam: 1965, *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Massachusetts.

Danos, V. and L. Regnier: 1989, ‘The structure of multiplicatives’, *Archive for Mathematical Logic* **28**, 181–203.

Došen, K.: 1992, ‘A brief survey of frames for the Lambek calculus’, *Zeitschrift für mathematische Logik und Grundlage der Mathematik* **38**, 179–187.

de Groote, Philippe and Christian Retoré: 1996, ‘On the Semantic Readings of Proof-Nets’, in G.-J. Kruijff, G. Morrill and R. T. Oehrle (eds.) *Proceedings of formal Grammar 1996*, Prague, 57–70.

Girard, Jean-Yves: 1987, ‘Linear Logic’, *Theoretical Computer Science* **50**, 1–102.

Hendriks, Herman: 1993, *Studied Flexibility: Categories and Types in Syntax and Semantics*, Ph.D. dissertation, Universiteit van Amsterdam.

Hepple, Mark: 1990, ‘Normal form theorem proving for the Lambek calculus’, in H. Karlgren (ed.), *Proceedings of COLING 1990*, Stockholm.

Hepple, Mark and Glyn Morrill: 1989, 'Parsing and Derivational Equivalence', in *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*, Manchester.

Kimball, John: 1973, 'Seven principles of surface structure parsing in natural language', *Cognition* **2**, 15–47.

König, E.: 1989, 'Parsing as natural deduction', in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Vancouver.

Lambek, J.: 1958, 'The mathematics of sentence structure', *American Mathematical Monthly* **65**, 154–170, also in Buszkowski, W., W. Marciszewski, and J. van Benthem (eds.): 1988, *Categorial Grammar*, Linguistic & Literary Studies in Eastern Europe Volume 25, John Benjamins, Amsterdam, 153–172.

Lecomte, Alain and Christian Retoré: 1995, 'An alternative categorial grammar', in G. Morrill and R. T. Oehrle (eds.) *Proceedings of formal Grammar 1995*, Barcelona.

Moortgat, Michael: 1988, *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*, Foris, Dordrecht.

Moortgat, Michael: 1997, 'Categorial Type Logics', in J. van Benthem and A. ter Meulen (eds.) *Handbook of Logic and Language*, Elsevier, Amsterdam, 93–177.

Morrill, Glyn: 1994, *Type Logical Grammar: Categorial Logic of Signs*, Kluwer Academic Publishers, Dordrecht.

Pentus, M.: 1994, 'Language completeness of the Lambek calculus', *Proceedings of the Eighth Annual IEEE Symposium on Logic in Computer Science*, Paris.

Roorda, Dirk: 1991, *Resource Logics: proof-theoretical investigations*, Ph.D. dissertation, Universiteit van Amsterdam.

Steedman, Mark J.: 1997, *Surface Structure and Interpretation*, MIT Press, Cambridge, Massachusetts.

Zielonka, W.: 1981, 'Axiomatizability of Ajdukiewicz-Lambek calculus by means of cancellation schemes', *Zeitschrift für mathematische Logik und Grundlage der Mathematik* **27**, 215–224.

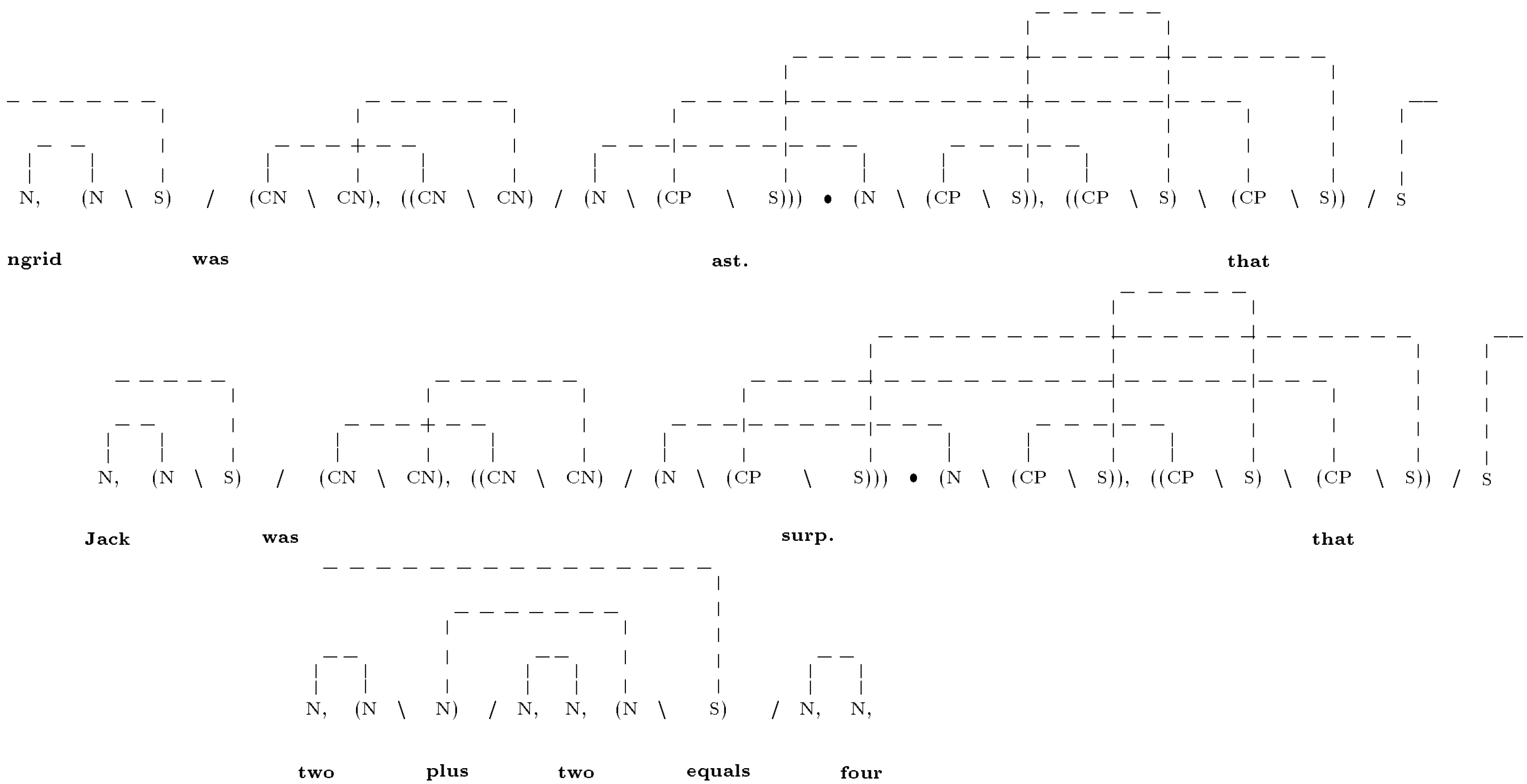


Figure 10
"Ingrid was astonished that Jack was surprised that two plus two equals four"

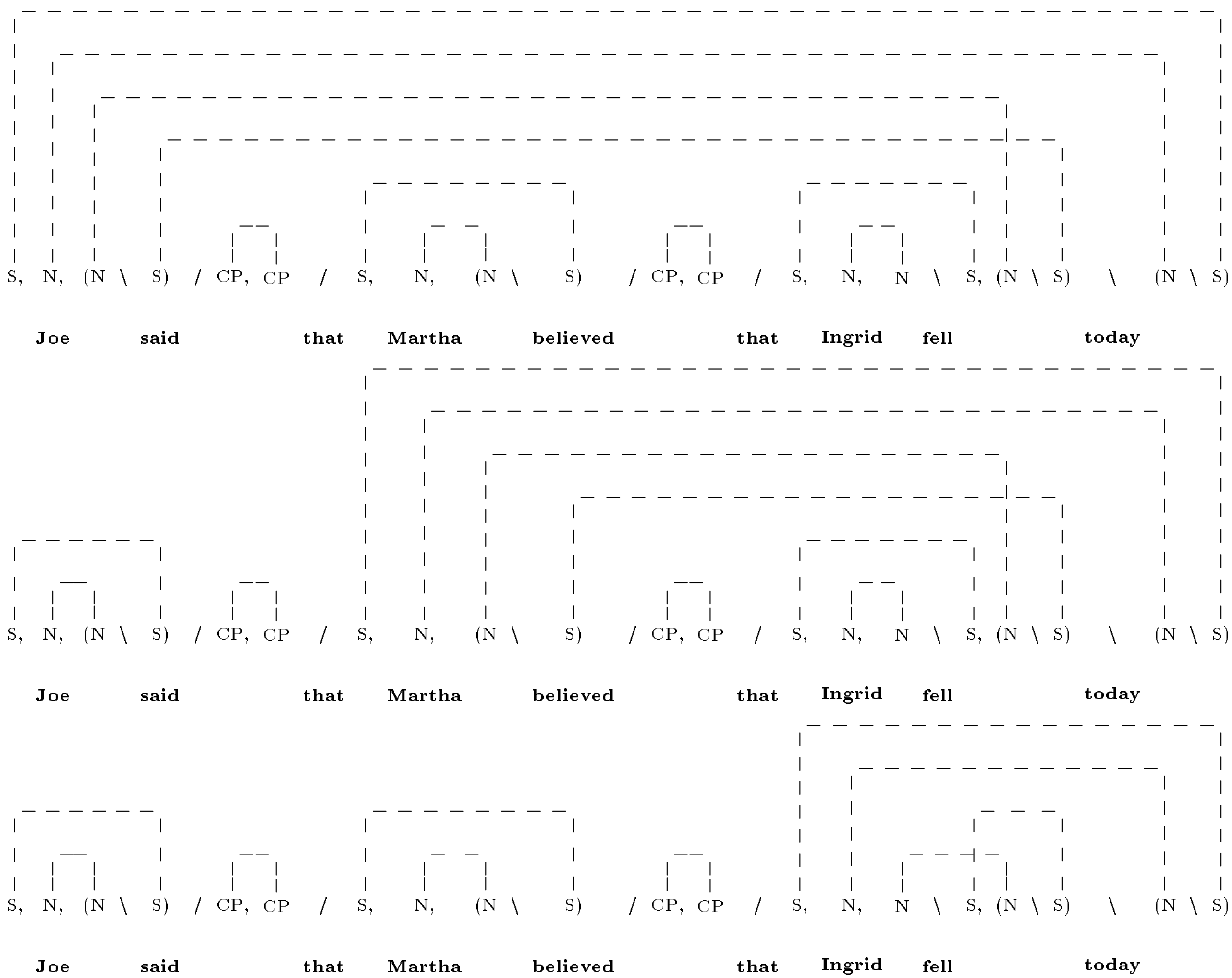


Figure 11
"Joe said that Martha believed that Ingrid fell today"

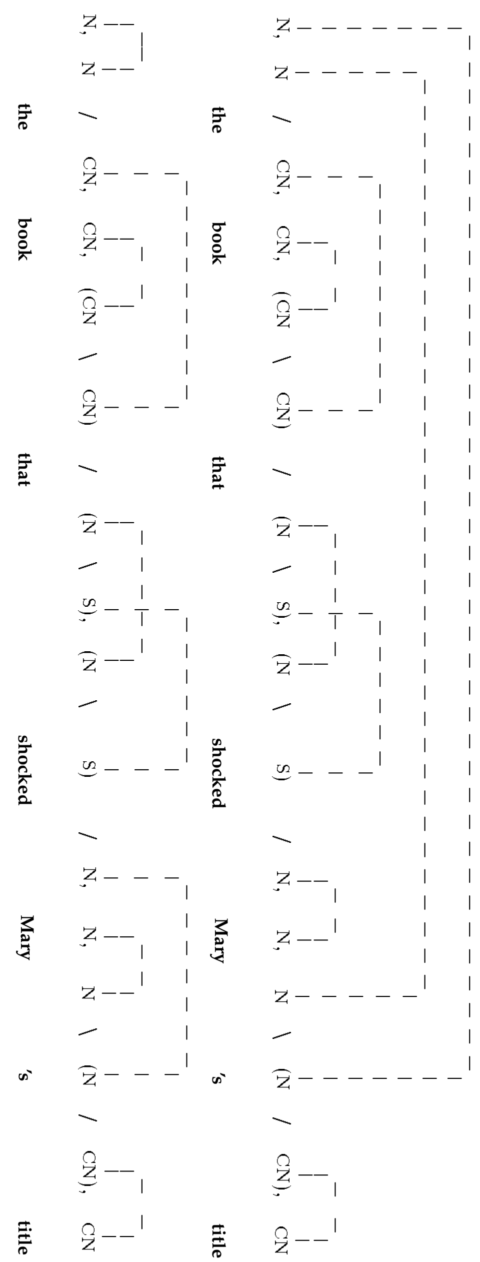


Figure 12
"the book that shocked Mary's title" (sensical and nonsensical)

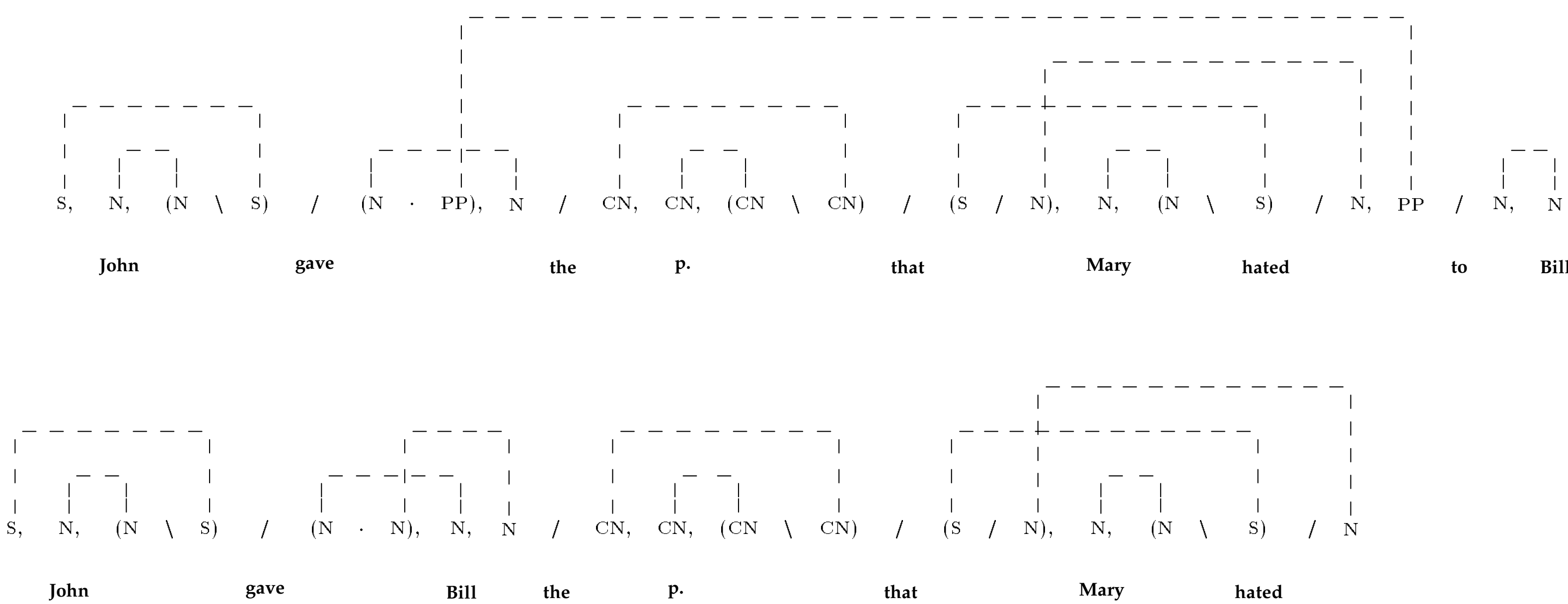


Figure 13
"John gave the painting that Mary hated to Bill" vs. "John gave Bill the painting that Mary hated"