From word to sentence:

a computational algebraic approach to grammar.

Joachim Lambek
McGill University,
Montreal, Canada.

Table of contents

0. Preface.

In 1957, Noam Chomsky published his pioneering booklet "Syntactic Structures", which offered a promising computational approach to grammar, based on what computer scientists call "rewrite rules" and what mathematicians describe as a "semi-Thue system", essentially a finitely presented semigroup. As a mathematical approach to linguistics, it greatly outperformed the categorial approach of Ajdukiewicz, Bar-Hillel and the present author, which in its polished form turned out to be based on so-called "residuated semigroups".

Chomsky refined and deepened his theory in subsequent years, publishing almost one book per annum. Yet, while going profoundly into the structure of natural languages, it became less accessible to computational implementation, at least in the opinion of some categorial linguists. Prompted by a suggestion by Claudia Casadio, I proposed a new computational algebraic approach in 1998, founded in the categorial tradition and based on a slight generalization of the ubiquitous concept of a group, called a "pregroup". In retrospect, this approach was already implicit in the work of Zellig Harris, Chomsky's teacher, and may be viewed as a refinement of the latter. It has now been tested on small fragments of several circum-Mediterranean and far Eastern languages. In the present monograph, I hope to apply the new approach to a more substantial portion of at least one language, English, the current lingua franca. To make it as widely accessible as possible, mathematical details have been suppressed, at least at the beginning, and may be skipped by the reader if she so wishes, and only a nodding acquaintance with modern linguistics is presupposed.

Whereas modern linguists claim to deal with the spoken language, the present enterprise follows more closely in the footsteps of traditional grammarians, who are mostly concerned with what has been or should be written. Grammatical sentences are idealized entities, which approximate spoken utterances, but they may serve as models for the latter when we wish to investigate natural language processing.

I. Introduction

1. Origin of languages.

While other animals have ways of communicating, language as we know it is specifically human. In fact, the language faculty may be the defining property of what makes us human. It gives us a competitive advantage over our relatives in the animal world. It allows useful knowledge to be passed from one generation to the next: What berries can be eaten safely? How does one make fire? Unfortunately, carried along with such useful knowledge there are also false beliefs: Which spirit is responsible for rain and how can he be bribed to end the drought?

As far as we know, spoken language was invented only once, at some time between 100,000 and 50,000 years ago, before humans moved out of Africa and spread all over the globe. It is believed by some linguists that all languages spoken today are descended from one protolanguage. Although this is contested by many other linguists, it is an opinion I share. What evidence is there for such a theory?

It has been known for over 200 years that there is a strong relationship between many languages spoken today and in classical times in a large corridor extending from Iceland to Ceylon. The similarity between the languages of this so-called Indo-European family is evident even to the casual observer, for example when looking at kinship terminology. Compare English *father, mother, son* and *daughter* with Sanskrit *pitr*, *mātr*, *sūnu* and *duhitr*. Of course, not every member of the Indo-European family need have preserved all these words. For example, Latin has substituted *filius* and *filia* for *son* and *daughter*.

More recently, Greenberg observed that Indo-European belongs to a bigger Eurasiatic superfamily, which also includes the Ural-Altaic languages and may stretch as far east as Japan. It is not so easy to find recognizably related words belonging to members of this superfamily, except for later borrowing. But there is other evidence based on morphemes. For example, all first person verb forms in Turkish end in *m*, and the same ending survives in some Indo-European languages. In Polish it is quite ubiquitous, in Latin it has been retained mostly for the past tense and in English only in the single word *am*.

As a rough guess, the Indo-European languages may have a common origin about 6,000 years ago and the Eurasiatic languages perhaps 20,000 years ago, and so it may no longer be easy to trace related words to a common origin. One problem is that words combine sound and meaning and both may change over time.

While sometimes both sound and meaning are still recognizably related after a long period of time, such as Sanskrit *duhitr* and English *daughter*, more often the meaning will change: for example, Hindi *dohiti* means *granddaughter*, though with a still recognizable related meaning. Sometimes the meaning is transformed beyond recognition. A striking example that comes to mind is the English word *lumberjack*. The first part *lumber* may be recognized in *long beard*, but the meaning was transformed in a relatively short time. The Germanic tribe of the long bearded *Langobardi* settled in Italy as *Lombards*. Lombards became bankers in Renaissance Europe; but loan institutions often degenerated into pawnshops, which ultimately also sold second-hand furniture. Discarded furniture was stored in the *lumber* room, which later also contained other pieces of wood. The second part *jack* is derived from the biblical name *Jacob*, although nowadays *Jack* is used as an alternative for *John*. The bible explains *Jacob* as "he who follows on the heel" in Hebrew, to make clear that Jacob was not the first-born of twins.

Often the meaning is preserved, but the sound is changed so completely that a common origin is no longer recognizable at first glace. For example, take English *five* and French *cinq*. We would not believe that these words are related, were it not known that English *five* corresponds to German *fünf* and that French *cinq* is derived from Latin *quinque*. Finally, both *fünf* and *quinque* are assumed to have evolved from a conjectured proto-Indo-European **pengue*.

When both sound and meaning have diverged too much, we may no longer recognize a common origin. One example that comes to mind is based on English *head* and *cattle*, the latter derived from Latin *caput*, and the former related to German *Haupt*, both meaning *head*. People will speak of "head of cattle", without realizing the tautologous repetition. (Incidentally, some Indonesian languages count cattle by their tails.)

The Afro-Asiatic superfamily of languages, also established by Greenberg, extends the previously recognized Hamito-Semitic family to include many African languages. It extends from the Atlantic ocean to the Tigris river. One may ask how closely it is related to Eurasiatic. At first sight, one is struck by the analogy between English *six*/*seven* and Hebrew *shesh*/$sheb^{h}a$. But it is now believed that this correspondence may be explained by an early loan from Semitic to Indo-European, say about 6,000 years ago. As far as I know, the two super-families are no more closely related than can be explained by the common origin of all languages [Ruhlen].

Let me add an observation of my own. In Indo-European languages there is a parallel development between words for *two* and *tooth*, as evidenced by the proportion

$$two/tooth \;=\; zwei/Zahn \;=\; deux/dent......$$

Surprisingly, this proportion can be extended to Hamito-Semitic. For example, Hebrew *shnayim* can mean "two" or a "pair of teeth". Can it be extended even further?

A conjectured proto-language [Bengtson and Ruhlen] with a surprising number of recognizable descendent words in modern English (*head, hound, man, milk, mind, queen, who*,......) also contains the reconstructed form **pal* for *two*. Although it does not list a word for *tooth*, such a word occurs in the rival proto-language Nostratic, namely $^{*}p^{h}al$. Of course this may just be a coincidence; but my point is that it may be easier to recognize relations between pairs of words than resemblances between the words themselves.

2. Grammar and mathematics.

Let us take a brief look at how language is processed in discourse. The speaker converts his intended meaning by a mysterious (to me) process into a sequence of words, inflected forms of lexical items in the mental dictionary, which is then translated into a sequence of sounds. The hearer is confronted with a continuous stream of sounds, which must be broken up into discrete phonemes, strings of which are collected into words. These are analyzed as inflected forms of lexical items, strings of which are analyzed hopefully as sentences. Sentences are translated into meaningful messages, which may give rise to a belief or to a behavioral response by the hearer. There may be feedback between any part of this process and any other part.

Of course, there is no guarantee that a complete utterance is a grammatical sentence. This is an idealized concept, originally confined to the written language. In fact, the word *grammar* is derived from a Greek word meaning "writing". Many people, with a memory of their school years, regard grammar as a tedious enterprise. Yet, at one time, the knowledge of grammar (meaning Latin grammar) was supposed to endow one with magical power, the power to enchant, hence our word *glamour*.

In this book, I will confine attention to a narrow part of the language process, the stringing together of words to form sentences. Aside from my lack of competence in phonology, this is because the mathematical techniques I feel confident in handling do not readily apply to the interface between syntax (wedded to morphology) and semantics (including pragmatics). Although it is only syntax and morphology that I treat formally using an algebraic approach, I will make occasional informal excursions into semantics and even into etymology, which deals with the origin of words. However, I shy away from phonology.

It should not come as a surprise that language processing involves computation. This may be obvious to all those (not so the present author) who use computers for word processing. But the connection between linguistics and mathematics is much older and embedded in folk memory. For example, the English words *talk, tell* and *tale* are all cognate with the German word *Zahl*, which means "number".

The ancients did not officially include grammar into what they considered to be mathematics. According to the Pythagorean school, mathematics was supposed to consist of Arithmetic, Geometry, Astronomy and Music. Much later, at the court of Theodoric in Ravenna, Boethius declared the same four subjects, collectively known as the "quadrivium", to be the core of a liberal education. Much later still, at medieval universities, it was realized that a more elementary "trivium" was required as a prerequisite, namely Logic, Grammar and Rhetoric. Logic was officially welcomed into the bosom of mathematics in the nineteenth century; think of Boole, Schroeder, Peirce and Frege, but Grammar only in the second half of the twentieth century.

Over the centuries, a number of mathematicians took a professional interest in language studies. Eratosthenes, known for his sieve of prime numbers and for being the first to calculate the circumference of the earth, and Wallis, an early contributor to the calculus, are both said to have written books on grammar. Grassmann is perhaps even better known for his work in philology (Grassmann's law) than for his pioneering ideas in mathematics. My own experience tells me that, in countries where high school students are faced with a choice between a scientific program and a humanistic one based on a study of classical languages, future mathematicians often arise from the latter stream.

3. An algebraic approach to grammar.

The present mathematical approach to grammar belongs to the tradition of *categorial grammar*. Here *category* means "type" and must be distinguished from the notion of a category introduced in 1945 by Eilenberg and Mac Lane into mathematics (although this notion too has been invoked in linguistics when taking grammatical derivatives seriously and interpreting them as morphisms of a category). The principal idea is to assign to each word in the mental dictionary one or more types. (Sometimes the types of inflected forms of lexical items may be calculated by certain "metarules".) These types are elements of an algebraic system or terms of a logical system and the grammatical status of a string of words is determined by a calculation on the corresponding types. Here are two examples:

$$\begin{array}{c} \textit{she will come} \ , \\ \pi_3 \ (\pi^r \mathbf{s}_1 \mathbf{j}^\ell) \ \mathbf{i} \end{array}$$

$$\begin{array}{c} \textit{she will see him} \\ \pi_3 \ (\pi^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{i}\mathbf{o}^\ell)\mathbf{o} \end{array} \ .$$

We have introduced the *basic types*

$\pi =$ subject,

$\pi_3 =$ third person singular subject,

$\mathbf{s}_1 =$ declarative sentence in present tense,

$\mathbf{i} =$ infinitive of intransitive verb,

$\mathbf{j} =$ infinitive of any complete verb phrase,

$\mathbf{o} =$ direct object.

We postulate

$$\pi_3 \to \pi, \quad \mathbf{i} \to \mathbf{j},$$

where the arrow denotes what mathematicians call a *partial order*. By this is meant a binary relation which is

reflexive: $x \to x$,

transitive: $x \to y$ and $y \to z$ implies $x \to y$, and

anti-symmetric: $x \to y$ and $y \to x$ implies $x = y$.

We may read $x \to y$ as saying that everything of type x is also of type y.

In the above examples, the pronoun *she* has type π_3 to indicate that it serves as a third person singular subject. It also has type π, since $\pi_3 \to \pi$. The instransitive verb *come* has type $\mathbf{i}$ and also $\mathbf{j}$, since $\mathbf{i} \to \mathbf{j}$. The transitive verb *see* has type $\mathbf{i}\mathbf{o}^\ell$, to indicate that it is looking for an object on the right. The superscript ℓ is to show that $\mathbf{o}^\ell$ cancels $\mathbf{o}$ from the *left*, that is, $\mathbf{o}^\ell\mathbf{o} \to 1$. The finite modal verb-form *will* here has type $\pi^r\mathbf{s}_1\mathbf{j}^\ell$ to indicate that it is looking for a subject on the left and for an infinitive on the right. The superscript r is to show that π^r cancels π from the *right*, hence also π_3, since

$$\pi_3\pi^r \to \pi\pi^r \to 1.$$

The superscript ℓ again reminds us that $\mathbf{j}^\ell$ cancels $\mathbf{j}$ from the *left*, hence also $\mathbf{i}$, since

$$\mathbf{j}^\ell\mathbf{i} \to \mathbf{j}^\ell\mathbf{j} \to 1.$$

The subscript 1 in the type $\pi^r \mathbf{s}_1 \mathbf{j}^\ell$ is to indicate that I consider the resulting sentence to be in the "present" tense. I have been told that this sentence is not in the present, but in the "future" tense. I could have evaded this objection, had I replaced *will* by *may*, since no one is likely to say that *she may come* is in the "permissive" tense. English, like German, but unlike French or Latin, has no "simple" future. Indeed, *will come* is often viewed as a "compound" future of *come*, expressed with the help of the present tense of the auxiliary verb *will* (even though I was taught in school to say *I shall come* in contrast to *she will come*). There are other ways for expressing the future, e.g. *I am going to come*, colloquially *I'm gonna come.* We may even say *I am coming tomorrow*, employing the continuous present of *come.*

Why is it necessary to distinguish between $\mathbf{i}$ and $\mathbf{j}$? For the moment, let me only point out that to negate sentences such as

she sleeps,   *she has slept*

different strategies are required. The former gives rise to

she does not sleep

and the latter to

she has not slept,

whereas

**she does not have slept*

is considered unacceptable. This shows that the infinitives *sleep* and *have slept* require different types, here types $\mathbf{i}$ and $\mathbf{j}$ respectively.

For purposes of calculation, we need only the associative law of concatenation

$$(xy)z = x(yz)$$

and the type 1 of the empty string, such that

$$1x = x = x1,$$

as well as the *contractions*

$$xx^r \to 1, \quad x^\ell x \to 1,$$

where r stands for *right* and ℓ for *left adjoint.* Thus, we might calculate as follows:

$$\begin{aligned} \pi_3(\pi^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{i}\mathbf{o}^\ell)\mathbf{o} &= [\pi_3\pi^r]\mathbf{s}_1\,[\mathbf{j}^\ell \mathbf{i}][\mathbf{o}^\ell \mathbf{o}] \\ &\to [\pi\pi^r]\mathbf{s}_1\,[\mathbf{j}^\ell \mathbf{j}]\,[\mathbf{o}^\ell \mathbf{o}] \\ &\to 1\mathbf{s}_1\,1\,1 \;=\; \mathbf{s}_1, \end{aligned}$$

although this is not exactly how a speaker or hearer would proceed, as we will see. The word "adjoint" here is borrowed from category theory and will be given a more precise meaning later. For the moment, we only require that x^r cancels x from the *right* and x^ℓ cancels x from the *left.*

With apologies to Chomsky, I would like to say that the parentheses (and) indicate the *surface* structure and the square brackets [and] the *deep* structure. In what follows I will usually leave out the square brackets and only retain the left square bracket [to avoid ambiguity when a contraction is to be postponed. Other left brackets will be restored in later chapters.

The idea behind the present approach can be traced back to Charles Sanders Peirce, who explored an analogy with Chemistry and would have said that the infinitive *see*, here of type $\mathbf{i}\mathbf{o}^{\ell}$, has an *unsaturated bond* or *valency* on the right, looking for an object, while *will*, of type $\pi^r\mathbf{s}_1\mathbf{j}^{\ell}$, has two unsaturated bonds, one on the left, looking for a subject, and one on the right, looking for an infinitive.

Zellig Harris made use of a notation similar to ours, although he would not have allowed our superscripts r and ℓ to be iterated, as we will do presently. First, consider the question

$$\begin{array}{l} \textit{will she see him ?} \\ (\mathbf{q}_1\mathbf{j}^{\ell}\pi^{\ell})\ \pi_3\ (\mathbf{i}\mathbf{o}^{\ell})\ \mathbf{o} \quad \rightarrow\ \mathbf{q}_1 \end{array}$$

Here we have introduced a new basic type

$\mathbf{q}_1$ = yes-or-no question in present tense

and assigned a new type $\mathbf{q}_1\mathbf{j}^{\ell}\pi^{\ell}$ to *will.*

A similar secondary assignment will apply to other modal and auxiliary verbs in English, whereas in German this can be done to any verb. Note that we have replaced the square brackets by *underlinks*, which were already used by Harris. In particular, the underlinks justify the contractions

$$\begin{array}{c} \pi^{\ell}\pi_3 \rightarrow \pi^{\ell}\pi \rightarrow 1, \\ \mathbf{j}^{\ell}\mathbf{i} \rightarrow \mathbf{j}^{\ell}\mathbf{j} \rightarrow 1 \end{array}$$

and

$$\mathbf{o}^{\ell}\mathbf{o} \rightarrow 1.$$

Now look at

$$\begin{array}{l} \textit{whom will she see} - ? \\ (\overline{\mathbf{q}}\mathbf{o}^{\ell\ell}\mathbf{q}^{\ell})(\mathbf{q}_1\mathbf{j}^{\ell}\pi^{\ell})\pi_3(\mathbf{i}\mathbf{o}^{\ell}) \quad \rightarrow\ \overline{\mathbf{q}} \end{array} \tag{3.1}$$

where we have introduced two new basic types

$\mathbf{q}$ = yes-or-no question in any tense,

$\overline{\mathbf{q}}$ = arbitrary question (including wh-question)

and postulated

$$\mathbf{q}_1 \ \rightarrow\ \mathbf{q}\ \rightarrow\ \overline{\mathbf{q}}.$$

The distinction between $\mathbf{q}$ and $\overline{\mathbf{q}}$ is necessary, since a question of type $\mathbf{q}$ can be preceded by *when* of type $\overline{\mathbf{q}}\mathbf{q}^{\ell}$ and a question of type $\overline{\mathbf{q}}$ cannot. The dash at the end of (3.1) was placed where Chomsky would have put a so-called *trace*. Traces are not needed in our approach and are only indicated to facilitate comparison with Chomsky's approach. It will turn out that a double superscript ℓ makes an appearance whenever Chomsky requires a trace.

Let me confess once and for all that I follow the late Inspector Morse in saying and writing *whom* for the object question pronoun (and for the object relative pronoun), as reported by Colin Dexter, while Chomsky and Pinker follow Sergeant Lewis instead, saying *who*, at least in most contexts. As Steven Pinker points out: "In the U.S. *whom* is used consistently only by careful writers and pretentious speakers." I apologize for being a pretentious speaker, since English is not my native language.

In anticipation of what we will need later, we will replace the type of *whom* by $\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell}$, where

$$\hat{\mathbf{o}} \to \mathbf{o} \not\to \hat{\mathbf{o}}.$$

The reason for this need not concern us now, but the contraction $\hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^{\ell} \to 1$ required for the computation of (3.1) can still be justified, since $\hat{\mathbf{o}} \to \mathbf{o}$ implies $\mathbf{o}^{\ell} \to \hat{\mathbf{o}}^{\ell}$ (as we will see in Chapter 4) and so

$$\hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^{\ell} \to \hat{\mathbf{o}}^{\ell\ell}\hat{\mathbf{o}}^{\ell} \to 1.$$

Moreover, since $\hat{\mathbf{o}}^{\ell\ell} \to \mathbf{o}^{\ell\ell}$, the original type $\overline{\mathbf{q}}\mathbf{o}^{\ell\ell}\mathbf{q}^{\ell}$ for *whom* is still correct.

Let me point out how a sentence such as (3.1) is analyzed in real time, as one proceeds from one word to the next:

$$\begin{aligned}
\textit{whom}: &\quad \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell},\\
\textit{whom will}: &\quad \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{q}^{\ell}\mathbf{q}_1}\mathbf{j}^{\ell}\pi^{\ell} \to \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{j}^{\ell}\pi^{\ell},\\
\textit{whom will she}: &\quad \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{j}^{\ell}\pi^{\ell}\pi_3 \to \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{j}^{\ell},\\
\textit{whom will she see}: &\quad \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{j}^{\ell}\,\mathbf{i}}\,\mathbf{o}^{\ell} \to \overline{\mathbf{q}}\underbrace{\hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^{\ell}} \to \overline{\mathbf{q}}.
\end{aligned}$$

It is tempting to identify the *simple* types $\overline{\mathbf{q}}$, $\hat{\mathbf{o}}^{\ell\ell}$, $\mathbf{q}^{\ell}$ etc with Miller's *chunks of information.* George A. Miller had asserted that the short term memory employed in language processing can hold no more than seven (plus or minus two) chunks of information. Keeping track of the number of chunks in the above calculations, we have

$$3,\ 6 \to 4,\ 5 \to 3,\ 5 \to 3 \to 1$$

chunks at successive stages. We will meet other examples illustrating Miller's thesis later.

While the formalism proposed by Harris will suffice for many languages, it seems that the description of modern European languages can profit from double superscripts ℓ or r, wherever Chomsky's approach postulates a trace. I am tempted to conjecture that such constructions first appeared in medieval or Renaissance Europe, but I must leave this question to experts in historical linguistics.

People often ask me how I arrive at the types assigned to words. My only answer is: "by trial and error". All type assignments are tentative and, like scientific hypotheses, must be revised as soon as new evidence shows up. Many of the type assignments I have proposed in earlier publications have been abandoned by now. Presumably, many of the types proposed here will also be revised in the future. More formally, "learning algorithms" have been devised by Buszkowski, Kanazawa and Forêt for related categorial grammars, which permit discovery of appropriate types from exposure to given texts.

4. Mathematical interlude.

Let me summarize what has been done so far. We presuppose a given partially ordered set of *basic types*, with the partial order denoted by an arrow. From any basic type a we may construct *simple types*

$$\cdots a^{\ell\ell},\ a^{\ell},\ a,\ a^{r},\ a^{rr},\ \cdots$$

by iterating superscripts ℓ and r. The partial order may be extended from basic to simple types by postulating *contravariance*:

$$\text{if } x \to y \text{ then } y^{\ell} \to x^{\ell} \text{ and } y^{r} \to x^{r}.$$

Repeating this process, we obtain $x^{\ell\ell} \to y^{\ell\ell}$, $y^{\ell\ell\ell} \to x^{\ell\ell\ell}$ etc. We also postulate

$$x^{\ell r} = x = x^{r\ell}.$$

By a *compound type*, or just a *type*, we will mean a string $x_1 \cdots x_n$ of simple types with $n \geq 0$. The empty string, with $n = 0$, will be denoted by 1. Strings are multiplied by juxtaposition (concatenation):

$$(x_1 \cdots x_m)\ (y_1 \cdots y_n)\ =\ x_1 \cdots x_m\, y_1 \cdots y_n$$

and

$$x1 = x = 1x.$$

The types constructed so far constitute what mathematicians would call the *free monoid* generated by the set of simple types. But we can also extend the partial order to compound types:

$$\text{if } x \to y \text{ and } x' \to y' \text{ then } xx' \to yy'$$

or equivalently,

$$\text{if } x \to y \text{ then } uxv \to uyv.$$

Moreover, the superscripts ℓ and r may also be applied to compound types if we require

$$(xy)^{\ell} = y^{\ell}x^{\ell},\ (xy)^{r} = y^{r}x^{r},\ 1^{\ell} = 1 = 1^{r}.$$

We call x^{ℓ} the *left adjoint* and x^{r} the *right adjoint* of x, the terminology being borrowed from category theory.

We postulate the *contraction rules*

$$x^{\ell}x \to 1, \qquad xx^{r} \to 1$$

for simple types x. But the same rules can then be shown to apply to compound types, e.g.

$$(xy)^{\ell}(xy) = (y^{\ell}x^{\ell})(xy) = y^{\ell}(x^{\ell}x)y \to y^{\ell}1y = y^{\ell}y \to 1.$$

In addition to contractions we also have *expansions*

$$1 \to xx^{\ell}, \quad 1 \to x^{r}x.$$

For example, from $xx^{r} \to 1$ we can infer that

$$1 = 1^{\ell} \to (xx^{r})^{\ell} = x^{r\ell}x^{\ell} = xx^{\ell}.$$

Finally, we postulate for types:

$$\text{if } x \to y \text{ and } y \to x \text{ then } x = y.$$

Of course, the anti-symmetry rule is already known for basic types.

The system of types described so far has been called the *free pregroup* generated by the partially ordered set of basic types. A *pregroup grammar* of a language is obtained from an assignment of types, that is, elements of the free pregroup generated by the language specific set of basic types, to the words of the language.

To accommodate readers with a phobia for mathematics, I defer the precise definition and a more thorough discussion of pregroups to a later chapter (Chapter 27). Readers who feel more comfortable with mathematics may consult this now.

The next five chapters will be devoted to a discussion of basic word classes in English, interrupted only by a short philosophical digression.

II. Basic Word Classes.

5. Nouns.

In addition to names, such as *John* and *Napoleon*, there are three kinds of nouns in English:

count nouns: *bean, pig, idea, delivery*, ...

mass nouns: *rice, pork, water, sadness*, ...

plurals: *beans, pigs*, ..., *police, cattle*, ..., *pants, scissors*, ...

All count nouns have plurals. With a small number of exceptions, these are formed by adding the morpheme $+s$. However, not all plurals are derived from count nouns. There is no such thing as **one pant*. Mass nouns have no plurals, so we cannot say **rices* or **porks*.

We will adopt the following basic types:

$$\begin{aligned} \mathbf{n} &= \text{ name,} \\ \mathbf{n}_0 &= \text{ mass noun,} \\ \mathbf{n}_1 &= \text{ count noun,} \\ \mathbf{n}_2 &= \text{ plural.} \end{aligned}$$

The four kinds of nouns differ in the determiners they admit or require. Names usually occur without determiner, as in

$$\begin{array}{cc} \textit{Napoleon slept}, & \textit{I see John} \\ \mathbf{n}\ (\pi_3^r \mathbf{s}_2) \to \mathbf{s}_2 & \pi_1\ (\pi_1^r \mathbf{s}_1\, \mathbf{o}^\ell)\mathbf{n} \to \mathbf{s}_1 \end{array}$$

where we postulate

$$\mathbf{n} \to \pi_3,\ \mathbf{o}$$

and calculate e.g.

$$\mathbf{n}\pi_3^r \to \pi_3\pi_3^r \to 1.$$

Although verbs and pronouns will only be considered later, we like to have some freedom to produce examples, so we have made use of some additional types already:

$$\begin{aligned} \mathbf{s}_2 &= \text{ declarative sentence in past tense,} \\ \pi_1 &= \text{ first person subject,} \\ \pi_2 &= \text{ plural or second person subject.} \end{aligned}$$

For the record, we also postulate

$$\pi_k \to \pi \quad (k = 1, 2, 3).$$

Mass nouns do not require determiners, but they admit certain ones, such as *much*:

$$\begin{array}{cccc} \textit{water is wet}, & & \textit{I drank much water.} & \\ \mathbf{n}_0\ (\pi_3^r \mathbf{s}_1\, \bar{\mathbf{a}}^\ell)\mathbf{a} & \to\ \mathbf{s}_1 & \pi_1\ (\pi^r\, \mathbf{s}_2\, \mathbf{o}^\ell)(\bar{\mathbf{n}}_0\ \mathbf{n}_0^\ell)\ \mathbf{n}_0 & \to\ \mathbf{s}_2 \end{array}$$

Here we have introduced some new basic types:

$$\begin{aligned} \mathbf{a} &= \text{ predicative adjective,} \\ \bar{\mathbf{a}} &= \text{ predicative adjectival phrase,} \\ \bar{\mathbf{n}}_i &= \text{ complete noun phrase.} \end{aligned}$$

(Conceivably, $\overline{\mathbf{n}}_0 = \overline{\mathbf{n}}_1$.) We postulate

$$\begin{aligned}&\mathbf{a} \to \overline{\mathbf{a}} \not\to \mathbf{a},\\ &\mathbf{n}_0 \to \overline{\mathbf{n}}_0 \to \pi_3,\ \mathbf{o},\\ &\mathbf{n}_2 \to \overline{\mathbf{n}}_2 \to \pi_2,\ \mathbf{o},\\ &\overline{\mathbf{n}}_1 \to \pi_3,\ \mathbf{o}.\end{aligned}$$

(The reason for distinguishing between $\mathbf{a}$ and $\overline{\mathbf{a}}$ will become clear later. When writing $x \not\to y$, I do not wish to postulate the negation of $x \to y$, but to negate postulating $x \to y$.)

On the other hand, count nouns do require determiners, such as articles:

$$\begin{array}{c} \textit{the king is dead,} \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\ \mathbf{n}_1\ (\pi_3^r\, \mathbf{s}_1\, \overline{\mathbf{a}}^{\ell})\mathbf{a} \end{array} \to \mathbf{s}_1$$

$$\begin{array}{c} \textit{she met a boy.} \\ \pi_3\ (\pi^r\, \mathbf{s}_2\, \mathbf{o}^{\ell})(\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\mathbf{n}_1 \end{array} \to \mathbf{s}_2$$

Note that the definite article *the* is universal, having type $\overline{\mathbf{n}}_i\mathbf{n}_i^{\ell}$ for $i = 0, 1, 2$, but the indefinite article $a(n)$ does not apply to mass nouns or plurals, hence has type $\overline{\mathbf{n}}_1\mathbf{n}_1^{\ell}$ only.

Plural nouns do not require determiners, but they admit certain ones, such as *many*:

$$\begin{array}{c} \textit{kings will die,} \\ \mathbf{n}_2\ (\pi^r\, \mathbf{s}_1\, \mathbf{j}^{\ell})\mathbf{i} \end{array} \to \mathbf{s}_1$$

since

$$\mathbf{n}_2 \to \overline{\mathbf{n}}_2 \to \pi_2 \to \pi,$$

$$\begin{array}{c} \textit{they have many pigs .} \\ \pi_2\ (\pi_2^r\, \mathbf{s}_1\, \mathbf{o}^{\ell})(\overline{\mathbf{n}}_2\, \mathbf{n}_2^{\ell})\ \mathbf{n}_2 \end{array} \to \mathbf{s}_1$$

There is some flexibility between the four kinds of nouns. People may say

the real McCoy, *a good beer*,
an interesting phenomena,

using a name, a mass noun or a plural as a count noun. One can imagine saying

cannibals prefer man to pork,

using a count noun as a mass noun. When the doctor says

Nurse is absent today,

he is using a count noun as a name.

Some transitions from one kind of noun to another have been codified by the passage of time. For example, the old mass noun *pease* is now seen as the plural *peas* of a newly conceived count noun *pea*. I have asked my students whether *rice* will ever be viewed as a plural like *mice*.

The distinction between count nouns and mass nouns may play a rôle in how we analyze a phrase. Consider, for example:

(a) *a man-eating tiger*,

(b) *a man eating soup.*

These have completely different structure. In (a), the count noun *tiger*, lacking an immediately preceding determiner, cannot be the object of *eating.* In fact the indefinite article at the beginning of (a) modifies the compound noun phrase. On the other hand, the mass noun *soup* cannot be linked to the indefinite article at the beginning of (b), so *soup* must be interpreted as the object of *eating.* (Note that the hyphen in (a) is not audible.)

However, the noun *fish* can serve either as a mass noun or as a count noun, hence the phrase

(c) *a man eating fish*

is ambiguous without the inaudible hyphen. Here are two ways of analyzing it:

$$\begin{array}{c} a\ man\ eating\ fish, \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\, [\mathbf{n}_1\, (\mathbf{n}_1^{r}\, \mathbf{n}_1\, \mathbf{o}^{\ell})\, \mathbf{n}_0 \end{array} \quad \rightarrow\ \overline{\mathbf{n}}_1$$

where *eating fish* is viewed as modifying *man* from the right, and the left square bracket reminds us to defer the contraction $\mathbf{n}_1^{\ell}\mathbf{n}_1 \rightarrow 1$, and

$$\begin{array}{c} a\ (man - eating)\ fish, \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\ (\mathbf{n}_1\, \mathbf{n}_1^{\ell})\ \mathbf{n}_1 \end{array} \quad \rightarrow\ \overline{\mathbf{n}}_1$$

where *man-eating* is viewed as a compound attributive adjective. Thus, the noun phrase (c) could refer to a male diner or to a shark.

Fortunately, unlike French or German, English does not have a grammatical gender. Why should a table be denoted by a feminine noun in French, but by a masculine one in German? In English, gender is based on sex only, with minor exceptions, as when sailors refer to a ship as *she.*

A curious development has taken place with the word *person.* Although feminine in both French and German, it was originally neutral in English and could therefore be substituted for *man* in many contexts, for the sake of political correctness. As a result, *person Friday* or *chairperson* are now often understood to refer to females! In French, the cognate *personne* can exceptionally be masculine, but then it means "nobody".

English, unlike some languages, has no explicit dual forms, but allows us to speak of "a pair of scissors". But even the distinction between singular and plural is sometimes blurred. I would say

$$a\ number\ of\ persons\ were\ in\ the\ room,$$

but

$$the\ number\ of\ persons\ in\ the\ room\ was\ decreasing.$$

In addition to gender and number, many languages distinguish different cases of nouns, adjectives and pronouns. Latin has six cases; the word *case* meant "fall" and hints at the six ways a die can fall. English has preserved two cases, distinguishable in pronouns only, corresponding to Latin nominative and accusative, for the subject and direct object respectively. English also has an invisible dative, for the indirect object, which manifests itself syntactically as distinct from the direct object.

6. Philosophical digression.

The distinction between count nouns and mass nouns was presumably recognized in ancient Greece and may have played an important rôle in early Greek philosophy and mathematics. Observation revealed that concrete objects in the world could either be counted, like sheep, or subdivided without loss of identity, like water. The earliest Greek philosophers held different opinions as to which kind of objects was more basic. Thales said that all things are water; but some of his followers replaced or augmented water by other substances, namely earth, air and, ultimately, fire. On the other hand, Democritus claimed that everything was made up of indivisible units, which he called "atoms".

In mathematics too, there were geometrical quantities to be measured and arithmetical entities to be counted. The legendary Pythagoras was even alleged to have said "everything is number", having in mind positive integers or ratios of such. Unfortunately, his followers were reluctantly forced to recognize that the diagonal of a unit square could not be measured by such a ratio.

Aristotle was to resolve the philosophical problem by introducing a distinction between "matter" and "forms", the former to be measured and the latter to be counted. You could measure water by counting cups, as in "two cups of water". The mathematical dichotomy had already been resolved by Eudoxus and Theaetetus, who explained ratios of geometric quantities (now called positive real numbers) in terms of what we now call "Dedekind cuts" and "continued fractions" respectively.

It is widely believed that nouns, at least concrete nouns, denote things, or perhaps kinds of things. What is not evident is whether the things were there before they were named or whether the name created the thing. This question has been a subject of intense debate by philosophers, old and new, and, more recently, by psychologists. Surely, pigs and water existed before anyone uttered the word *pig* or *water*. (A squirrel can tell a pig from a tree and water from grass, but can it tell a pig from a sheep?) But what about *uncle* and *king*? These concepts depend on human social organization. The problem becomes even more acute when we look at abstract nouns, such as *time, hope, charity*, etc.

Philo of Alexandria tried to reconcile biblical theology with Greek philosophy. Wondering about the nature of Plato's *ideas* (meant to be "ideal objects", not "thoughts"), he decided that they should be identified with words, namely with nouns. Undoubtedly, he was influenced by the observation that *word* and *thing* share the same translation into Hebrew. His position was that words create things, hence the gospel of John later said "the Word (logos) became flesh". This suggested identification of the Logos with a divine creator was presumably influenced by an earlier proposal by Heraclitus, for whom the Logos was the law (of nature).

Modern physics has resolved the ancient philosophical controversy in favour of indivisible units, although no longer referred to as "atoms", as against substances. Physicists nowadays recognize two kinds of elementary particles: *fermions*, including electrons, neutrinos and quarks, and *bosons*, including photons, weak vector bosons and gluons. The existence of other elementary particles, such as the graviton and the Higgs, has been conjectured, but has not yet been verified experimentally. However, most present day physicists agree with the ancient philosopher Zeno, who had shown in one of his paradoxes that time and space do not come in discrete atoms. They still retain the real numbers for describing time and space. Other paradoxes by Zeno had exposed some problems with the continuity of motion as well, but these don't seem to bother practicing physicists. On the other hand, mathematicians have overcome these problems with the help of the concept of limit, and logicians have done so using non-standard analysis.

7. Adjectives.

Adjectives in English play two distinct rôles. They can serve as predicates, as in

$$(a) \qquad \underset{\pi_3\,(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)\mathbf{a}}{\textit{she is poor,}} \quad \rightarrow \ \mathbf{s}_1$$

or as attributes modifying nouns, as in

$$(b) \qquad \underset{(\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell)(\mathbf{n}_1\,\mathbf{n}_1^\ell)\mathbf{n}_1}{\textit{the poor man,}} \ \rightarrow \ \overline{\mathbf{n}}_1 \quad \underset{(\mathbf{n}_2\,\mathbf{n}_2^\ell)\mathbf{n}_2}{\textit{poor men,}} \ \rightarrow \ \mathbf{n}_2, \quad \underset{(\mathbf{n}_0\,\mathbf{n}_0^\ell)\mathbf{n}_0}{\textit{poor wine,}} \ \rightarrow \ \mathbf{n}_0$$

Thus, we assign to adjectives the type $\mathbf{a}$ as well as $\mathbf{n}_i\mathbf{n}_i^\ell$, where $i = 0, 1$ or 2. We will ignore the marginal *poor Jane*, where the adjective modifies a name.

English speakers are lucky that adjectives need not bear assorted features. In German, the attributive adjective carries an ending which encodes gender, number and case, and these features must agree with those of the noun it modifies. The Latin adjective displays similar features even when used as a predicate.

Our notation explains why you can say

$$(c) \qquad \underset{(\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell)(\mathbf{n}_1\,\mathbf{n}_1^\ell)(\mathbf{n}_1\,\mathbf{n}_1^\ell)\mathbf{n}_1}{\textit{the poor old man}} \quad \rightarrow \ \overline{\mathbf{n}}_1$$

but not

$$(d) \qquad \underset{\pi_3(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)\mathbf{a}\ \mathbf{a}}{{}^*\textit{she is poor old,}} \quad \rightarrow \ \mathbf{s}_1\mathbf{a} \not\rightarrow \mathbf{s}_1.$$

I confess that it won't explain why *poor old* is preferable to *old poor*.

Now let us modify the adjective by an adverb such as *very*. This is easy for (a), if we assign the type $\mathbf{a}\mathbf{a}^\ell$ to the adverb *very*, yielding

$$(e) \qquad \underset{(\pi_3\,(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)\,(\mathbf{a}\,\mathbf{a}^\ell)\mathbf{a}}{\textit{she is very poor}} \quad \rightarrow \ \mathbf{s}_1$$

but not for (b). The tentative type

$$(\mathbf{n}_i\mathbf{n}_i^\ell)(\mathbf{n}_i\mathbf{n}_i^\ell)^\ell = \mathbf{n}_i\mathbf{n}_i^\ell\mathbf{n}_i^{\ell\ell}\mathbf{n}_i^\ell \rightarrow \ \mathbf{n}_i\mathbf{n}_i^\ell$$

easily leads to wrong predictions. My solution is to provide the attributive adjective with an invisible ending of type $\mathbf{a}^r\mathbf{n}_i\mathbf{n}_i^\ell$, so that *poor* will have the additional type

$$\mathbf{a}\,\mathbf{a}^r\mathbf{n}_i\mathbf{n}_i^\ell \ \rightarrow \ \mathbf{n}_i\mathbf{n}_i^\ell,$$

but now it is easily modified by *very*:

$$\underset{(\mathbf{a}\,\mathbf{a}^\ell)(\mathbf{a}\,\mathbf{a}^r\,\mathbf{n}_i\mathbf{n}_i^\ell)}{\textit{very poor}} \quad \rightarrow \ \mathbf{n}_i\mathbf{n}_i^\ell.$$

Introducing invisible endings may be a mathematical trick, but German attributive adjectives do carry visible endings

$$+e,\ +em,\ +en,\ +er,\ +es,$$

which encode information not only about the features gender, number and case, but also whether the adjective is strong or weak, which correlates with whether the noun phrase containing it is definite or indefinite.

The present participles of intransitive verbs and the past participles of transitive ones behave much like adjectives. We adopt the basic types

$$\begin{aligned}\mathbf{p}_1 &= \text{ present participle,}\\ \mathbf{p}_2 &= \text{ past participle,}\end{aligned}$$

both for intransitive verbs, hence we assign the types $\mathbf{p}_j\mathbf{o}^\ell$, with $j = 1$ or 2, to participles of transitive verbs. The present participles of intransitive verbs and the past participles of transitive ones may be used attributively and may then carry invisible endings of types $\mathbf{p}_1^r\mathbf{n}_i\mathbf{n}_i^\ell$ and $\mathbf{o}\mathbf{p}_2^r\mathbf{n}_2\mathbf{n}_2^\ell$ respectively. Thus we have

$$\begin{array}{c}\textit{sleeping dogs,}\\ (\mathbf{p}_1\mathbf{p}_1^r\,\mathbf{n}_2\,\mathbf{n}_2^\ell)\,\mathbf{n}_2\end{array} \rightarrow \mathbf{n}_2$$

$$\begin{array}{c}\textit{beaten dogs.}\\ (\mathbf{p}_2\,\mathbf{o}^\ell\mathbf{o}\,\mathbf{p}_2^r\,\mathbf{n}_2\,\mathbf{n}_2^\ell)\mathbf{n}_2\end{array} \rightarrow \mathbf{n}_2$$

Yet *sleeping* and *beaten* are not genuine adjectives, since one does not say

$$^*\textit{very sleeping}, \qquad ^*\textit{very beaten}.$$

(Note that German participles carry visible endings like German adjectives.)

However, there are many verbs whose participles, both present and past, are indeed adjectives and are usually, even if not always, listed as such in standard dictionaries. Here is a partial list:

(7.1)
amuse, *annoy*, *charm*, *convince*,
depress, *disappoint*, *discourage*, *distress*,
disturb, *excite*, *fascinate*, *frighten*, *interest*,
intimidate, *intoxicate*, *please*, *satisfy*,
surprise, ...

For example, we can say

very disappointed readers did not find the book very interesting.

Must English speakers memorize the long list of verbs whose participles double as adjectives or is there some criterion they can use? It seems that all these verbs are transitive and that they carry a meaning which refers to the *causation of an emotional or mental state*. Thus

(7.2) X *frightens* Y means X *causes* Y *to be afraid.*

This semantic criterion does not tell the whole story, but it goes a long way towards summarizing numerous cases and excluding others.

There are a number of verbs for which only one of the two participles is a genuine adjective, e.g.

(7.3) *heat*, *promise*, ...

and these must be annotated accordingly in the dictionary.

Here is a word about the syntax-semantics boundary concerning one of the verbs on our list (7.1). As Chomsky was fond of pointing out,

sincerity frightens John

is more grammatical than

John frightens sincerity.

However, any reader who says that the latter is ungrammatical because "sincerity cannot be frightened" should think again: she has just made a statement which she believes to be ungrammatical.

It may be of interest to find out whether people with Asperger's syndrome, alleged to be unaware of other people's mental or emotional states, will come up with the same list (7.1).

Some adjectives require or admit complements. We will take a look at two of Chomsky's favourite examples: the adjectives *eager* and *easy*. In addition to the usual type $\mathbf{a}$ they also have types $\mathbf{a}\bar{\mathbf{j}}^{\ell}$ and $\mathbf{a}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{j}}^{\ell}$ respectively, where

$$\bar{\mathbf{j}} = \text{ complete infinitive with } \mathit{to}.$$

Thus we have

$$\begin{array}{c} \textit{eager to come} \\ (\mathbf{a}\,\bar{\mathbf{j}}^{\ell})\;(\bar{\mathbf{j}}\,\mathbf{j}^{\ell})\,\mathbf{i} \end{array} \quad \to \; \mathbf{a}$$

and

$$\begin{array}{c} \textit{easy to dislike} \\ (\mathbf{a}\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^{\ell})\;\bar{\mathbf{j}}\,\mathbf{j}^{\ell})\;(\mathbf{i}\,\mathbf{o}^{\ell}) \end{array} \quad \to \; \mathbf{a}$$

As Chomsky has pointed out, the word *too* may turn any adjective into an adjectival phrase of the same type as *eager* or *easy*. Therefore, *too* may be assigned the type $\bar{\mathbf{a}}\bar{\mathbf{j}}^{\ell}\mathbf{a}^{\ell}$ or $\bar{\mathbf{a}}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{j}}^{\ell}\mathbf{a}^{\ell}$, as well as the simpler $\bar{\mathbf{a}}\bar{\mathbf{a}}^{\ell}$, where

$$\mathbf{a} \;\to\; \bar{\mathbf{a}} \;\not\to\; \mathbf{a}.$$

This will account for

$$\begin{array}{c} \textit{too frightened} \\ (\bar{\mathbf{a}}\,\bar{\mathbf{a}}^{\ell})\;\mathbf{a} \end{array} \quad \to \; \bar{\mathbf{a}},$$

$$\begin{array}{c} \textit{too frightened to flee} \\ (\bar{\mathbf{a}}\,\bar{\mathbf{j}}^{\ell}\,\mathbf{a}^{\ell})\;\mathbf{a}\;(\bar{\mathbf{j}}\;\mathbf{i}^{\ell})\;\mathbf{i} \end{array} \quad \to \; \bar{\mathbf{a}}$$

$$\begin{array}{c} \textit{too frightened to avoid} - \\ (\bar{\mathbf{a}}\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^{\ell}\,\mathbf{a}^{\ell})\ \mathbf{a}\ (\bar{\mathbf{j}}\,\mathbf{i}^{\ell})\ (\mathbf{i}\,\mathbf{o}^{\ell}) \end{array} \quad \to\ \bar{\mathbf{a}}$$

without permitting

$$\begin{array}{c} {}^{*}\textit{very too frightened} \\ (\mathbf{a}\mathbf{a}^{\ell})\ (\overline{\mathbf{a}}\overline{\mathbf{a}}^{\ell})\ \mathbf{a} \end{array}$$

Unfortunately, our typing also permits the debatable *too very frightened.*

However, concerning the types of *easy* and *too*, this is not the complete story, since *easy to dislike* may be conceived as elliptical for

$$\textit{easy [for me] to dislike},$$

and its ultimate analysis should await the discussion of such quasi-sentences as

$$\textit{[for] me to dislike him.}$$

8. Verbs.

The English verb has four simple tenses. To keep our discussion within reasonable bounds, we will not discuss the two almost obsolete subjunctives (nor the related imperative). That leaves the so-called *present* and *past*, although, in line with some non-European languages, these would have been better named "imperfective" and "perfective" respectively. The former often refers to the future and sometimes even to the past, as in the works of Damon Runyan. Whereas most European languages have six persons, three singular and three plural, in English the three persons of the plural are indistinguishable and have even absorbed the second person singular, since the ancient *thou* has disappeared. To simplify matters, we will let the second person singular include the three plural ones.

Thus, with each verb V we may associate a 2 by 3 matrix $C_{jk}V$, with $j = 1$ or 2 referring to the tense and $k = 1, 2$ or 3 to the person. All except the modal verbs are represented by their infinitive, since the ancient infinitives of *will, can* etc have disappeared. It is convenient to regard *would* and *could* formally as the past tenses of *will* and *can*, even if this is no longer justified on semantic grounds.

For example,

$$\begin{aligned} C_{jk}\ be\ &\to \begin{pmatrix} am\ are\ is \\ was\ were\ was \end{pmatrix}, \\ C_{jk}\ go\ &\to \begin{pmatrix} go\ go\ goes \\ went\ went\ went \end{pmatrix}, \\ C_{jk}\ will\ &\to \begin{pmatrix} will\ will\ will \\ would\ would\ would \end{pmatrix}. \end{aligned}$$

In addition to the infinitives, except for modals, there are two participles

$$\begin{aligned} P_j\ be\ &\to \begin{pmatrix} being \\ been \end{pmatrix}, \\ P_j\ go\ &\to \begin{pmatrix} going \\ gone \end{pmatrix}, \end{aligned}$$

again except for modals. I call C_{jk} and P_j *inflectors*. They act on the verb to generate the finite forms and the participles.

We distinguish between *main* verbs and *auxiliary* verbs, the latter including modals. Main verbs such as *sleep* and *go*, said to be instransitive, will have an infinitive of type $\mathbf{i}$. Main verbs such as *like* and *frighten*, said to be transitive, require a direct object complement and will have an infinitive of type $\mathbf{io}^\ell$. Main verbs with other complements will be discussed as the occasion arises.

Here is a sample calculation:

$$\begin{aligned} C_{13}\ go\ &\to\ goes. \\ (\pi_3^r \mathbf{s}_1 \mathbf{j}^\ell)\mathbf{i}\ &\to\ \pi_3^r \mathbf{s}_1 \end{aligned}$$

Note that the upper arrow generates the verb form and the lower arrow represents the partial order in the pregroup of types. We have assigned the same type $\pi_k^r \mathbf{s}_j \mathbf{j}^\ell$ to the inflector C_{jk} as to the modal *will/would.* When $j = 2$, the person is irrelevant, so we might have written

$$\begin{aligned} C_{2k}\ go\ &\to\ went, \\ (\pi^r \mathbf{s}_2 \mathbf{j}^\ell)\ \mathbf{i}\ &\to\ \pi^r \mathbf{s}_2 \end{aligned}$$

where $\pi_k \to \pi$. Here are two more examples:

$$\begin{array}{rcl} C_{11}\ \textit{like} & \to & \textit{like}, \\ (\pi_1^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) & \to & \pi_1^r \mathbf{s}_1 \mathbf{o}^\ell \end{array}$$

$$\begin{array}{rcl} C_{13}\ \textit{like} & \to & \textit{likes}. \\ (\pi_3^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) & \to & \pi_3^r \mathbf{s}_1 \mathbf{o}^\ell \end{array}$$

It must be pointed out that many verbs can be either transitive or intransitive. For example, *see* and *kill* have types $\mathbf{i}$ as well as $\mathbf{io}^\ell$. In fact, *see* can also have type $\mathbf{is}^\ell$, as in

$$\begin{array}{l} \textit{I see you saw her}, \\ \pi_1\,(\pi_1^r\,\mathbf{s}_1\,\mathbf{s}^\ell)\,\pi_2\,(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell)\mathbf{o} \quad \to \quad \mathbf{s}_1 \end{array}$$

recalling that $\pi_2 \to \pi$ and postulating $\mathbf{s}_2 \to \mathbf{s}$, where

$$\mathbf{s} = \text{declarative sentence.}$$

When assigning types to the inflectors P_j we have to be careful, since the auxiliary verbs *be* and *have*, when used to form compound tenses, don't necessarily admit all participles:

**she is being coming*,
**she is having come*,
**she has had come*.

We resolve this problem by introducing two intermediate types between $\mathbf{i}$ and $\mathbf{j}$:

$$\mathbf{i} \to \mathbf{i}' \to \mathbf{j}' \to \mathbf{j}$$

and by assigning the following types to the inflectors P_j:

$$P_1 : \mathbf{p}_1\mathbf{i}'^\ell, \quad P_2 : \mathbf{p}_2\mathbf{j}'^\ell.$$

At the same time, we assign types to the auxiliary verbs forming the *progressive* tense, the *perfect* tense and the *passive* voice:

$$\begin{array}{rl} \textit{be}_{\text{prog}} : & \mathbf{j}'\mathbf{p}_1^\ell, \\ \textit{have}_{\text{perf}} : & \mathbf{j}\mathbf{p}_2^\ell, \\ \textit{be}_{\text{pass}} : & \mathbf{i}'\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell. \end{array}$$

Note that the passive can also be formed with

$$\textit{get} : \mathbf{i}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell.$$

Here are some illustrations:

$$\begin{array}{l} \textit{they arrive}\ , \\ \pi_2(\pi_2^r\,\mathbf{s}_1\,\mathbf{j}^\ell\mathbf{i}) \quad \to \quad \mathbf{s}_1 \end{array}$$

they will arrive ,

$$\pi_2\ (\pi^r \mathbf{s}_1 \mathbf{j}^\ell)\ \mathbf{i} \quad \to\ \mathbf{s}_1$$

they have arrived ,

$$\pi_2\ (\pi_2^r \mathbf{s}_1 \mathbf{j}^\ell \mathbf{j}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_1$$

they are arriving ,

$$\pi_2\ (\pi_2^r \mathbf{s}_1 \mathbf{j}^\ell \mathbf{j}'\, \mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_1$$

they do arrive ,

$$\pi_2\ (\pi_2^r \mathbf{s}_1 \mathbf{i}^\ell)\ \mathbf{i} \quad \to\ \mathbf{s}_1$$

they are seen − ,

$$\pi_2(\pi_2^r \mathbf{s}_1 \mathbf{j}^\ell \mathbf{i}'\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}\, \mathbf{o}^\ell) \quad \to\ \mathbf{s}_1$$

they get seen − .

$$\pi_2(\pi_2^r \mathbf{s}_1 \mathbf{j}^\ell \mathbf{i}\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}\, \mathbf{o}^\ell) \quad \to\ \mathbf{s}_1$$

However, we are prevented from forming the following:

**she is being coming,*

$$(\mathbf{p}_1\ \mathbf{i}'^\ell \mathbf{j}'\ \mathbf{p}_1^\ell)$$

**she is having come,*

$$(\mathbf{p}_1\ \mathbf{i}'^\ell \mathbf{j}\ \mathbf{p}_2^\ell)$$

**she has had come,*

$$(\mathbf{p}_2\ \mathbf{j}'^\ell \mathbf{j}\ \mathbf{p}_2^\ell)$$

as long as $\mathbf{j}' \not\to \mathbf{i}',\ \mathbf{j} \not\to \mathbf{i}',\ \mathbf{j} \not\to \mathbf{j}'$.

Our type assignments also permit some multiple compound tenses. Varying person and tense we obtain:

I have been arriving,

$$\pi_1(\pi_1^r \mathbf{s}_1 \mathbf{j}^\ell \mathbf{j}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{j}'\, \mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_1$$

you will have arrived,

$$\pi_2\ (\pi^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{j}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_1$$

she will be arriving,

$$\pi_3(\pi^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{j}'\, \mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_1$$

we would have been arriving,

$$\pi_2(\pi^r \mathbf{s}_2 \mathbf{j}^\ell)(\mathbf{j}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{j}'\, \mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}) \quad \to\ \mathbf{s}_2$$

you would be seen −,

$$\pi_2(\pi^r \mathbf{s}_2 \mathbf{j}^\ell)(\mathbf{i}'\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}\, \mathbf{o}^\ell) \quad \to\ \mathbf{s}_2$$

and somewhat reluctantly:

$$\begin{array}{c} \textit{they would have been being seen} - , \\ \pi_2(\pi^r \mathbf{s}_2 \mathbf{j}^\ell)(\mathbf{j}\,\mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{j}'\,\mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}'\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}\,\mathbf{o}^\ell) \;\rightarrow\; \mathbf{s}_2 \end{array}$$

although this would become more acceptable if we replace *being seen* by

$$\begin{array}{c} \textit{getting seen} . \\ (\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2 \mathbf{j}'^\ell \mathbf{i}\,\mathbf{o}^\ell) \;\rightarrow\; \mathbf{p}_1 \end{array}$$

At this stage we can also assign a type to the copula *be* when it is used for introducing predicative adjectives of type $\mathbf{a}$. We put

$$be_{\text{cop}} : \; \mathbf{i}'\overline{\mathbf{a}}^\ell$$

and recall that $\mathbf{a} \rightarrow \overline{\mathbf{a}}$ to account for

$$\begin{array}{c} \textit{she will be good} , \\ \pi_3\ (\pi^r \mathbf{s}_1 \mathbf{j}^\ell)(\mathbf{i}'\,\overline{\mathbf{a}}^\ell)\ \mathbf{a} \;\rightarrow\; \mathbf{s}_1 \end{array}$$

$$\begin{array}{c} \textit{she is being good} , \\ \pi_3\ (\pi_3^r \mathbf{s}_1\,\mathbf{p}_1^\ell)(\mathbf{p}_1 \mathbf{i}'^\ell \mathbf{i}'\,\overline{\mathbf{a}}^\ell)\ \mathbf{a} \;\rightarrow\; \mathbf{s}_1 \end{array}$$

$$\begin{array}{c} {}^*\textit{she does be good} , \\ \pi_3\ (\pi_3^r \mathbf{s}_1 \mathbf{i}^\ell)(\mathbf{i}'\overline{\mathbf{a}}^\ell)\ \mathbf{a} \;\not\rightarrow\; \mathbf{s}_1 \end{array}$$

$$\begin{array}{c} \textit{she is too good} . \\ \pi_3\ (\pi_3^r \mathbf{s}_1\,\overline{\mathbf{a}}^\ell)(\overline{\mathbf{a}}\,\mathbf{a}^\ell)\mathbf{a} \;\rightarrow\; \mathbf{s}_1 \end{array}$$

9. Adverbs.

It is frequently asserted that there are four basic word classes: nouns, adjectives, verbs and prepositions. The way I see it, prepositions may be viewed as adverbs requiring an object complement, just like transitive verbs. Similarly, subordinate conjunctions may be regarded as adverbs requiring a sentence complement.

The category of adverbs is a catchall for many kinds of words. My favourites are temporal and spatial adverbs like *today* and *here*, because they occupy more restricted positions in a sentence than many other adverbs, thus simplifying our task of assigning types to them. Apparently, for this very reason, McCawley denies that they are "true" adverbs and views them as prepositional phrases without prepositions. It seems to me that this is putting the cart before the horse. Would one deny that *furiously* is a true adverb because its position in a sentence is similarly restricted?

I might have avoided controversy by calling adverbs with the restricted distribution, like *today* and *furiously*, "adverbials". They occur preferably at the end of a sentence, where they may be thought of as modifying the verb and may be assigned the type $\mathbf{i}^r\mathbf{i}$. Less naturally, adverbials may also modify a sentence from the beginning and be assigned type $\mathbf{s}\mathbf{s}^\ell$. Hence prepositions should be assigned types $\mathbf{i}^r\mathbf{i}\mathbf{o}^\ell$ and $\mathbf{s}\mathbf{s}^\ell\mathbf{o}^\ell$, and subordinate conjunctions the types $\mathbf{i}^r\mathbf{i}\mathbf{s}^\ell$ and $\mathbf{s}\mathbf{s}^\ell\mathbf{s}^\ell$. Here are some illustrations, where the left bracket guards against premature contraction:

I will meet you tomorrow,

$$\pi_1(\pi^r\,\mathbf{s}_1\,\mathbf{j}^\ell)([\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o}(\mathbf{i}^r\,\mathbf{i}) \quad \to \quad \mathbf{s}_1$$

she had met him in the rain,

$$\pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell[\mathbf{j}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{j}'^\ell\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o}(\mathbf{i}^r\,\mathbf{i}\,\mathbf{o}^\ell)(\overline{\mathbf{n}}_1^\ell\,\mathbf{n}_1)\mathbf{n}_1 \quad \to \quad \mathbf{s}_2$$

he was met when it rained.

$$\pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell[\mathbf{i}'\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{o}^\ell)(\mathbf{i}^r\,\mathbf{i}\,\mathbf{s}^\ell)\,\pi_3(\pi^r\,\mathbf{s}_2) \quad \to \quad \mathbf{s}_2$$

Here we have assigned to *met* and *rained* the already contracted types:

$$\mathbf{p}_2\,\mathbf{j}'^\ell\mathbf{i}\,\mathbf{o}^\ell \to \mathbf{p}_2\mathbf{o}^\ell, \qquad \pi^r\mathbf{s}_2\,\mathbf{j}^\ell\mathbf{i} \to \pi^r\mathbf{s}_2.$$

We will follow this policy frequently from now on.

More easily typed, but less acceptable stylistically, are the following variants:

tomorrow, I will meet you,
in the rain, she had met him,
when it rained, he was met.

I would regard

I will meet you yesterday

well-formed syntactically, even if jarring semantically. Still, it admits a meaningful interpretation in a science fiction novel dealing with time travel.

Sometimes several interpretations are possible. For example, in the sentence

I promised to see her today

two possible interpretations depend on whether *today* modifies *see* or *promise*. This is reflected by two distinct calculations on types, as dictated by the placing of the left bracket:

$$\begin{array}{c} \textit{I promised to see her today,} \\ \pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell\mathbf{i}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{j}^\ell)([\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o}(\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \rightarrow \ \mathbf{s}_2$$

$$\begin{array}{c} \textit{I promised to see her today.} \\ \pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell[\mathbf{i}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o}(\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \rightarrow \ \mathbf{s}_2$$

We are now ready to examine Chomsky's [1956] famous example:

$$\begin{array}{c} \textit{colourless green ideas sleep furiously.} \\ (\mathbf{a}\mathbf{a}^r\,\mathbf{n}_2\,\mathbf{n}_2^\ell)(\mathbf{a}\mathbf{a}^r\,\mathbf{n}_2\,\mathbf{n}_2^\ell)\,\mathbf{n}_2\,(\pi_2\,\mathbf{s}_1\,\mathbf{j}^\ell[\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \rightarrow \ \mathbf{s}_1$$

Here *furiously* of type $\mathbf{i}^r\mathbf{i}$ can be replaced by

$$\begin{array}{c} \textit{under the cover} \\ (\mathbf{i}^r\mathbf{i}\,\mathbf{o}^\ell)(\bar{\mathbf{n}}\,\mathbf{n}_1^\ell)\mathbf{n}_1 \end{array} \quad \rightarrow \ \mathbf{i}^r\mathbf{i}$$

or by

$$\begin{array}{c} \textit{when it rains.} \\ (\mathbf{i}^r\mathbf{i}\,\mathbf{s}^\ell)\,\pi_3(\pi_3^r\,\mathbf{s}_1) \end{array} \quad \rightarrow \ \mathbf{i}^r\mathbf{i}.$$

The "genuine" adverbs considered by McCawley include the words *very*, already discussed, and *not*, still to be discussed. But most adverbs are obtained from adjectives by adding the suffix +*ly*. Presumably, there are general principles which explain why we have

furiously, *gladly*, *presumably*, ...

but not

**bigly*, **redly*, ...;

why we have *fearfully* but not **afraidly*. If such principles exist, I haven't discovered them.

What all adverbs or adverbials have in common is that they can modify something other than nouns, without changing its type. In our terminology, this means that they have a type xx^ℓ or x^rx with $x \neq \mathbf{n}_i$. Usually, but not always, x may be taken to be a basic type. We have already met *very* of type $\mathbf{a}\mathbf{a}^\ell$ and *today* of type $\mathbf{i}^r\mathbf{i}$.

McCawley considers the different positions an adverb may occupy and discusses six places in a sample sentence:

1, *the enemy* 2 *will* 3 *have* 4 *destroyed the village* 5, 6.

He points out that different adverbs may occupy different places. In particular, *completely* may occupy positions 4 and 5 and perhaps 3, *intentionally* may occupy positions 3, 4 and 5, and *probably* may occupy positions 1, 2, 3 and 6 and perhaps 4. In our terminology, adverbs occupying these six positions will have types

$$\mathbf{s}\mathbf{s}^\ell,\ \pi^r\mathbf{s}\mathbf{s}^\ell\pi,\ \mathbf{j}\mathbf{j}^\ell,\ \mathbf{p}_2\mathbf{p}_2^\ell,\ \mathbf{i}^r\mathbf{i},\ \mathbf{s}^r\mathbf{s}$$

respectively. Note that

$$\pi^r \mathbf{s}\mathbf{s}^\ell \pi = (\pi^r \mathbf{s})(\pi^r \mathbf{s})^\ell,$$

so *probably* in position 2 may be thought of as modifying the inflector C_{jk}.

We had decided that prepositional phrases should be regarded as adverbials of type $\mathbf{i}^r\mathbf{i}$, but they can also modify nouns from the right, so they should also be assigned types $\mathbf{n}_i^r\mathbf{n}_i$, when $i = 0, 1$ or 2, as in

water under the bridge	$(i = 0)$,
the cat in the hat	$(i = 1)$,
people without money	$(i = 2)$.

Thus prepositions can have type $\mathbf{n}_i^r\mathbf{n}_i\mathbf{o}^\ell$ in addition to $\mathbf{i}^r\mathbf{i}\mathbf{o}^\ell$.

Prepositional phrases governed by the preposition *by* are frequently used in passive constructions to refer to the agent, as in

he was seen by her.

In principle, there is no difference between this construction and that of

he was seen with her,

or when *by* is replaced by any other preposition, although there may be some difficulty in interpreting

he was seen for her,

for example.

Finally, let us point out that prepositional phrases can also take the place of predicative adjectives, as in

she will be without money.

Hence prepositions may provisionally be assigned the additional type $\bar{\mathbf{a}}\mathbf{o}^\ell$.

III. English Sentence Structure.

10. Negative and interrogative sentences.

The easiest way to negate a declarative sentence in English is to insert the adverb *not* in an appropriate place, namely after an auxiliary or modal verb. Thus, we assign to *not* the type xx^ℓ, where $x = \mathbf{j}, \mathbf{i}, \mathbf{p}_1, \mathbf{p}_2$ or $\overline{\mathbf{a}}$, as illustrated by the following examples:

$$\begin{array}{c} \textit{she would not come,} \\ \pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell)(\mathbf{j}\,\mathbf{j}^\ell)\,\mathbf{i} \end{array} \quad \to \quad \mathbf{s}_2$$

$$\begin{array}{c} \textit{she does not come,} \\ \pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{i}^\ell)\,\mathbf{i} \end{array} \quad \to \quad \mathbf{s}_1$$

$$\begin{array}{c} \textit{she is not coming,} \\ \pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{p}_1^\ell)(\mathbf{p}_1\,\mathbf{p}_1^\ell)\,\mathbf{p}_1 \end{array} \quad \to \quad \mathbf{s}_1$$

$$\begin{array}{c} \textit{she has not come,} \\ \pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{p}_2^\ell)\,\mathbf{p}_2 \end{array} \quad \to \quad \mathbf{s}_1$$

$$\begin{array}{c} \textit{she was not seen,} \\ \pi_3(\pi_3^r\,\mathbf{s}_2\,\mathbf{j}^\ell\mathbf{i}'\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{o}^\ell) \end{array} \quad \to \quad \mathbf{s}_2$$

$$\begin{array}{c} \textit{she is not very good,} \\ \pi_3(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)(\overline{\mathbf{a}}\,\overline{\mathbf{a}}^\ell)(\mathbf{a}\,\mathbf{a}^\ell)\mathbf{a} \end{array} \quad \to \quad \mathbf{s}_1$$

$$\begin{array}{c} \textit{she is not too good,} \\ \pi_3(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)(\overline{\mathbf{a}}\,\overline{\mathbf{a}}^\ell)(\overline{\mathbf{a}}\,\overline{\mathbf{a}}^\ell)\mathbf{a} \end{array} \quad \to \quad \mathbf{s}_1$$

but

$$\begin{array}{c} {}^*\textit{she is very not good.} \\ \pi_3(\pi_3^r\mathbf{s}_1\overline{\mathbf{a}}^\ell)(\mathbf{a}\mathbf{a}^\ell)(\overline{\mathbf{a}}\overline{\mathbf{a}}^\ell)\mathbf{a} \end{array} \quad \not\to \quad \mathbf{s}_1$$

The negated auxiliary and modal verb forms are frequently contracted to verb forms of the same type. Thus we have

$$\begin{array}{ccc} \textit{will not} & \to & \textit{won't} \\ (\pi^r\mathbf{s}_1\,\mathbf{j}^\ell)(\mathbf{j}\,\mathbf{j}^\ell) & \to & \pi^r\mathbf{s}_1\mathbf{j}^\ell \end{array},$$

$$\begin{array}{ccc} \textit{does not} & \to & \textit{doesn't} \\ (\pi_3^r\mathbf{s}_1\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{i}^\ell) & \to & \pi_3^r\mathbf{s}_1\mathbf{i}^\ell \end{array},$$

$$\begin{array}{ccc} \textit{is not} & \to & \textit{isn't} \\ (\pi_3^r\mathbf{s}_1\,x^\ell)(x\,x^\ell) & \to & \pi_3^r\mathbf{s}_1 x^\ell \end{array},$$

when $x = \mathbf{p}_2$ or $\overline{\mathbf{a}}$,

$$\begin{array}{ccc} \textit{is not} & \to & \textit{isn't} \\ (\pi_3^r\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{p}_2^\ell) & \to & \pi_3^r\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell \end{array}.$$

We can now generate yes-or-no questions by assigning new types to finite forms of modal and auxiliary verbs, both positive and negative. We introduce the basic types

$\mathbf{q}$ = yes-or-no question,

$\mathbf{q}_1$ = yes-or-no question in present tense,

$\mathbf{q}_2$ = yes-or-no question in past tense,

and postulate

$$\mathbf{q}_j \to \mathbf{q} \qquad (j = 1, 2)\ .$$

In order not to overload the dictionary, it will be convenient to introduce a metarule.

Metarule 1. If the finite verb form of a modal or auxiliary verb has the type $\pi_k^r \mathbf{s}_j x^\ell$, with $x = \mathbf{j}, \mathbf{i}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_2\hat{\mathbf{o}}^\ell$ or $\overline{\mathbf{a}}$ in a declarative sentence, then it has type $\mathbf{q}_j x^\ell \pi_k^\ell$ in a question, and similarly when the subscript j or k is omitted.

In Elizabethan English or in modern German, such a metarule would also apply to main verbs, as in *comes she*. Here are some examples illustrating the modern metarule with $x = \mathbf{i}, \mathbf{p}_1$ and $\mathbf{p}_2$:

does she not see me ?

$$(\mathbf{q}_1\, \mathbf{i}^\ell\, \pi_3^\ell)\, \pi_3\, (\mathbf{i}\, \mathbf{i}^\ell)(\mathbf{i}\, \mathbf{o}^\ell)\mathbf{o} \quad \to \quad \mathbf{q}_1$$

doesn't she see me ?

$$(\mathbf{q}_1\, \mathbf{i}^\ell\, \pi_3^\ell)\, \pi_3\, (\mathbf{i}\, \mathbf{o}^\ell)\, \mathbf{o} \quad \to \quad \mathbf{q}_1$$

and even

doesn't she not see me ?

We also have the passive question

am I being seen − ?

$$(\mathbf{q}_1\, \mathbf{p}_1^\ell\, \pi_1^\ell)\pi_1\, (\mathbf{p}_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{p}_2^\ell)(\mathbf{p}_2\, \mathbf{o}^\ell) \quad \to \quad \mathbf{q}_1$$

We use Metarule 1 with $x = \mathbf{p}_2\hat{\mathbf{o}}^\ell$ and $x^\ell = \hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell$ to obtain

was I seen − ?

$$(\mathbf{q}_2\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{p}_2^\ell\, \pi_1^\ell)\, \pi_1\, (\mathbf{p}_2\, \mathbf{o}^\ell) \quad \to \quad \mathbf{q}_2$$

11. Direct Wh-questions.

In addition to questions that can be answered with *yes* or *no*, there are those asking for more specific information. These are introduced by question words, almost all starting with *wh*. The easiest way is to place the question word in situ (and raising one's voice), as in

she went where ?

$$\pi_3\,(\pi^r\,\mathbf{s}_2)(\mathbf{s}^r\,\overline{\mathbf{q}}) \quad\to\quad \overline{\mathbf{q}}$$

or

she saw whom ?

$$\pi_3(\pi^r\,\mathbf{s}_1\mathbf{o}^\ell)(\mathbf{o}\mathbf{s}^r\,\overline{\mathbf{q}}) \quad\to\quad \overline{\mathbf{q}}$$

thus assigning the types $\mathbf{s}^r\overline{\mathbf{q}}$ and $\mathbf{o}\mathbf{s}^r\overline{\mathbf{q}}$ to *where* and *whom* respectively. We have introduced a new basic type

$$\overline{\mathbf{q}} = \text{ question}$$

and we postulate

$$\mathbf{q} \to \overline{\mathbf{q}} \not\to \mathbf{q}.$$

Another simple way to handle Wh-questions starting with *where, when, why* or *how* (the sole exception), asking for place, time, reason or manner respectively, is to assign to them the type $\overline{\mathbf{q}}\mathbf{q}^\ell$. Thus we have

where did she go ?

$$(\overline{\mathbf{q}}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_3\,\mathbf{i} \quad\to\quad \overline{\mathbf{q}}$$

but

**why where* did she go?

$$(\overline{\mathbf{q}}\mathbf{q}^\ell)\ (\overline{\mathbf{q}}\mathbf{q}^\ell)$$

On second thought, one feels that such questions look for an answer with an adverbial phrase of type $\mathbf{i}^r\mathbf{i}$, such as

she went to Rome.

$$\pi_3\,(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell\,[\mathbf{i})\,(\mathbf{i}^r\,\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \quad\to\quad \mathbf{s}_2.$$

It might seem more honest to assign to the adverbial question word the type $\overline{\mathbf{q}}\mathbf{i}^\ell\mathbf{i}^{\ell\ell}\mathbf{q}^\ell$, so that we obtain

where did she go – ?

$$(\overline{\mathbf{q}}\,\mathbf{i}^\ell\,\mathbf{i}^{\ell\ell}\,\mathbf{q}^\ell)\,(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\,\pi_3\ \ \mathbf{i} \quad\to\quad \overline{\mathbf{q}}$$

Chomskyan linguists might say that the question word has been moved from the end to the beginning of the question, and I have placed a dash at the end to represent what they call a trace. Traces are not really needed for our approach, their place being taken by double adjoints.

Next, let us turn to questions asking for the object of a transitive verb or verb phrase. These are usually introduced by *whom* or *what*, depending on whether the expected answer refers to a person or something other than a person.

whom / what did she see – ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)\ (\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\ \pi_3\ (\mathbf{i}\,\mathbf{o}^\ell) \quad\to\quad \overline{\mathbf{q}}$$

Again, the dash at the end represents a Chomskyan trace, put in for comparison with the literature.

When I first proposed this approach in 1998, Michael Moortgat challenged me by asking: but what if there is an adverb such as *today* at the end of the question? My present answer is to introduce a metarule:

Metarule 2. In the type of a transitive verb or in a preposition (= transitive adverb), $\mathbf{o}^\ell$ may be replaced by $\hat{\mathbf{o}}^\ell\mathbf{i}^\ell\mathbf{i}$.

We can then analyze

whom did she see today ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)\ (\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\ \pi_3\ (\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{i}^\ell\,\mathbf{i})\ (\mathbf{i}^r\,\mathbf{i}) \quad \to \quad \overline{\mathbf{q}}$$

We are finally in the position to justify having introduced the hat on $\hat{\mathbf{o}}^\ell$ and $\hat{\mathbf{o}}^{\ell\ell}$: it guards against

**did she see today him,*

$$(\mathbf{q}_2\mathbf{i}^\ell\pi^\ell)\pi_3(\mathbf{i}\hat{\mathbf{o}}^\ell\,\mathbf{i}^\ell\,\mathbf{i})(\mathbf{i}^r\,\mathbf{i})\mathbf{o}$$

as long as $\mathbf{o} \not\to \hat{\mathbf{o}}$.

In the above analysis, the adverb *today* can be replaced by any prepositional phrase or subordinate clause of type $\mathbf{i}^r\mathbf{i}$, such as *at the conference* or *when it rained.* Also the transitive verb *see* can be replaced by a verb cum preposition, as in

whom did she go with yesterday ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_3\,\mathbf{i}\,(\mathbf{i}^r\,\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{i}^\ell\,\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \quad \to \quad \overline{\mathbf{q}}$$

If we wish to be pedantic, we might say instead

with whom did she go yesterday ?

Then *with whom* should have the same type as *where*, either $\overline{\mathbf{q}}\mathbf{q}^\ell$ or $\overline{\mathbf{q}}\mathbf{i}^\ell\mathbf{i}^{\ell\ell}\mathbf{q}^\ell$, let me say the former. Keeping the type of *whom* unchanged, the type of *with* will involve a triple left adjoint:

with whom

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell\ell}\,\overline{\mathbf{q}}^\ell)(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell) \quad \to \quad \overline{\mathbf{q}}\mathbf{q}^\ell$$

The object being asked for could be a noun modified by a determiner, so *whom* or *what* might be replaced by *whose brother* or *which books.* Then *whose* and *which* will have type $\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell\mathbf{n}_i^\ell$, with $i = 0, 1$ or 2. Here is just one example, applying Metarule 2 to the type $\mathbf{i}^r\mathbf{i}\mathbf{o}^\ell$ of *in*:

whose bed did she sleep in yesterday ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell\,\mathbf{n}_1^\ell)\,\mathbf{n}_1(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_3\,\mathbf{i}\,(\mathbf{i}^r\,\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{i}^\ell\,\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \quad \to \quad \overline{\mathbf{q}}$$

Finally, let us look at questions asking for the subject of a sentence. The question word is *who* or *what*, depending on whether the expected answer refers to a human or not. We may handle these words in situ by assigning to them the type $\overline{\mathbf{q}}\mathbf{s}^\ell\pi_3$, as in

$$\begin{array}{c} \textit{who came yesterday ?} \\ (\overline{\mathbf{q}}\,\mathbf{s}^\ell\,\pi_3)(\pi_3^r\,\mathbf{s}_2\,\mathbf{j}^\ell[\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \quad \overline{\mathbf{q}}$$

More honestly, such questions can be derived from "pseudo-questions" (as I believe is implicit in Chomsky [1958]) such as

$$\begin{array}{c} {}^*\textit{came yesterday she ?} \\ (\mathbf{q}_2\hat{\pi}_3^\ell\,\mathbf{j}^\ell\,[\mathbf{i})(\mathbf{i}^r\,\mathbf{i})\pi_3 \end{array} \qquad \not\to \quad \mathbf{q}_2$$

where the hat guards against the contraction $\hat{\pi}_3^\ell\pi_3 \to 1$, as long as

$$\hat{\pi}_3 \;\to\; \pi_3 \;\not\to\; \hat{\pi}_3.$$

Thus, the type of *who/what* should be $\overline{\mathbf{q}}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell$, as in

$$\begin{array}{c} \textit{who came yesterday ?} \\ (\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\pi_3^\ell\,\mathbf{j}^\ell[\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \quad \overline{\mathbf{q}}$$

since $\pi_3^\ell \to \hat{\pi}_3^\ell$ follows from $\hat{\pi}_3 \to \pi_3$. To justify the above procedure, we adopt

Metarule 3. In direct questions derived from pseudo-questions, the inflector C_{jk}, normally of type $\pi_k^r\mathbf{s}_j\mathbf{j}^\ell$, may be assigned the new type $\mathbf{q}_j\hat{\pi}_k^\ell\mathbf{j}^\ell$, and similarly when the subscript j or k is omitted. The same goes for modals.

Although we have only used this metarule with $k = 3$, other uses may come up later.

In the same vein, from the pseudo-question

$$\begin{array}{c} {}^*\textit{was seen yesterday she ?} \\ (\mathbf{q}_2\hat{\pi}_3^\ell\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\hat{\mathbf{o}}^\ell\,\mathbf{j}^\ell\,[\mathbf{j})(\mathbf{j}^r\,\mathbf{j})\pi_3 \end{array} \quad \not\to \quad \mathbf{q}_2$$

we can derive the passive Wh-question

$$\begin{array}{c} \textit{who was seen yesterday ?} \\ (\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\hat{\pi}_3^\ell\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\hat{\mathbf{o}}^\ell\,\mathbf{j}^\ell\,[\mathbf{j})(\mathbf{j}^r\,\mathbf{j}) \end{array} \quad \to \quad \overline{\mathbf{q}}$$

When it comes to compound question words such as *whose friends* or *which wine*, we are led to assign to *whose/which* the type $\overline{\mathbf{q}}\hat{\pi}_k^{\ell\ell}\mathbf{q}^\ell\mathbf{n}_i^\ell$, where $(k,i) = (3,0),\ (3,1)$ or $(2,2)$, as in

$$\begin{array}{c} \textit{which wine will please you ?} \\ (\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell\,\mathbf{n}_0^\ell)\,\mathbf{n}_0\,(\mathbf{q}_1\,\hat{\pi}_3^\ell\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \end{array} \quad \to \quad \overline{\mathbf{q}}$$

or

$$\begin{array}{c} \textit{whose books interest you ?} \\ (\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell\,\mathbf{n}_2^\ell)\,\mathbf{n}_2\,(\mathbf{q}_1\,\hat{\pi}_2^\ell\,\mathbf{j}^\ell\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \end{array} \quad \to \quad \overline{\mathbf{q}}$$

The reader may have noticed the absence of the dash representing a trace in the last few examples. This is because I would have placed it at the end of the questions, whereas mainstream linguists might have inserted the trace after the subject question word.

12. Indirect questions.

Whereas direct Wh-questions of type $\overline{\mathbf{q}}$ are formed from yes-or-no questions of type $\mathbf{q}$, indirect questions are formed from declarative sentences of type $\mathbf{s}$. We will adopt the basic type

$$\mathbf{t} = \text{ indirect question.}$$

Indirect questions analogous to yes-or-no direct ones are introduced by *whether/if* of type $\mathbf{t}\mathbf{s}^\ell$, as illustrated by the following examples:

I wonder whether she is coming.

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{j}^\ell\mathbf{i}\,\mathbf{t}^\ell)(\mathbf{t}\,\mathbf{s}^\ell)\,\pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{p}_1^\ell)\mathbf{p}_1 \quad\to\quad \mathbf{s}_1$$

does she know if I expect her ?

$$(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_3^\ell)\pi_3(\mathbf{i}\,\mathbf{t}^\ell)(\mathbf{t}\,\mathbf{s}^\ell)\,\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{o}^\ell)\mathbf{o} \quad\to\quad \mathbf{q}_1$$

The verbs *wonder* and *know* may require an indirect question as a complement, hence they have been assigned the type $\mathbf{i}\mathbf{t}^\ell$. More generally, we adopt the following

Metarule 4. Indirect questions are formed by replacing the type of the direct question word $\overline{\mathbf{q}}\cdots\mathbf{q}^\ell\cdots$ by $\mathbf{t}\cdots\mathbf{s}^\ell\cdots$.

For example, *when* and *where* of type $\overline{\mathbf{q}}\mathbf{i}^\ell\mathbf{i}^{\ell\ell}\mathbf{q}^\ell$ may be assigned the new type $\mathbf{t}\mathbf{i}^\ell\mathbf{i}^{\ell\ell}\mathbf{s}^\ell$. Thus we have

she asked me when I would come.

$$\pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell\mathbf{i}\,\mathbf{t}^\ell\,\mathbf{o}^\ell)\mathbf{o}(\mathbf{t}\,\mathbf{i}^\ell\,\mathbf{i}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell)\mathbf{i} \quad\to\quad \mathbf{s}_2$$

Here *ask* has two complements: a direct object of type $\mathbf{o}$ and an indirect question of type $\mathbf{t}$. Note that $\mathbf{j}^\ell \to \mathbf{i}^\ell$ since $\mathbf{i} \to \mathbf{j}$, hence $\mathbf{i}^{\ell\ell}\mathbf{j}^\ell \to \mathbf{i}^{\ell\ell}\mathbf{i}^\ell \to 1$.

Here are some other examples of indirect questions:

I wonder where I will go tomorrow,

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\mathbf{i}^\ell\,\mathbf{i}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_1\,\mathbf{j}^\ell)[\,\mathbf{i}(\mathbf{i}^r\,\mathbf{i}) \quad\to\quad \mathbf{s}_1$$

she doesn't know what I like,

$$\pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{o}^\ell) \quad\to\quad \mathbf{s}_1$$

I wonder which books I should read.

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell\,\mathbf{n}_2^\ell)\mathbf{n}_2\,\pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) \quad\to\quad \mathbf{s}_1$$

The indirect version of *who/what* has to be of type $\mathbf{t}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell$, as in

I ask who is coming tomorrow,

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\hat{\pi}_3^\ell\,\mathbf{p}_1^\ell)(\mathbf{p}_1\,\mathbf{i}'^\ell\,[\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \quad\to\quad \mathbf{s}_1$$

and the subject version of *whose/which* in indirect questions should have type $\mathbf{t}\hat{\pi}_k^{\ell\ell}\mathbf{q}^\ell\mathbf{n}_i$ subject to $(k,i) = (3,0),\ (3,1)$ or $(2,2)$ as in

I know which girl likes you.

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell\,\mathbf{n}_1)\mathbf{n}_1(\mathbf{q}_1\,\hat{\pi}_3^\ell\,\mathbf{o}^\ell)\mathbf{o} \quad \rightarrow \quad \mathbf{s}_1$$

However, the simpler alternative type $\mathbf{t}\mathbf{s}^\ell\pi_3$ will also do for *who/what*, as in

I ask who is coming,

and *whose/which* will then have type $\mathbf{t}\mathbf{s}^\ell\pi_k\mathbf{n}_i^\ell$.

Question words other than those asking for the subject, occurring in indirect questions, can also govern complete infinitives with *to*:

I wonder where to go,

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{i}^\ell)\mathbf{i} \quad \rightarrow \quad \mathbf{s}_1$$

I wonder whom to see tomorrow.

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{i}^\ell)(\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{j}^\ell\,[\mathbf{i})(\mathbf{i}^r\,\mathbf{i}) \quad \rightarrow \quad \mathbf{s}_1$$

Here *whether, where, when* and *how* can also receive the type $\mathbf{t}\bar{\mathbf{j}}^\ell$ and *whom/what* the type $\mathbf{t}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{j}}^\ell$. Finally, consider:

I know which books to read.

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^\ell\,\mathbf{n}_2^\ell)\mathbf{n}_2(\bar{\mathbf{j}}\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) \quad \rightarrow \quad \mathbf{s}_1$$

Again, *which books* should have the same type as *what*, and so *which* must have the type $\mathbf{t}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{j}}^\ell\mathbf{n}_i^\ell$, here with $i = 2$.

13. Relative clauses.

Most of the question words discussed in the last chapter can also serve as relative pronouns, except that *what* must be replaced by *which* and the adverbial question words *where, when, why* and *how* are subject to a semantic constraint, inasmuch as they should only attach to nouns that refer to place, time, reason or manner respectively.

It is customary to distinguish between *restrictive* relative clauses modifying nouns and *non-restrictive* ones modifying complete noun phrases. Both are introduced by the same relative pronouns; but, in the restrictive case, the subject or object relative pronoun can be replaced by *that*. Orthography requires that a non-restrictive relative clause is preceded by a comma. McCawley also considers some other kinds, e.g. infinitival and free relative clauses. These we will only touch upon briefly at the end of this chapter.

Here are some examples of restrictive relative clauses with an object relative pronoun:

the wine which/[that] I drank − ,
the man whom/[that] I saw − ,
the apple which/[that] I ate − .

These relative clauses modify nouns of type $\mathbf{n}_i$, where $i = 0, 1$ or 2 respectively. Therefore, *whom/which/[that]* should all be assigned type $\mathbf{n}_i^r \mathbf{n}_i \hat{\mathbf{o}}^{\ell\ell} \mathbf{s}^\ell$. For example,

$$\begin{array}{c} \textit{the man whom I saw} \; - \; . \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^\ell)\, [\mathbf{n}_1\, (\mathbf{n}_1^r\, \mathbf{n}_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^\ell)\, \pi_1 (\pi^r\, \mathbf{s}_2\, \mathbf{o}^\ell) \quad \rightarrow \quad \overline{\mathbf{n}}_1 \end{array}$$

Note that in these examples, the object relative pronoun *that* can be omitted, so we obtain

the wine ∅ *I drank*,
the man ∅ *I saw*,
the apple ∅ *I ate*,

and we should attach the type $\mathbf{n}_i^r \mathbf{n}_i \hat{\mathbf{o}}^{\ell\ell} \mathbf{s}^\ell$ to the empty string, here denoted by ∅. This may be easy for the speaker, but how does the hearer know where to interpolate a typed empty string? Elsewhere [2007], I have suggested that a noun of type $\mathbf{n}_i$ may have an optional invisible ending of type $\hat{\mathbf{o}}^{\ell\ell} \mathbf{s}^\ell$, but this proposal brings up some other complications, which I will not pursue here.

Anyway, ellipsis of object relative pronouns may render apparently well-formed sentences quite incomprehensible. My favourite example is

police police police police police ,

which can be interpreted to mean

police [that] police control control police

or

police control police [that] police control.

Readers interested in mathematical puzzles may amuse themselves by calculating the number of interpretations of *police* repeated $2n + 1$ times ($n \geq 1$) as a declarative sentence.

Now look at the subject relative pronouns *who* / *which* / *that*, as in

the wine which tastes good,
the man who saw me,
the apples that were eaten.

The easiest way is to assign to them the type $\mathbf{n}_i^r \mathbf{n}_i \mathbf{s}^\ell \pi_k$, where $(i,k) = (0,3),(1,3)$ or $(2,2)$. Thus we have the noun phrases

the wine which tastes good,
$$(\overline{\mathbf{n}}_0\, \mathbf{n}_0^\ell)\, [\mathbf{n}_0 (\mathbf{n}_0^r\, \mathbf{n}_0\, \mathbf{s}^\ell\, \pi_3)(\pi_3^r\, \mathbf{s}_1\, \mathbf{a}^\ell)\mathbf{a} \;\to\; \overline{\mathbf{n}}_0$$

the men who saw me.
$$(\overline{\mathbf{n}}_2\, \mathbf{n}_2^\ell)\, [\mathbf{n}_2 (\mathbf{n}_2^r\, \mathbf{n}_2\, \mathbf{s}^\ell\, \pi_2)(\pi^r\, \mathbf{s}_2\, \mathbf{o}^\ell)\mathbf{o} \;\to\; \overline{\mathbf{n}}_2$$

The compound relative pronominal expression *whose books* should have the same type as *which*, namely $\mathbf{n}_i^r \mathbf{n}_i \hat{\mathbf{o}}^{\ell\ell} \mathbf{s}^\ell$ when modifying an object and $\mathbf{n}_i^r \mathbf{n}_i \mathbf{s}^\ell \pi_k$ when modifying a subject. Thus we have

the author whose books I have read −
$$(\overline{\mathbf{n}}_1\, \mathbf{n}_1^\ell)\, [\mathbf{n}_1\, (\mathbf{n}_1^r\, \mathbf{n}_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^\ell\, \mathbf{n}_2^\ell)\mathbf{n}_2\, \pi_1 (\pi_1^r\, \mathbf{s}_1\, \mathbf{p}_2^\ell)(\mathbf{p}_2\, \mathbf{o}^\ell) \;\to\; \overline{\mathbf{n}}_1$$

the author whose books entertain me.
$$(\overline{\mathbf{n}}_1\, \mathbf{n}_1^\ell)\, [\mathbf{n}_1\, (\mathbf{n}_1^r\, \mathbf{n}_1\, \mathbf{s}^\ell\, \pi_2\, \mathbf{n}_2^\ell)\mathbf{n}_2 (\pi_2^r\, \mathbf{s}_1\, \mathbf{o}^\ell)\mathbf{o} \;\to\; \overline{\mathbf{n}}_1$$

Now let us consider some adverbial relative clauses:

the question whether/if she snores,
the city where I was born,
the year when/[that] we met,
the reason why/[that] she drinks,
the way how/[that] he walks.

The relative pronouns here all have type

$$\mathbf{n}_i^r \mathbf{n}_i\, \mathbf{i}^\ell \mathbf{i}\, \mathbf{s}^\ell \;\to\; \mathbf{n}_i^r \mathbf{n}_i \mathbf{s}^\ell,$$

as in

the year when we met in Paris.
$$(\overline{\mathbf{n}}_1\, \mathbf{n}_1^\ell)\, [\mathbf{n}_1\, (\mathbf{n}_1^r\, \mathbf{n}_1\, \mathbf{s}^\ell)\, \pi_2 (\pi^r\, \mathbf{s}_2\, \mathbf{j}^\ell\, [\mathbf{i})(\mathbf{i}^r\, \mathbf{i}\, \mathbf{o}^\ell)\mathbf{o} \;\to\; \overline{\mathbf{n}}_1$$

On purely semantical grounds, we are not likely to say

the year why we met,
the reason where she drinks,

but I am willing to consider these expressions syntactically acceptable. Note that the alternative *that* may be omitted altogether, raising again the question whether its type should be attached to the empty string or to an invisible ending of the modified noun.

Let us take a brief look at relative pronouns governed by prepositions, such as

$$\begin{array}{c} \textit{the house which/[that] he will go into } - \ . \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\, [\mathbf{n}_1\, (\mathbf{n}_1^{r}\, \mathbf{n}_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^{\ell})\, \pi_3(\pi^{r}\, \mathbf{s}_1\, \mathbf{j}^{\ell})\, [\mathbf{i}(\mathbf{i}^{r}\, \mathbf{i}\, \mathbf{o}^{\ell}) \quad \rightarrow \ \overline{\mathbf{n}}_1 \end{array}$$

Alternatively, we may place the preposition before the relative pronoun, so that *into which* will have the same type as *where* in

$$\begin{array}{c} \textit{the house where he will go.} \\ (\overline{\mathbf{n}}_1\, \mathbf{n}_1^{\ell})\, [\mathbf{n}_1\, (\mathbf{n}_1^{r}\, \mathbf{n}_1\, \mathbf{s}^{\ell})\, \pi_3(\pi^{r}\, \mathbf{s}_1\, \mathbf{j}^{\ell})\mathbf{i} \quad \rightarrow \ \mathbf{s}_1 \end{array}$$

We still have to decide whether to change the type of *into* or that of *which*. One solution involving triple adjoints leaves the type of *which* unchanged:

$$\begin{array}{c} \textit{into} \quad \textit{which .} \\ (\mathbf{n}_1^{r}\mathbf{n}_1\, \hat{\mathbf{o}}^{\ell\ell\ell}\, \mathbf{n}_1^{\ell}\, \mathbf{n}_1)(\mathbf{n}_1^{r}\, \mathbf{n}_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^{\ell}) \quad \rightarrow \ \mathbf{n}_1^{r}\mathbf{n}_1\mathbf{s}^{\ell} \end{array}$$

Non-restrictive relative pronouns are handled differently. For example, the object relative pronoun *whom* can have type $\pi_k^r\pi_k\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}$ when modifying the subject, or $\mathbf{o}^r\mathbf{o}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}$ when modifying the object, as in

$$\begin{array}{c} \textit{I, whom she liked} \\ \pi_1\, (\pi_1^{r}\, \pi_1\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^{\ell})\, \pi_3(\pi^{r}\, \mathbf{s}_2\, \mathbf{o}^{\ell}) \quad \rightarrow \ \pi_1 \end{array}$$

$$\begin{array}{c} \textit{me, whom she liked} \\ \mathbf{o}\, (\mathbf{o}^{r}\, \mathbf{o}\, \hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^{\ell})\, \pi_3(\pi^{r}\, \mathbf{s}_2\, \mathbf{o}^{\ell}) \quad \rightarrow \ \mathbf{o}. \end{array}$$

Note that the object relative pronoun *that* is never omitted in non-restrictive relative clauses. (Indeed, there is no immediately preceding noun to which the invisible ending of type $\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}$ could be attached.) The subject relative pronoun *who* can have type $\pi_k^r\pi_k\mathbf{s}^{\ell}\pi_k$ when modifying the subject, as in

$$\begin{array}{c} \textit{I, who like her.} \\ \pi_1\, (\pi_1^{r}\, \pi_1\, \mathbf{s}^{\ell}\, \pi_1)\ (\pi_1^{r}\, \mathbf{s}_1\, \mathbf{o}^{\ell})\mathbf{o} \quad \rightarrow \ \pi_1 \end{array}$$

There is a small problem when *who* modifies an object pronoun, as in

them, who like her.

Once we assign the type $\mathbf{o}$ to *them*, we no longer remember that *like* should be in the plural. Let us introduce the inflector

$$Acc = \text{accusative}$$

of type $\mathbf{o}\pi^{\ell}$, so that

$$\begin{array}{c} \textit{Acc they} \ \rightarrow \ \textit{them,} \\ (\mathbf{o}\, \pi^{\ell})\pi_2 \ \rightarrow \ \mathbf{o} \end{array}$$

where the upper arrow belongs to the generative morphology and the lower arrow denotes the partial order of the pregroup. Thus we may analyze

$$\begin{array}{c} \textit{them, who like her .} \\ (\mathbf{o}\,\pi^\ell)[\pi_2\,(\pi_2^r\,\pi_2\,\mathbf{s}^\ell\,\pi_2)\;(\pi_2^r\,\mathbf{s}_1\,\mathbf{o}^\ell)\mathbf{o} \end{array} \;\rightarrow\; \mathbf{o}$$

The same device could be exploited to reduce the two above mentioned types of *whom* to one.

Adverbial relative clauses can also modify complete noun phrases, hence may be non-restrictive:

Montreal, where my sons were born,

the year 2005, when I visited Rome.

Here *when* and *where* have type

$$\pi_i^r\pi_i\mathbf{i}^\ell\mathbf{i}\mathbf{s}^\ell \;\rightarrow\; \pi_i^r\pi_i\mathbf{s}^\ell.$$

Let me briefly mention infinitival relative clauses:

$$\begin{array}{c} \textit{the person whom to consult} - , \\ (\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell)\,[\mathbf{n}_1\,(\mathbf{n}_1^r\,\mathbf{n}_1\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) \end{array} \;\rightarrow \overline{\mathbf{n}}_1$$

$$\begin{array}{c} \textit{the person with whom to speak.} \\ (\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell)\,[\mathbf{n}_1\,(\mathbf{n}_1^r\,\mathbf{n}_1\,\hat{\mathbf{o}}^{\ell\ell\ell}\,\mathbf{n}_1^\ell\,\mathbf{n}_1)(\mathbf{n}_1^r\,\mathbf{n}_1\,\hat{\mathbf{o}}^{\ell\ell}\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{i}^\ell)\mathbf{i} \end{array} \;\rightarrow \overline{\mathbf{n}}_1$$

Looking at the last three chapters, we see that *whom* (which I prefer to distinguish from *who*) has many different types. It may have other types as yet when referring to an indirect object, which we have not yet considered. In Standard German, historically related to English, the corresponding question words take two distinct forms:

wen (asking for the direct object),

wem (asking for the indirect object),

while the relative pronoun *whom* translates into German

den (masculine accusative),
die (feminine or plural accusative),
das (neuter accusative),
dem (masculine or neuter dative),
der (feminine dative),
denen (plural dative).

IV. Other verbal complements.

14. Doubly transitive verbs.

A number of verbs require or admit two complements, the most familiar being those that are normally accompanied by a direct as well as an indirect object, as illustrated by the infinitival phrases

give her a book,
show her a picture,
tell her a story.

Traditionally, *her* in these examples is viewed as an indirect object, say of type $\mathbf{o}'$, originally expressed in the now obsolete dative case, which has become indistinguishable from the accusative case in modern English. Thus, any noun phrase of type $\mathbf{o}$ also has type $\mathbf{o}'$, so we postulate

$$\mathbf{o} \to \mathbf{o}'.$$

English not being my first language, I must confess to a tendency of generating sentences not acceptable to native speakers, and I may have done so in earlier publications. The way I see it now, the direct object of a doubly transitive verb like *give* should not be a pronoun.

Let us introduce a new type $\overline{\mathbf{n}}$:

$$\overline{\mathbf{n}} = \text{complete noun phrase},$$

subject to

$$\overline{\mathbf{n}}_i \to \overline{\mathbf{n}} \to \mathbf{o} \not\to \overline{\mathbf{n}} \qquad (i = 0, 1, 2),$$

and assign to doubly transitive verbs the type $\mathbf{i}\overline{\mathbf{n}}^\ell\mathbf{o}'^\ell$. This will allow

$$\begin{array}{c} \textit{give her flowers,} \\ (\mathbf{i}\,\overline{\mathbf{n}}^\ell\,\mathbf{o}'^\ell)\mathbf{o}\;\overline{\mathbf{n}}_2 \end{array} \quad \to \mathbf{i}$$

but not

$$\begin{array}{c} {}^*\textit{give her them,} \\ (\mathbf{i}\overline{\mathbf{n}}^\ell\,\mathbf{o}'^\ell)\mathbf{o}\;\mathbf{o} \end{array} \quad \to \mathbf{i}$$

using $\mathbf{o} \to \mathbf{o}'$ and $\overline{\mathbf{n}}_2 \to \overline{\mathbf{n}}$, but $\mathbf{o} \not\to \overline{\mathbf{n}}$. In both examples, the pronoun *her* may be replaced by the noun phrase *the girl* of type $(\overline{\mathbf{n}}_1\mathbf{n}_1^\ell)\mathbf{n}_1 \to \overline{\mathbf{n}}_1 \to \mathbf{o}$.

In British novels, I have come across the different word order

give them her,

but the present type assignment will not account for this. I believe most people prefer

$$\begin{array}{c} \textit{give them to her,} \\ (\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o}\,(\mathbf{i}^r\,\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \end{array} \quad \to \mathbf{i}$$

where *give* is treated as a simply transitive verb, supplemented by a prepositional phrase. In principle, this is not very different from

give them for her

or even

give them over the counter.

In my opinion, any verb phrase can be accompanied by a prepositional phrase, even if a semantic interpretation is not always easy to find, as in

give them at her.

It has been observed by many people that messages on Chinese fortune cookies can always be supplemented by

in bed

or

under the blanket.

Here are two examples that I might have admitted, had I not been told otherwise by native English speakers:

$$\begin{array}{c} {}^{*}\textit{whom did he give flowers?} \\ (\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}_2^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_3\,(\mathbf{i}\,\overline{\mathbf{n}}^{\ell}\,\mathbf{o}'^{\ell})\overline{\mathbf{n}}_2 \quad \not\to \quad \overline{\mathbf{q}} \end{array}$$

$$\begin{array}{c} {}^{*}\textit{they were given her.} \\ \pi_2\ (\pi_2^{r}\,\mathbf{s}_2\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})(\mathbf{p}_2\,\overline{\mathbf{n}}^{\ell}\,\mathbf{o}'^{\ell})\mathbf{o} \quad \not\to \quad \mathbf{s}_2 \end{array}$$

To avoid these sentences, we must insist that $\overline{\mathbf{n}}^{\ell} \not\to \hat{\mathbf{o}}^{\ell}$, that is, $\hat{\mathbf{o}} \not\to \overline{\mathbf{n}}$.

The doubly transitive *give* may be viewed as a causative of the simply transitive *have*, so that *give her flowers* can be understood as meaning

let her have flowers,

even when *have* is used metaphorically; so we can say

give her a cold.

Other doubly transitive verbs, such as *show* and *tell*, may be similarly construed as causatives of *see* and *hear* respectively. Thus,

show her the picture = *let her see the picture*

and

tell her a story = *let her hear a story.*

These verb phrases may also be transformed into

show the pictures to her, *tell the story to her*

respectively.

When we replace doubly transitive verbs by simply transitive ones, accompanied by a prepositional phrase with *to*, questions and passives may be handled as follows:

$$\begin{array}{c} \textit{did he give books to her ?} \\ (\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_3(\,[\mathbf{i}\,\mathbf{o}^{\ell})n_2(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\,\mathbf{o} \quad \to \quad \mathbf{q}_2 \end{array}$$

(using $\mathbf{n}_2 \to \overline{\mathbf{n}}_2 \to \mathbf{o}$),

what did he give to her ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_3(\mathbf{i}\,\hat{\mathbf{o}}^{\ell}\,\mathbf{i}^{\ell}\,[\mathbf{i})(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o} \quad \to \quad \overline{\mathbf{q}}$$

(using Metarule 2 for the type of *give*),

who gave flowers to her ?

$$(\overline{\mathbf{q}}\,\hat{\pi}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\hat{\pi}^{\ell}\,\mathbf{j}^{\ell}\,[\mathbf{i}\,\mathbf{o}^{\ell})n_2(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\,\mathbf{o} \quad \to \quad \overline{\mathbf{q}}$$

(using Metarule 3 for the inflector of *gave*),

flowers were given to her,

$$\mathbf{n}_2\,(\pi_2^{r}\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})(\mathbf{p}_2\,\hat{\mathbf{o}}^{\ell}\,\mathbf{i}^{\ell}\,[\mathbf{i})(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o} \quad \to \quad \mathbf{s}_2$$

(using Metarule 2 to replace $\mathbf{p}_2\mathbf{o}^{\ell}$ by $\mathbf{p}_2\hat{\mathbf{o}}^{\ell}\mathbf{i}^{\ell}\,[\mathbf{i}$,

were flowers given to her ?

$$(\mathbf{q}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell}\,\pi_2^{\ell})\mathbf{n}_2\,(\mathbf{p}_2\,\hat{\mathbf{o}}^{\ell}\,\mathbf{i}^{\ell}\,[\mathbf{i})(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o} \quad \to \quad \mathbf{q}_2$$

(using Metarule 1 with $x = \mathbf{p}_2\mathbf{o}^{\ell}$),

what was given to her ?

$$(\overline{\mathbf{q}}\,\mathbf{s}^{\ell}\,\pi_3)(\pi_3^{r}\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})(\mathbf{p}_2\,\hat{\mathbf{o}}^{\ell}\,\mathbf{i}^{\ell}\,[\mathbf{i})(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o} \quad \to \quad \overline{\mathbf{q}}$$

(using the in situ typing of *what*).

From the pseudo-question

**did he give to her flowers ?*

$$(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_3\,(\mathbf{i}\,\hat{\mathbf{o}}^{\ell}\,\mathbf{i}^{\ell}\,[\mathbf{i})(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o}\,\mathbf{o} \quad \not\to \quad \mathbf{q}_2$$

we may derive

what (did he give to her) ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})\,(\mathbf{q}_2\,\hat{\mathbf{o}}^{\ell}) \quad \to \quad \overline{\mathbf{q}}$$

Our grammar also predicts the awkward

she was given flowers to,

$$\pi_3\,(\pi_3^{r}\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})(\mathbf{p}_2\,\mathbf{j}^{\ell}\,[\mathbf{i}\,\mathbf{o}^{\ell})\mathbf{o}\,(\mathbf{i}^{r}\,\mathbf{i}\,\mathbf{o}^{\ell}) \quad \to \quad \mathbf{s}_2$$

but this is better replaced by

she was given flowers,

where we have taken advantage of the fact that English differs from other European languages in allowing the indirect object to turn into the subject of a passive construction. We will explain this with the help of Metarule 5 in the next chapter.

15. Sentential complements.

Direct or indirect sentences may appear as first or second complements of certain verbs. For example, verbs like

believe, *know*, *say*, *think*, ...

admit complements of type $\mathbf{s}$ or

$$\bar{\mathbf{s}} = \text{ indirect statement,}$$

and verbs like

know, *wonder*, *ask*, ...

admit complements of type

$$\mathbf{t} = \text{ indirect question.}$$

The former verbs have type $\mathbf{i}\mathbf{s}^\ell$ or $\mathbf{i}\bar{\mathbf{s}}^\ell$ and the latter type $\mathbf{i}\mathbf{t}^\ell$.

For example, in the sentence

$$\begin{array}{l} \textit{she knew [that] she loved him} \\ \pi_3\ (\pi^r\,\mathbf{s}_2\,\bar{\mathbf{s}}^\ell)(\bar{\mathbf{s}}\,\mathbf{s}^\ell)\ \pi_3\ (\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell)\mathbf{o} \end{array} \quad \to\ \mathbf{s}_2$$

the complementizer *that* is optional. In its absence the type of *knew* changes to $\pi^r\mathbf{s}_2\mathbf{s}^\ell$.

We now look at verbs with two complements: an object and an indirect sentence. I will treat the object as a direct one of type $\mathbf{o}$, even if it may have carried a dative rather than an accusative case in the distant past. Many of these verbs also incorporate causation:

$$\begin{aligned} \textit{tell} = &\quad \textit{let know} \text{ of type } \mathbf{i}\mathbf{s}^\ell\mathbf{o}^\ell, \mathbf{i}\bar{\mathbf{s}}^\ell\mathbf{o}^\ell, \mathbf{i}\mathbf{t}^\ell\mathbf{o}^\ell, \mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell, \\ \textit{persuade} = &\quad \textit{make believe} \text{ of type } \mathbf{i}\bar{\mathbf{s}}^\ell\mathbf{o}^\ell, \mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell, \end{aligned}$$

where

$$\bar{\mathbf{j}} = \text{ complete infinitive with } \textit{to}.$$

Here are some sample sentences illustrating the first:

she told him [*that*] *she loved him*,

she told him whom she loved -,

she told him whether she loved him,

she told him whom to love -,

she told him [*for him*] *to love her.*

(Here the optional [*for him*] would be replaced by Chomsky's PRO. See Chapter 17 below.)

The passive of these verbs is formed in partial analogy to that of doubly transitive ones. But, to form the passive of a Wh-question asking for the object, we require another metarule.

Metarule 5. If a verb has type $\mathbf{i}x^\ell\mathbf{o}^\ell$, with $x = \mathbf{s}, \bar{\mathbf{s}}, \mathbf{t}, \bar{\mathbf{j}}$ or $\bar{\mathbf{n}}$, then it also has type $\mathbf{i}\hat{\mathbf{o}}^\ell x^\ell$.

(With a little extra work, one might derive Metarule 2 from Metarule 5 by allowing the case $x = \mathbf{i}^r\mathbf{i}$.) Here are some illustrations:

he was told [that] she loved him -,
he was told whom she loved - - ,
he was told whether she loved him - ,
he was told whom to love - - ,
he was told to love her - .

Note the two consecutive traces in the second and fourth of these sentences. For example, we analyze the second one as follows:

$$\begin{array}{c} \textit{he was told whom she loved} \;-\;-\;. \\ \pi_3\,(\pi^r\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\hat{\mathbf{o}}^\ell\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_3\,(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \to \quad \mathbf{s}_2 \end{array}$$

Verbs of type $\mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell$ similarly receive the alternative type $\mathbf{i}\hat{\mathbf{o}}^\ell\bar{\mathbf{j}}^\ell$ by Metarule 5, e.g.

$$\begin{array}{c} \textit{he was told to come} \;-\;, \\ \pi_3\,(\pi^r\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\hat{\mathbf{o}}^\ell\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{i}^\ell)\mathbf{i} \quad \to \quad \mathbf{s}_2 \end{array}$$

Similarly, we may treat Wh-questions asking for the object:

whom did she tell that she loved him - ?
whom did she tell whom she loved - - ?
whom did she tell whether she loved him - ?
whom did she tell whom to love - - ?
whom did she tell to love her - ?

For example, we invoke Metarule 5 with $x = \mathbf{t}$ to analyze the second question above, after looking at a declarative sentence leading up to it:

$$\begin{array}{c} \textit{she told him whom she loved} \;-\;, \\ \pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{t}^\ell\,\mathbf{o}^\ell)\mathbf{o}\,(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_3\,(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \to \quad \mathbf{s}_2 \end{array}$$

$$\begin{array}{c} \textit{whom did she tell whom she loved} \;-\;-\;? \\ (\bar{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_3(\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \to \quad \bar{\mathbf{q}} \end{array}$$

We may also apply Metarule 5 in the case $x = \overline{\mathbf{n}}$ to explain the last sentence of Chapter 14. The metarule allows us to assign to the infinitive *give*, originally of type

$$\mathbf{i}\overline{\mathbf{n}}^\ell\mathbf{o}'^\ell \to \mathbf{i}\overline{\mathbf{n}}^\ell\mathbf{o}^\ell,$$

the new type $\mathbf{i}\hat{\mathbf{o}}^\ell\overline{\mathbf{n}}^\ell$, hence to the past participle *given* the new type

$$\mathbf{p}_2\mathbf{j}'^\ell\mathbf{i}\hat{\mathbf{o}}^\ell\overline{\mathbf{n}}^\ell \to \mathbf{p}_2\hat{\mathbf{o}}^\ell\overline{\mathbf{n}}^\ell.$$

We then have

$$\begin{array}{c} \textit{was she given flowers} \;-\;? \\ (\mathbf{q}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell\,\pi_3^\ell)\pi_3(\mathbf{p}_2\,\hat{\mathbf{o}}^\ell\,\overline{\mathbf{n}}^\ell)\mathbf{n}_2 \quad \to \quad \mathbf{q}_2 \end{array}$$

This will not yet explain

$$\begin{array}{c} \textit{what was she given} \ - \ - \ ? \\ (\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell}\,\pi_3^{\ell})\pi_3(\mathbf{p}_2\,\hat{\mathbf{o}}^{\ell}\,\overline{\mathbf{n}}^{\ell}) \quad \rightarrow \quad \mathbf{q}_2 \end{array}$$

as long as $\overline{\mathbf{n}}^{\ell} \not\rightarrow \hat{\mathbf{o}}^{\ell}$, that is, $\hat{\mathbf{o}} \not\rightarrow \overline{\mathbf{n}}$.

The way around this objection is to recall from Chapter 7 that past participles of transitive verbs behave much like adjectives. We may take advantage of this observation and assign to them the type $\overline{\mathbf{a}}$, which admits them as complements of the copula *be* of type $\mathbf{i}^{\ell}\overline{\mathbf{a}}^{\ell}$. More generally, we adopt the following

Metarule 6. In the type of a past participle, $\mathbf{p}_2\overline{\mathbf{n}}^{\ell}$ may be replaced by $\overline{\mathbf{a}}$.

We may now re-analyze the above question as follows:

$$\begin{array}{c} \textit{what was she given} \ - \ ? \\ (\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\overline{\mathbf{a}}^{\ell}\,\pi_3^{\ell})\pi_3(\overline{\mathbf{a}}\,\mathbf{o}^{\ell}) \quad \rightarrow \quad \overline{\mathbf{q}} \end{array}$$

Metarule 6 is often useful when an indirect sentence appears in subject position with obligatory complementizer. Consider, for example,

$$\begin{array}{c} \textit{that she loved him is true} \\ (\overline{\mathbf{s}}\,\mathbf{s}^{\ell})\ \pi_3(\pi^r\,\mathbf{s}_2\,\mathbf{o}^{\ell})\ \mathbf{o}\ (\pi_3\,\mathbf{s}_1\,\overline{\mathbf{a}}^{\ell})\mathbf{a} \quad \rightarrow \quad \mathbf{s}_1 \end{array}$$

where we postulate

$$\overline{\mathbf{s}} \rightarrow \pi_3.$$

Here *true* may be replaced by *known* of type $\mathbf{p}_2\mathbf{o}^{\ell} \rightarrow \mathbf{p}_2\overline{\mathbf{n}}^{\ell}$, hence of type $\overline{\mathbf{a}}$ by Metarule 6.

The postulate $\overline{\mathbf{s}} \rightarrow \pi_3$ also allows

$$\begin{array}{c} (\textit{that she loved him})\ \textit{surprises John} \\ \overline{\mathbf{s}} \qquad (\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^{\ell})\ \mathbf{o} \qquad\qquad \rightarrow \quad \mathbf{s}_1 \end{array}$$

where *surprises* can be replaced by *annoys*, or even by any transitive verb, although it may not always be easy to interpret the resulting sentence.

It is customary to replace such sentences by constructions involving the impersonal pronoun *it*, e.g.

$$\begin{array}{c} \textit{it surprises John (that she loved him)} \\ (\mathbf{s}\,\overline{\mathbf{s}}^{\ell}\,\mathbf{s}^{\ell}\,\pi_3)\ (\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^{\ell})\mathbf{o}\ \overline{\mathbf{s}} \qquad \rightarrow \quad \mathbf{s} \end{array}$$

where *it* has type $\mathbf{s}\overline{\mathbf{s}}^{\ell}\mathbf{s}^{\ell}\pi_3$, or even type $\mathbf{s}_j\overline{\mathbf{s}}^{\ell}\mathbf{s}_j^{\ell}\pi_3$ $(j = 1, 2)$ if we want to keep track of the tense. Note that the simpler type π_3 won't do, because *it* cannot be replaced by *he*.

16. Discontinuous dependencies.

In this chapter, we look at three kinds of verbs which require as a second complement a bare infinitive (without *to*) or a prepositional adverb.

16.1 Verbs of causation.

The verbs *have, help, let* and *make* (in alphabetic order) may denote different degrees of causation. While *make* denotes strict causation, *let* may denote permission, *help* involves cooperation of the subject and *have* may invoke a third person as an intermediary. All these verbs require a bare infinitival phrase as a second complement, as in

$$\begin{array}{c} (\textit{have/help/let/make})\ \textit{her}\ (\textit{do the dishes}) \\ (\mathbf{i}\,\mathbf{i}'^{\ell}\,\mathbf{o}^{\ell})\ \mathbf{o}\ \ \mathbf{i} \end{array} \quad \to \ \mathbf{i}$$

The reason for $\mathbf{i}'^{\ell}$ instead of $\mathbf{i}^{\ell}$ is to allow for

$$\begin{array}{c} \textit{let her be prepared} - \\ (\mathbf{i}\,\mathbf{i}'^{\ell}\,\mathbf{o}^{\ell})\mathbf{o}\,(\mathbf{i}'\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\,\mathbf{o}^{\ell}) \end{array} \quad \to \ \mathbf{i}$$

Sometimes, but not always, the order of the two complements may be reversed:

let the girl go,
let her go,
let go the girl,
**let go her.*

We try to account for this by citing a variant of Metarule 5 (for the moment only in case $x = \mathbf{i}'$).

Metarule 7. If a verb has type $\mathbf{i}x^{\ell}\mathbf{o}^{\ell}$, then it also has type $\mathbf{i}\overline{\mathbf{n}}^{\ell}x^{\ell}$, where $x = \mathbf{i}',\ \overline{\mathbf{a}}, \delta, \cdots$.

Here we recall from Chapter 14 that

$$\overline{\mathbf{n}} = \text{ complete noun phrase,}$$

subject to the postulates

$$\overline{\mathbf{n}}_i \to \overline{\mathbf{n}} \to \mathbf{o}\ \not\to\ \overline{\mathbf{n}}\ (i = 0, 1, 2).$$

Thus, *let* of type $\mathbf{i}\mathbf{i}'^{\ell}\mathbf{o}^{\ell}$ also has type $\mathbf{i}\overline{\mathbf{n}}^{\ell}\mathbf{i}'^{\ell}$ to account for

$$\begin{array}{c} \textit{let go the girl,} \\ (\mathbf{i}\,\overline{\mathbf{n}}^{\ell}\,\mathbf{i}'^{\ell})\mathbf{i}\,(\overline{\mathbf{n}}_1\,\mathbf{n}_1^{\ell})\mathbf{n}_1 \end{array} \quad \to \mathbf{i}$$

but

$$\begin{array}{c} {}^{*}\textit{let go her} \\ (\mathbf{i}\overline{\mathbf{n}}^{\ell}\,\mathbf{i}'^{\ell})\,\mathbf{i}\ \ \mathbf{o} \end{array} \ \not\to\ \mathbf{i}.$$

The verb *make* can also have type $\mathbf{i}\overline{\mathbf{a}}^{\ell}\mathbf{o}^{\ell}$, hence $\mathbf{i}\overline{\mathbf{n}}^{\ell}\overline{\mathbf{a}}^{\ell}$, by Metarule 7 with $x = \overline{\mathbf{a}}$. This will account for

make the promise good,
make it good,
make good the promise,
**make good it.*

16.2 Verbs of perception.

Verbs of perception, such as *see* and *hear*, may also admit bare infinitival complements and are typed like verbs of causation, as in

$$\begin{array}{c} \textit{see her be kissed } - \ . \\ (\mathbf{i}\,\mathbf{i}'^{\ell}\,\mathbf{o}^{\ell})\,\mathbf{o}\,(\mathbf{i}'\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell})(\mathbf{p}_2\,\mathbf{o}^{\ell}) \end{array} \rightarrow \mathbf{i}$$

The analogy between verbs of perception and verbs of causation may be due to a kind of folk philosophy, which imagines the act of perception as bringing the perceived object into being. While this view contradicts the insights of modern science, it was at one time defended by the eminent philosopher George Berkeley. Even long before his time, the ancient philosopher Empedocles believed that light rays emanate from the eye of the beholder.

16.3 Transitive phrasal verbs.

"Phrasal verbs" are compounds of verbs with certain particles, called "prepositional adverbs", usually taking the form of prepositions. We are here interested in situations where an additional object is required, that is, where the phrasal verb is transitive. Consider

they brought the criminal in,
they brought him in,
they brought in the criminal,
**they brought in him.*

Although we feel that *bring in* is a compound verb, these examples suggest that there is a discontinuity between *bring* and *in*, which may be partially bridged with the help of Metarule 7, taking $x = \delta$, where δ is the type of the particle *in*.

Although *in* looks like a preposition, it does not function like one in the present context. Assigning to *bring* the type $\mathbf{i}\delta^{\ell}\mathbf{o}^{\ell}$, we obtain a secondary type $\mathbf{i}\overline{\mathbf{n}}^{\ell}\delta^{\ell}$ by Metarule 6. The primary type will account for

$$\begin{array}{c} \textit{they brought him in} \\ \pi_2(\pi^r\,\mathbf{s}_2\,\delta^{\ell}\,\mathbf{o}^{\ell})\,\mathbf{o}\;\;\delta \end{array} \rightarrow \mathbf{s}_2$$

and the secondary type for

$$\begin{array}{c} \textit{they brought in the criminal}, \\ \pi_2\,(\pi^r\,\mathbf{s}_2\,\overline{\mathbf{n}}^{\ell}\,\delta^{\ell})\delta(\overline{\mathbf{n}}_1\,\mathbf{n}_1^{\ell})\mathbf{n}_1 \end{array} \rightarrow \mathbf{s}_2$$

but

$$\begin{array}{c} {}^{*}\textit{they brought in him}, \\ \pi_2\,(\pi^r\,\mathbf{s}_2\overline{\mathbf{n}}^{\ell}\,\delta^{\ell})\,\delta\;\;\mathbf{o} \end{array} \not\rightarrow \mathbf{s}_2$$

since

$$\overline{\mathbf{n}}_1 \rightarrow \overline{\mathbf{n}} \rightarrow \mathbf{o} \not\rightarrow \overline{\mathbf{n}}.$$

Similarly, we have

I turn the light off,
I turn it off,
I turn off the light,
**I turn off it.*

To account for these examples, we may as well assign the same type δ to the particle *off*. In a more precise analysis of English syntax, we may wish to replace δ by $\delta_1, \delta_2, \cdots$ to take account of which prepositional adverbs go with which verbs.

17. Quasi-sentences.

Consider the sentence

Jane wanted him to kiss her.

At first sight, *want* has two complements: a direct object *him* and a complete infinitive *to kiss her*. However, we may expand the above sentence by introducing the complementizer *for* to obtain

Jane wanted for him to kiss her,

and we then tend to think of *want* as having only one complement, the indirect quasi-sentence

[*for*] *him to kiss her.*

Indeed, if we replace *want* by *hope*, the complementizer becomes obligatory:

Jane hoped for him to kiss her,

but

**John hoped him to kiss her.*

In all the above examples *her* could refer to Jane, but conceivably also to Alice. However, in the sentence

(17.1) *Jane hoped for her to kiss him*

the prepositional phrase *for her* can be omitted altogether to yield

Jane hoped to kiss him,

but this assumes that *her* in (17.1) refers to Jane and not to Alice.

Similarly,

Jane persuaded John [*for him*] *to kiss her*

has two complements: the object *John* and the quasi-sentence

[*for him*] *to kiss her*,

where again *for him* can be omitted, provided that *him* refers to John and not to Bob. We will return to these problems of internal reference in a later chapter devoted to anaphora.

Note that the complementizer *for* is obligatory when the quasi-sentence appears in subject position:

(17.2) *for him to kiss her would be nice*,

**him to kiss her would be nice.*

Here also the whole prepositional phrase may be omitted to yield

to kiss her would be nice.

We could also rephrase (17.2) more elegantly as

$$it\ would\ be\ nice\ [for\ him]\ to\ kiss\ her.$$

This is a construction to which we will return later.

Quasi-sentences may also be formed from participles instead of infinitives. The rôle of the complementizer is then taken by the inflector *Gen* (= genetive), where

$$Gen\ John\ \rightarrow\ John's,\ Gen\ him\ \rightarrow\ his.$$

To make the analogy complete, I have assumed that *Gen* operates on *him* rather than on *he*. For example,

Jane recalled his/him kissing her,
Jane recalled Alice kissing him,
Jane recalled [*her*] *kissing him*,

where *her* may refer to Jane; but in

Jane recalled kissing her

her cannot refer to Jane.

Again, the complementizer *Gen* is obligatory in subject position:

his kissing her would be nice
**him kissing her would be nice.*

The quasi-sentence

[*Gen*] *John kissing Alice*

should not be confused with the gerund noun phrase

$$John's\ kissing\ of\ Alice.$$

To account for quasi-sentences in our pregroup grammar we will introduce the basic types

ϕ = quasi-sentence formed from infinitive,

ψ = quasi-sentence formed from participle, and postulate

$$\phi, \psi, \bar{\mathbf{j}}, \mathbf{p}_1 \rightarrow \pi_3.$$

We can then assign appropriate types to the words *want, hope, recall* and *persuade* as follows:

want: $\mathbf{i}\phi^\ell, \mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell, \mathbf{i}\bar{\mathbf{j}}^\ell$;

hope: $\mathbf{i}\phi^\ell, \mathbf{i}\bar{\mathbf{j}}^\ell$;

recall: $\mathbf{i}\psi^\ell, \mathbf{i}\mathbf{p}_1^\ell\mathbf{o}^\ell, \mathbf{i}\mathbf{p}_1^\ell$;

persuade: $\mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell$.

The proper types for the complementaries are then as follows:

for: $\phi\bar{\mathbf{j}}^\ell\mathbf{o}^\ell$,

Gen: $\psi\mathbf{p}_1^\ell\mathbf{o}^\ell$.

Here are some illustrations:

want for him to come, *want him to come*, *want to come*;
hope [for him] to come, **hope him to come*;
recall his coming, *recall him coming*, *recall coming*;
persuade him to come, **persuade [for him] to come.*

Much has been written about the contraction

$$\textit{want to} \;\to\; \textit{wanna}$$

originally in connection with justifying Chomsky's traces. The way I see it, the contraction is permitted for *want* of type $\mathbf{i}\bar{\mathbf{j}}^\ell$, but not for *want* of type $\mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell$, which also has type $\mathbf{i}\hat{\mathbf{o}}^\ell\bar{\mathbf{j}}^\ell$ by Metarule 5. Thus we have

you wanna kiss Jane,

$$\pi_2(\pi_2^r\,\mathbf{s}_1\,\bar{\mathbf{j}}^\ell\,\bar{\mathbf{j}}\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\;\mathbf{o} \quad \to \mathbf{s}_1$$

whom do you wanna kiss $-$?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_2^\ell)\pi_2(\mathbf{i}\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell) \quad \to \overline{\mathbf{q}}$$

whom do you want to kiss Jane ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_2^\ell)\pi_2(\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\bar{\mathbf{j}}^\ell)(\bar{\mathbf{j}}\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \quad \to \overline{\mathbf{q}}$$

whom do you **wanna kiss Jane* ?

$$(\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_2)\;\pi_2\,(\mathbf{i}\,\hat{\mathbf{o}}^\ell\,\mathbf{j}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{o} \quad \to \overline{\mathbf{q}}$$

In other words, *wanna* has type $\mathbf{i}\mathbf{j}^\ell$, but not type $\mathbf{i}\bar{\mathbf{j}}^\ell\mathbf{o}^\ell$ or $\mathbf{i}\hat{\mathbf{o}}^\ell\mathbf{j}^\ell$.

The situation is quite analogous with *gonna*, an abbreviation of *going to*, where traces won't serve as an explanation why we can say

he is gonna come

$$\pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{p}_1^\ell)(\mathbf{p}_1\,\mathbf{j}^\ell)\mathbf{i} \quad \to \mathbf{s}_1$$

but not

he is **gonna town.*

$$\pi_3(\pi_3^r\mathbf{s}_1\mathbf{p}_1^\ell)(\mathbf{p}_1\mathbf{o}^\ell)\mathbf{o}$$

Here *gonna* may have type $\mathbf{p}_1\mathbf{j}^\ell$ but not $\mathbf{p}_1\mathbf{o}^\ell$.

V. Unbounded Dependencies and Constraints.

18. Unbounded dependencies.

Pregroup grammars easily account for long distance dependencies, as in

whom did I say [that] I think [that] I saw − ?

We note that the verbs *think* and *say* admit direct or indirect sentences as complements, hence have type $\mathbf{i}\sigma^\ell$, where

$$\sigma = \text{ direct or indirect sentence.}$$

To start with, we postulate

$$\mathbf{s} \to \sigma \not\to \mathbf{s}, \qquad \overline{\mathbf{s}} \to \sigma \not\to \overline{s}.$$

Thus we can analyze the question

whom do you think [that] I saw − ?

$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_2^\ell)\pi_2\,(\mathbf{i}\,\sigma^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \to \quad \overline{\mathbf{q}}$$

with the complementizer *that* being optional.

Similarly we analyze the indirect question in

I know whom you think [that] I saw −

$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)\,\pi_2(\pi_2^r\,\mathbf{s}_1\,\sigma^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi_1^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \to \quad \mathbf{s}_1$$

and the relative clause in

the girl whom (you think [that] I saw −) .

$$(\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell)[\mathbf{n}_1(\mathbf{n}_1^r\,\mathbf{n}_1\,\hat{\mathbf{o}}^{\ell\ell}\,\sigma^\ell)(\mathbf{s}_2\,\mathbf{o}^\ell) \qquad \to \quad \overline{\mathbf{n}}_1$$

Note however that the complementizer *that* becomes obligatory when the indirect sentence is used as subject:

that I saw her is nice,

$$(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell)\mathbf{o}\,(\pi_3^r\,\mathbf{s}_1\,\overline{\mathbf{a}}^\ell)\mathbf{a} \quad \to \quad \mathbf{s}_1$$

if we postulate

$$\overline{\mathbf{s}} \to \pi_3.$$

Recall from Chapter 11 that the question

who saw me ?

$$(\overline{\mathbf{q}}\,\mathbf{s}^\ell\,\pi_3)(\pi_3^r\,\mathbf{s}_2\,\mathbf{o}^\ell)\mathbf{o} \quad \to \quad \overline{\mathbf{q}}$$

could also be analyzed as follows:

$$(\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\hat{\pi}_3^\ell\,\mathbf{o}^\ell)\mathbf{o} \quad \to \quad \overline{\mathbf{q}}$$

Postulating

$$\mathbf{q} \rightarrow \sigma$$

we can analyze

$$\textit{who do you think saw me} - ?$$
$$(\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_2^\ell)\pi_2(\mathbf{i}\,\sigma^\ell)(\mathbf{q}_2\,\hat{\pi}_3^\ell\,\mathbf{o}^\ell)\mathbf{o}$$

However, we cannot justify

$$^*\textit{who do you think that } \text{saw me?}$$
$$(\overline{\mathbf{s}}\mathbf{s}^\ell)\ (\mathbf{q}\hat{\pi}_3^\ell\mathbf{o}^\ell)$$

as long as $\mathbf{q} \not\rightarrow \mathbf{s}$.

The verb *say* also has types $\mathbf{i}\sigma^\ell$, so we can justify the following examples with optional *that* included:

$$\textit{who did I say [that] I think saw me} - ?$$
$$(\overline{\mathbf{q}}\,\hat{\pi}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_1(\mathbf{i}\,\sigma^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi_1^r\,\mathbf{s}_1\,\sigma^\ell)(\mathbf{q}_2\,\hat{\pi}^\ell\,\mathbf{o}^\ell)\mathbf{o} \quad \rightarrow \quad \overline{\mathbf{q}}$$

$$\textit{whom did I say [that] I think [that] I saw} - ?$$
$$(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_1(\mathbf{i}\,\sigma^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_1\,\sigma^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\pi_1(\pi^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \rightarrow \quad \overline{\mathbf{q}}$$

We can even justify the following sentence which McCawley considers to be ungrammatical:

$$\textit{I wonder whom that John was promoted} - \textit{pleased} -$$
$$\pi_1(\pi_1^r\,\mathbf{s}_1\,\mathbf{t}^\ell)(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\ \mathbf{n}\,(\pi^r\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{o}^\ell)(\pi_3^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \quad \rightarrow \quad \mathbf{s}_1$$

The problem here lies with the short-term memory. Indeed, proceeding step by step, we calculate

(1) $\pi_1\ (\pi_1^r\mathbf{s}_1\mathbf{t}^\ell) \ \rightarrow\ \mathbf{s}_1\mathbf{t}^\ell,$

(2) $\mathbf{s}_1\,\mathbf{t}^\ell(\mathbf{t}\,\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell) \ \rightarrow\ \mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell,$

(3) $(\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell)(\overline{\mathbf{s}}\,\mathbf{s}^\ell)\,\mathbf{n}(\pi^r\,\mathbf{s}_2\,\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell) \rightarrow \mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\overline{\mathbf{s}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell,$

(4) $(\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\overline{\mathbf{s}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^\ell)(\mathbf{p}_2\,\mathbf{o}^\ell) \rightarrow \mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\overline{\mathbf{s}},$

(5) $(\mathbf{s}_1\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{s}^\ell\,\overline{\mathbf{s}})(\pi_3^r\,\mathbf{s}_2\,\mathbf{o}^\ell) \rightarrow \mathbf{s}_1,$

making use of the postulate $\overline{\mathbf{s}} \rightarrow \pi_3$.

In the third line of this calculation, the speaker or hearer must temporarily hold ten simple types, here identified with Miller's chunks of information, in her short-term memory, thus exceeding Miller's limit of seven plus or minus two. Fortunately, the above sentence can be paraphrased more elegantly by

I wonder whom it pleased (*that John was promoted*);

but the type we assigned to *it* at the end of chapter 15 won't work here. We must defer a discussion of this to another occasion.

Unfortunately, like mainstream grammars, ours also predicts some impossible sentences, such as the question

$$\begin{array}{c} {}^{*}\textit{whom did I see him and} - ? \\ (\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_1(\mathbf{i}\,[\mathbf{o}^{\ell})\;\;[\mathbf{o}\,(\mathbf{o}^{r}\,\mathbf{o}\,\mathbf{o}^{\ell}) \quad \rightarrow \; \overline{\mathbf{q}} \end{array}$$

Both left square brackets are left over from the "deep structure" discussed in Chapter 3. The first square bracket ensures that the question does not end after *see* and the second square bracket indicates that *did I see him* is not a constituent.

Although Anne Preller [2004] has shown that such sentences can be avoided by clever type assignments, I believe that they are not strictly ungrammatical, but should be ruled out for causing processing difficulty. I have suggested elsewhere that they can be avoided by what I call the "**block constraint**", which forbids compound types of the form $[x[y$, x and y being simple types arising in the calculation.

Let us look at another example:

$$\begin{array}{c} {}^{*}\textit{whom should I say who saw} - ? \\ (\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{j}^{\ell}\,\pi^{\ell})\pi_1(\mathbf{i}\,\overline{\mathbf{q}}^{\ell})(\overline{\mathbf{q}}\,\hat{\pi}_3^{\ell\ell}\,\mathbf{q}^{\ell})((\mathbf{q}_2\,\hat{\pi}_3^{\ell}\,\mathbf{o}^{\ell}) \;\; \rightarrow \overline{\mathbf{q}} \end{array}$$

Again, our grammar wrongly predicts that this is a well-formed question. Elsewhere I have suggested another constraint, which disallows two double adjoints to appear successively at any stage of the calculation, at least when each double adjoint was originally preceded by a left bracket. This may be incorporated into our block constraint by retaining left brackets immediately preceding double adjoints, whenever they occurred there in what I have called the deep structure. In the present case, we thus calculate,

$$(\overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{j}^{\ell}\,\pi^{\ell})\pi_1(\mathbf{i}\,\overline{\mathbf{q}}^{\ell})(\overline{\mathbf{q}}[\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell}).....$$

$$\rightarrow \;\; \overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}[\hat{\pi}_3^{\ell\ell}......$$

and the block constraint tells us to abort the calculation. However, there is nothing to prevent us from processing

$$\begin{array}{c} \textit{to whom was she promised} - ? \\ (\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell\ell}\,\overline{\mathbf{q}}^{\ell})(\overline{\mathbf{q}}\,\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{p}_2^{\ell}\,\pi_3^{\ell})\pi_3\,(\mathbf{p}_2\,\mathbf{o}^{\ell}) \quad \rightarrow \; \overline{\mathbf{q}} \end{array}$$

since the first occurrence of $\hat{\mathbf{o}}^{\ell\ell}$ here was never preceded by a left bracket in the first place.

19. Unblocking maneuver.

Apparently, the following will be ruled out by the block constraint:

whom did he paint pictures of − ?

$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\,\pi_3(\mathbf{i}\,[\mathbf{o}^{\ell})\,[\mathbf{n}_2\ (\mathbf{n}_2^{r}\mathbf{n}_2\mathbf{o}^{\ell}) \ \to\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{o}^{\ell}.....$$

Here the first square bracket belongs to the double adjoint, the second bracket ensures that the question does not end after *paint* and the third bracket prevents *did he paint pictures* from being a constituent.

Fortunately, we can replace $[\mathbf{o}^{\ell}$ by $\mathbf{n}_2^{\ell}$, without left bracket, since $\mathbf{n}_2 \to \mathbf{o}$ and $\hat{\mathbf{o}} \not\to \mathbf{n}_2$, hence $\mathbf{n}_2^{\ell} \not\to \hat{\mathbf{o}}^{\ell}$. Thus we arrive at

$$\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{n}_2^{\ell}\,[\mathbf{n}_2\mathbf{n}_2^{r}\,\mathbf{n}_2\,\mathbf{o}^{\ell} \ \to\ \overline{\mathbf{q}}$$

and the calculation is successful after all.

I will call this trick the "unblocking maneuver". It also applies to

what does she like reading books about − ?

$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_1\,\mathbf{i}^{\ell}\,\pi_3^{\ell})\,\pi_3(\mathbf{i}\,\mathbf{p}_1^{\ell})(\mathbf{p}_1\,[\mathbf{o}^{\ell})\,[\mathbf{n}_2\ (\mathbf{n}_2^{r}\mathbf{n}_2\mathbf{o}^{\ell}) \ \to\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{o}^{\ell}.....$$

Again, we may replace $[\mathbf{o}^{\ell}$ by $\mathbf{n}_2^{\ell}$ and calculate

$$\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{n}_2^{\ell}\,[\mathbf{n}_2\mathbf{n}_2^{r}\,\mathbf{n}_2\,\mathbf{o}^{\ell} \ \to\ \overline{\mathbf{q}}.$$

The unblocking maneuver will not help to dismiss

**whom does she like reading books which discuss* − − ?

$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_1\,\mathbf{i}^{\ell}\,\pi_3^{\ell})\,\pi_3(\mathbf{i}\,\mathbf{p}_1^{\ell})(\mathbf{p}_1\,[\mathbf{o}^{\ell})\,[\mathbf{n}_2\ (\mathbf{n}_2^{r}\mathbf{n}_2\mathbf{s}^{\ell}\pi_2)(\pi_2^{r}\mathbf{s}_1\mathbf{o}^{\ell})$$

Having replaced $[\mathbf{o}^{\ell}$ by $\mathbf{n}_2^{\ell}$, where $\mathbf{n}_2^{\ell} \not\to \hat{\mathbf{o}}^{\ell}$ and so $\hat{\mathbf{o}}^{\ell\ell}\mathbf{n}_2^{\ell} \not\to 1$, we obtain

$$\longrightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{n}_2^{\ell}\,[\mathbf{n}_2\mathbf{n}_2^{r}\ \mathbf{n}_2\ \mathbf{s}^{\ell}\,\pi_2\,\pi_2^{r}\ \mathbf{s}_1\,\mathbf{o}^{\ell} \ \to\ \overline{\mathbf{q}}$$

To salvage our procedure, we will retain the original left bracket in front of $\mathbf{s}^{\ell}$ in the type of *which* after all, thus arriving at

$$\longrightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{s}^{\ell}......$$

which is ruled out by the block constraint.

Similarly, in

**what did you see the girl who drank* − ?

$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\,\pi_2(\mathbf{i}\,[\mathbf{o}^{\ell})\,(\overline{\mathbf{n}}_1\mathbf{n}_1^{\ell})\,[\mathbf{n}_1(\mathbf{n}_1^{r}\mathbf{n}_1\,[\mathbf{s}^{\ell}\pi_3)(\pi^{r}\mathbf{s}_2\mathbf{o}^{\ell})$$

$$\longrightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{o}^{\ell}......$$

we run into the block constraint, but may apply the unblocking process, replacing $[\mathbf{o}^\ell$ by $\overline{\mathbf{n}}_1^\ell$, then the calculation proceeds to

$$\longrightarrow \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\overline{\mathbf{n}}_1^\ell\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell\,[\mathbf{n}_1\mathbf{n}_1^r\,\mathbf{n}_1\,[\mathbf{s}^\ell.....$$

$$\longrightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{s}^\ell\cdots\cdots$$

and our calculation is aborted by the block constraint.

We are thus forced to retain the left bracket in front of $\mathbf{s}^\ell$ in the type $\mathbf{n}_i^r\mathbf{n}_i\,[\mathbf{s}^\ell\pi_k$ of the subject relative pronoun *who* / *which* / *that*. The same applies to the type $\mathbf{n}_i^r\mathbf{n}_i\,[\mathbf{s}^\ell\pi_1\mathbf{n}_i^\ell$ of *whose* (see Chapter 13).

The unblocking procedure may have to be applied twice in the same sentence:

whom did I see copies of pictures of $-$?

$$(\overline{\mathbf{q}}\,[_1\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\,\pi_1(\mathbf{i}\,[_2\mathbf{o}^\ell)\,[_3\mathbf{n}_2\ (\mathbf{n}_2^r\mathbf{n}_2\,[_4\mathbf{o}^\ell)\,[_5\mathbf{n}_2(\mathbf{n}_2^r\mathbf{n}_2\mathbf{o}^\ell)$$

where I have labeled the left brackets 1 to 5:

1 belongs to a double adjoint,

2 prevents the question from ending after *see*,

3 ensures that *did I see copies* is not a constituent,

4 prevents the question from ending after the first *of*,

5 ensures that *copies of pictures* is not a constituent.

The unblocking maneuver may be applied to the second and fourth brackets, yielding

$$\begin{aligned}&\rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{n}_2^\ell\,[\mathbf{n}_2\mathbf{n}_2^r\,\mathbf{n}_2\ \ \mathbf{n}_2^\ell\,[\mathbf{n}_2\mathbf{n}_2^r\,\mathbf{n}_2\,\mathbf{o}^\ell\\ &\rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^\ell\ \rightarrow\ \overline{\mathbf{q}}\end{aligned}$$

I must confess that I am not too happy about the unblocking maneuver. All I can say is that it works, but I have not calculated the cost it takes in time and effort to carry it out and what new processing difficulties it might pose.

Next, consider

**whom did you sleep and hear* $-$?

$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_2\,[\mathbf{i}\ (\mathbf{i}^r\,\mathbf{i}\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\quad\rightarrow\ \overline{\mathbf{q}}$$

To eliminate this misleading calculation we need an algebraic equivalent of what mainstream grammarians call the "coordinate structure constraint". Before we do so, we will take a closer look at coordinate and subordinate conjunctions.

20. Coordinate and subordinate conjunctions.

Coordinate conjunctions typically combine two constituents of type x (say) to form another expression of the same type, hence they should have polymorphic type $x^r x x^\ell$. There are exceptions: *she and I* is a plural noun phrase, though both constituents are singular. Hence *and* here has type $\pi^r \pi_2 \pi^\ell$.

The story is even more complicated with the conjunction *or*. (Never mind that logicians call this a "disjunction".) Here is what the American Heritage Dictionary of English says, as was pointed out by Fisher [1986]:

"When two elements [connected by *or*] do not agree in number, or when one or both of them is a personal pronoun, the verb is governed by the element which is nearer."

The astute reader is invited to discover that this sentence breaks the rule which it states.

Let us return to the conjunction *and* of almost polymorphic type $x^r x x^\ell$. For example, when $x = \mathbf{o}$, we have

did you see her and me ?

$$(\mathbf{q}_2\, \mathbf{i}^\ell\, \pi^\ell)\pi_2(\mathbf{i}\, \mathbf{o}^\ell)\ [\mathbf{o}\, (\mathbf{o}^r\, \mathbf{o}\, \mathbf{o}^\ell)\mathbf{o} \quad \to \quad \mathbf{q}_2$$

(I must confess that I am completely mystified why many people say *her and I*. I can imagine reasons for replacing *me* by *I*, but why only in the second conjunct, and why not replace *her* by *she*? Logic would seem to suggest the alternatives **I and her* or **me and she*?)

In the type $x^r x x^\ell$, x need not be simple. For example, when $x = \mathbf{i}\mathbf{o}^\ell$ (hence $x^\ell = \mathbf{o}^{\ell\ell}\mathbf{i}^\ell$), we have

whom can you see and hear − ?

$$(\overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{q}^\ell)(\mathbf{q}_1\, \mathbf{j}^\ell\, \pi^\ell)\, \pi_2([\, \mathbf{i}\, \mathbf{o}^\ell)\, (\mathbf{o}\, \mathbf{i}^r\, \mathbf{i}\, \mathbf{o}^\ell\, [\mathbf{o}^{\ell\ell}\, \mathbf{i}^\ell)(\mathbf{i}\, \mathbf{o}^\ell) \quad \to \quad \overline{\mathbf{q}}$$

We may even have $x = \pi\mathbf{i}\mathbf{o}^\ell$, as in

whom can (you see) and (I hear) − ?

However, some values of x run against the block constraint, for example, $x = \pi\mathbf{i}$. One does not say

**did he come and I see her .*

$$(\mathbf{q}_2\, \mathbf{i}^\ell\, \pi_3^\ell)\ [\pi_3\, [\mathbf{i}(\mathbf{i}^r\, \pi^r\, \pi\, \mathbf{i}\, \mathbf{i}^\ell\, \pi^\ell)\pi_1\, (\mathbf{i}\, \mathbf{o}^\ell)\mathbf{o} \quad \to \quad \mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell[\pi_3[\mathbf{i}.....$$

Hence neither does one say

**whom did he come and I see − ?*

To eliminate such expressions as

**whom did you sleep and hear − ?*

with $x = \mathbf{i}$, mainstream grammarians have introduced a "coordinate structure constraint". We can express this algebraically, by letting the type of the coordinate conjunction be $x^r y\, [x^\ell$. Thus

**whom did you sleep and ⋯ ?*

$$(\overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{q}^\ell)(\mathbf{q}_2\, \mathbf{i}^\ell\, \pi^\ell)\pi_2\, [\mathbf{i}\, (\mathbf{i}^r\, \mathbf{i}\, [\mathbf{i}^\ell) \cdots \quad \to \quad \overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}[\mathbf{i}^\ell \cdots$$

runs against the block constraint. There are some exceptions to this analysis. In such expressions as

whom did you come and see − ?
whom did you go and hear − ?

the conjunction *and* may perhaps be thought of as replacing *to*.

For the case $x = \mathbf{s}\mathbf{o}^\ell$, hence $x^r = \mathbf{o}\mathbf{s}^r$ and $x^\ell = [\mathbf{o}^{\ell\ell}\mathbf{s}^\ell$, we may consider the noun phrase

the girl whom I saw and you heard −

$$(\overline{\mathbf{n}}_1\, \mathbf{n}_1^\ell)\, [\mathbf{n}_1(\mathbf{n}_1^r\, \mathbf{n}_1[\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{s}^\ell)\, \pi_1(\pi^r[\mathbf{s}_2\, \mathbf{o})(\mathbf{o}\, \mathbf{s}^r\, \mathbf{s}\, \mathbf{o}^\ell[\mathbf{o}^{\ell\ell}\, \mathbf{s}^\ell)\, \pi_2(\pi^r\, \mathbf{s}_2\, \mathbf{o}^\ell) \quad \to \quad \overline{\mathbf{n}}_1$$

and the calculation runs smoothly.

On the other hand, in case $x = \mathbf{q}$, consider

**whom shall I sing or will you dance with − ?*

$$(\overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{q}^\ell)([\mathbf{q}_1\, \mathbf{j}^\ell\, \pi^\ell)\pi_1\, \mathbf{i}(\mathbf{q}^r\, \mathbf{q}\, [\mathbf{q}^\ell) \cdots \qquad \to \quad \overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}[\mathbf{q}^\ell \cdots$$

where we run against the block constraint.

Finally, let us look at a more complicated example with six left square brackets, labeled with numerical subscripts from 1 to 6:

**whom did I see photos of friends and paintings of − ?*

$$(\overline{\mathbf{q}}[_1\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{q}^\ell)(\mathbf{q}_2\, \mathbf{i}^r\, \pi^r)\pi_1(\mathbf{i}[_2\mathbf{o}^\ell)[_3\mathbf{n}_2(\mathbf{n}_2^r\mathbf{n}_2[_4\mathbf{o}^\ell)[_5\mathbf{n}_2)(\mathbf{o}^r\mathbf{o}[_6\mathbf{o}^\ell)\mathbf{n}_2(\mathbf{n}_2^r\mathbf{n}_2\mathbf{o}^\ell)$$

Here the first bracket belongs to the double adjoint. The second bracket prevents the question from ending after *see*; but it can be removed by the unblocking maneuver, replacing $[\mathbf{o}^\ell$ by $\mathbf{n}_2^\ell$. The third bracket ensures that *did I see photos* is not a constituent. The fourth bracket prevents the question from ending after the first occurrence of *of*; it too can be unblocked, replacing $[\mathbf{o}^\ell$ by $\mathbf{n}_2^\ell$. The fifth bracket ensures that *did I see photos of friends* is not a constituent. The sixth square bracket is due to our type of *and*, which could also be replaced by $\mathbf{n}_2^r\mathbf{n}_2[\mathbf{n}_2^\ell$. Anyway, after unblocking, the calculation leads to

$$\begin{aligned} &\to\ \overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\mathbf{n}_2^\ell\, [\mathbf{n}_2\mathbf{n}_2^r\mathbf{n}_2\mathbf{n}_2^\ell\, [\mathbf{n}_2\mathbf{n}_2^r\mathbf{n}_2\, [\mathbf{n}_2^\ell \cdots \\ &\to\ \overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\, [\mathbf{n}_2^\ell \cdots \end{aligned}$$

and we run against the block constraint.

Subordinate conjunctions introduce subordinate clauses of type $\mathbf{i}^r\mathbf{i}$ and are followed by declarative sentences, hence have type $\mathbf{i}^r\mathbf{i}\, [\mathbf{s}^\ell$, where again we have decided to retain the left square bracket in front of $\mathbf{s}^\ell$. This will ensure that

**whom did John come when I called − ?*

$$(\overline{\mathbf{q}}\, [\hat{\mathbf{o}}^{\ell\ell}\, \mathbf{q}^\ell)(\mathbf{q}_2\, \mathbf{i}^\ell\, \pi^\ell)\mathbf{n}\, [\mathbf{i}(\mathbf{i}^r\, \mathbf{i}\, [\mathbf{s}^\ell)\pi_1(\pi^r\mathbf{s}_2\mathbf{o}^\ell)$$

$$\to\ \overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}\, [\mathbf{s}^\ell \cdots$$

runs against the block constraint.

21. Summary of constraints.

In Chapter 3, we had briefly introduced square brackets to represent what corresponds to deep structure in our system. To simplify the notation, we might have omitted all the right square brackets, as these could easily be restored. For most of our discussion, square brackets were replaced by the visually more appealing underlinks; but left square brackets were systematically retained, as kinds of punctuation marks, to prevent premature contractions.

In more recent chapters, we decided to restore some other left brackets as well. Thus, in Chapter 18, we restored all left brackets in front of double adjoints, wherever they had appeared in the first place. (They did not occur when a double adjoint was to be canceled by a triple one.) Later, in Chapter 18 we restored also left brackets in types of the form $\mathbf{n}_i^r \mathbf{n}_i [\mathbf{s}^\ell$, which are assigned to subject relative pronouns *who/which/that* or which appear inside the type assigned to the possessive relative pronoun *whose.* Finally, in Chapter 20, we also restored left brackets in types of the form $x^r y[x^\ell$, which are assigned to coordinate conjunctions such as *and* or *or*, and in types of the form $\mathbf{i}^r \mathbf{i}[\mathbf{s}^\ell$, which are assigned to subordinate conjunctions such as *when.*

In summary, we have restored left brackets in types of the form $[x^{\ell\ell}$, $x^r x[\mathbf{s}^\ell$ and $x^r y[x^\ell$. Let us call the occurrence of a simple type "blocked", when it is preceded by a remaining or restored left bracket. The *block constraint* asserts:

BLOCK CONSTRAINT. When a calculation arrives at two consecutive blocked simple types $[x[y$, it should be aborted.

I believe that the remaining or restored left brackets should not be regarded as elements of a pregroup with additional structure, as was suggested by Fadda [2002], but that they serve as some kind of punctuation marks. If so, the block constraint should not be viewed as a grammatical rule, but as a device for avoiding a processing difficulty, like that arising when one exceeds the temporary storage of nine chunks of information. I am not quite sure why this should be the case and how much of the processing difficulty is confined to modern English. (For example, two consecutive double adjoints occur in French and other Romance languages, see e.g. Bargelli and Lambek [2001], Casadio and Lambek [2001].)

I can't honestly say that I have covered all the constraints suggested by linguists. McCawley (Chapter 15) discusses a number of constraints proposed by Ross [1967]. Although many of his examples can also be handled by our block constraint, there are problems. Here is just one example, in which *rumour* has been assigned the type $\mathbf{n}_1 \overline{\mathbf{s}}^\ell$, to indicate that it may require an indirect sentential complement

$$^*\textit{what did he report a rumour that she was paying} - ?$$
$$(\overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^\ell)(\mathbf{q}_2\,\mathbf{i}^\ell\,\pi^\ell)\pi_3(\mathbf{i}\,[\mathbf{o}^\ell)(\overline{\mathbf{n}}_1\mathbf{n}_1^\ell)(\mathbf{n}_1\overline{\mathbf{s}}^\ell)(\overline{\mathbf{s}}\mathbf{s}^\ell)\pi_3(\pi_3^r\mathbf{s}_2\mathbf{p}_1^\ell)(\mathbf{p}_1\mathbf{o}^\ell)$$

$$\rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,[\mathbf{o}^\ell......$$

running against the block constraint. Unfortunately, our unblocking maneuver allows $[\mathbf{o}^\ell$ to be replaced by $\overline{\mathbf{n}}_1^\ell$ without a left bracket, and then the calculation could continue:

$$\rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\,\overline{\mathbf{n}}_1^\ell\overline{\mathbf{n}}_1\,\mathbf{n}_1^\ell\mathbf{n}_1\,\overline{\mathbf{s}}^\ell\overline{\mathbf{s}}\,\mathbf{s}^\ell\,\pi_3\pi_3^r\,\mathbf{s}_2\,\mathbf{p}_1^\ell\mathbf{p}_1\,\mathbf{o}^\ell\ \rightarrow\ \overline{\mathbf{q}}$$

This would seem to make the above question acceptable after all!

Perhaps we were wrong in assigning the type $\mathbf{n}_1\overline{\mathbf{s}}^\ell$ to *rumour*, and its type should have been $[\mathbf{n}_1$, the left bracket ensuring that *a rumour* is not a constituent. The word *that* would then have to be assigned the type $\mathbf{n}_1^r\mathbf{n}_1\,[\mathbf{s}^\ell$, in accordance with our decision to restore left brackets in types of the form $x^r x\,[\mathbf{s}^\ell$. Moreover, the indefinite article *a* would have to have type $[\overline{\mathbf{n}}_1\mathbf{n}_1^\ell$, to ensure that *report a rumour* is not a constituent. The calculation would then continue as follows:

$$\rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\overline{\mathbf{n}}_1^\ell\,[\overline{\mathbf{n}}_1\,\underbrace{\mathbf{n}_1^\ell\,[\underbrace{\mathbf{n}_1\mathbf{n}_1^r}\mathbf{n}_1}\,[\mathbf{s}^\ell\ \rightarrow\ \overline{\mathbf{q}}\,[\hat{\mathbf{o}}^{\ell\ell}\overline{\mathbf{n}}_1^\ell\,[\overline{\mathbf{n}}_1\,[\mathbf{s}^\ell\ldots\ldots$$

and we would run against the block constraint.

Anyway, it seems to me that the objectionable question becomes almost acceptable if we replace *report* by *spread*:

what did he spread a rumour that she was paying $-$?

According to McCawley, even Ross is willing to admit the similarly constructed question:

whom does he hold the belief that he can see -

on the grounds that the verb phrase *hold the belief* can be treated like the transitive verb *believe* of type $\mathbf{i}\overline{\mathbf{s}}^\ell$. Perhaps *spread a rumour* could similarly be treated like the transitive verb *rumour.*

Finally, let me point out that my judgement does not always agree with that of others. It seemed to me that

whom does she know whether I expect $-$?

$$(\overline{\mathbf{q}}[\hat{\mathbf{o}}^{\ell\ell}\,\underbrace{\mathbf{q}^\ell)(\mathbf{q}_1}\,\mathbf{i}^\ell\,\underbrace{\pi_3^\ell)\pi_3}\,(\mathbf{i}\,\underbrace{\mathbf{t}^\ell)(\mathbf{t}}\,\mathbf{s}^\ell)\,\underbrace{\pi_1(\pi_1}\,{}^r\mathbf{s}_1\,\mathbf{o}^\ell)$$

should be acceptable, in analogy to

whom does she know [that] I expect $-$?

However, people have told me that it is not acceptable. If so, the type of *whether* should be $\mathbf{t}[\mathbf{s}^\ell$, forcing the calculation to be aborted by our block constraint. Evidently, the present discussion of constraints is not complete.

VI. Semantic Considerations.

22. Quantifiers.

Montague semantics starts with the basic assumption that types represent sets. While this is not easily justified by our choice of basic types, let me simplify matters for expository purposes and assume that there are only two basic types:

$N =$ set of human beings,

$S =$ set of truth values.

Given any types x and y, here not distinguished from the sets they represent, we interpret xy as the Cartesian product $x \times y$ and both $x^r y$ and yx^ℓ as the set of functions from x to y, which mathematicians denote by y^x.

A problem arises when we look at xyz^ℓ, which can be interpreted as $(x \times y)^z$ or as $x \times (y^z)$. The ambiguity disappears if we distinguish xyz^ℓ from $x[yz^\ell$, as we have often done in earlier chapters for syntactic reasons, but will do so now with a semantic motivation.

In our rudimentary pregroup grammar we analyze

$$\begin{array}{c} \textit{Alice likes Bob .} \\ N\,(N^r\,S\,N^\ell)\,N \;\to\; S \end{array}$$

The names *Alice* and *Bob* both represent elements of N, and the predicate *likes Bob* of type

$$(N^r S\, N^\ell)N \;\to\; N^r S$$

is interpreted as a function from N to S, hence as an element of S^N. Therefore *likes* represents a function from N to S^N, that is, an element of

$$(S^N)^N \simeq S^{N\times N},$$

usually called a "binary relation". In the language of mathematical logic, we abbreviate *Alice likes Bob* as LAB.

Now look at the following sentences:

everybody likes Bob,
Alice likes somebody,
everybody likes somebody.

Following a proposal by Claudia Casadio, we do not assign the type N to the subject *everybody* and to the object *somebody* (after all, they are not nouns), but the types $SS^\ell N$ and $NS^r S$ respectively. Thus we are led to analyze the first two sentences as follows:

$$\begin{array}{c} \textit{everybody likes Bob,} \\ (S\,S^\ell\,N)(N^r\,S\,N^\ell)N \;\to\; S \end{array}$$

$$\begin{array}{c} \textit{Alice likes somebody.} \\ N\,(N^r\,S\,N^\ell)\,(N\,S^r\,S) \;\to\; S \end{array}$$

However, the third sentence can be analyzed in two different ways:

$$\begin{array}{c} \textit{everybody likes somebody} \\ (S\, S^{\ell}\, N)(N^{r}\, S\, N^{\ell})(N\, S^{r}\, S) \end{array} \quad \to \; S$$

or

$$\begin{array}{c} \textit{everybody likes somebody}\ . \\ (S\, S^{\ell}\, N)(N^{r}\, [S\, N^{\ell})(N\, S^{r}\, S) \end{array} \quad \to \; S$$

In the language of mathematical logic, the first two sentences are translated as

$$\forall_{x \in N}\, LxB,$$

and

$$\exists_{y \in N}\, LAy.$$

respectively. The third sentence gives rise to two different translations:

$$\exists_{y \in N}\, \forall_{x \in N}\, Lxy,$$

and

$$\forall_{x \in N}\, \exists_{y \in N}\, Lxy,$$

associated to the two different ways of analyzing it. I have heard linguists assert that both interpretations are plausible, but mathematicians and teachers of elementary logic courses will insist that only the second interpretation is correct, even though it involves a deferred contraction, as indicated by the left bracket.

We can interchange the order of the two quantifiers and their types to obtain

$$\begin{array}{c} \textit{somebody likes everybody} \\ (S\, S^{\ell}\, N)(N^{r}\, [S\, N^{\ell})(N\, S^{r}\, S) \end{array}$$

with the translation

$$\exists_{x \in N}\, \forall_{y \in N}\, Lxy.$$

It seems to me that, if we want the translation

$$\forall_{y \in N}\, \exists_{x \in N}\, Lxy,$$

we should replace *everybody* by *anybody*. In other contexts, these two words may be synonymous, for example in

$$\begin{array}{l} \textit{Alice likes anybody}, \\ \textit{anybody likes Bob}, \\ \textit{anybody likes somebody} \end{array}$$

anybody can be replaced by *everybody* without change in meaning. The last of these sentences does not admit the translation with $\exists_{x \in N}$ in front.

Now let us abandon the rudimentary pregroup grammar and return to the more realistic pregroup grammar developed in this book. Then some of the sentences analyzed above should be re-analyzed as follows:

$$\begin{array}{c} \textit{Alice likes Bob,} \\ \mathbf{n}\,(\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^\ell)\,\mathbf{n} \end{array} \quad \to \mathbf{s}_1$$

$$\begin{array}{c} \textit{Alice likes somebody,} \\ \mathbf{n}\,(\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{s}^r\,\mathbf{s}) \end{array} \quad \to \mathbf{s}$$

or

$$\begin{array}{c} \textit{Alice likes somebody,} \\ \mathbf{n}\,(\pi_3^r\,\mathbf{s}_1\,\mathbf{j}^\ell\,[\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \mathbf{s}_1$$

$$\begin{array}{c} \textit{Alice likes everybody,} \\ \mathbf{n}\,(\pi_3^r\,\mathbf{s}_1\,\mathbf{j}^\ell\,[\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \mathbf{s}_1$$

$$\begin{array}{c} \textit{Alice likes anybody.} \\ \mathbf{n}\,(\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{s}^r\,\mathbf{s}) \end{array} \quad \to \mathbf{s}$$

Although the object quantifiers *everybody* and *anybody* here mean the same, they give rise to different interpretations in negative sentences and questions, as illustrated by the following examples:

$$\begin{array}{c} \textit{Alice doesn't like everybody,} \\ \mathbf{n}(\pi_3^r\,\mathbf{s}_1\,\mathbf{i}^\ell)(\,[\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \mathbf{s}_1$$

$$\begin{array}{c} \textit{Alice doesn't like anybody,} \\ \mathbf{n}(\pi_3^r\,\mathbf{s}_1\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{s}^r\,\mathbf{s}) \end{array} \quad \to \mathbf{s}$$

$$\begin{array}{c} \textit{does Alice like everybody ?} \\ (\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_3^\ell)\mathbf{n}(\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{i}^r\,\mathbf{i}) \end{array} \quad \to \mathbf{q}_1$$

$$\begin{array}{c} \textit{does Alice like anybody ?} \\ (\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_3^\ell)\mathbf{n}(\mathbf{i}\,\mathbf{o}^\ell)(\mathbf{o}\,\mathbf{q}^r\,\mathbf{q}) \end{array} \quad \to \mathbf{q}$$

Here *everybody* of type $\mathbf{oi}^r\mathbf{i}$ and *anybody* of type $\mathbf{os}^r\mathbf{s}$ or $\mathbf{oq}^r\mathbf{q}$ are both universal quantifiers in object position, but with different scope.

The type assignment is a bit more complicated for universal quantifiers in subject position:

$$\begin{array}{c} \textit{everybody / anybody likes Bob ,} \\ (\mathbf{s}\,\mathbf{s}^\ell\,\pi_3)(\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^\ell)\mathbf{n} \end{array} \quad \to \mathbf{s}$$

$$\begin{array}{c} \textit{does everybody like Bob ?} \\ (\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_3^\ell)(\pi_3\,\mathbf{i}\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{n} \end{array} \quad \to \mathbf{q}_1$$

$$\begin{array}{c} \textit{does anybody like Bob ?} \\ (\mathbf{q}_1\,\mathbf{i}^\ell\,\pi_3^\ell)(\pi_3\,\mathbf{i}\,\mathbf{q}^r\,\mathbf{q}\,\mathbf{i}^\ell)(\mathbf{i}\,\mathbf{o}^\ell)\mathbf{n} \quad \to \quad \mathbf{q} \end{array}$$

Thus *anybody* has type $\mathbf{os}^r\mathbf{s}$ or $\mathbf{oq}^r\mathbf{q}$ in object position and $\mathbf{ss}^\ell\pi_3$ or $\pi_3\mathbf{iq}^r\mathbf{qi}^\ell$ in subject position. Other types may still be required for different contexts, e.g. in

is it true that anybody likes Bob?

We shall not pursue this topic any further and leave it to the interested reader to re-examine

somebody likes anybody

etc.

Everybody always denotes a universal quantifier and so does *anybody*, although the latter may sometimes be mistaken for an existential one, since, for example,

does Alice like anybody

has the same meaning as

does Alice like somebody.

Teachers of elementary logic courses would surely insist that *somebody* always denotes an existential quantifier. This is not so in the conditional sentence

if somebody hates Bob, he likes Alice,

which is translated as

$$\forall_{x\in N}\ (HxB \Rightarrow LxA).$$

The semantics explored in this chapter won't explain this. As far as I know, the Discourse Representation of Hans Kamp will do so. (See also Groenendijk and Stabler [1991].)

If the subject quantifier *everybody* is assigned the type $\mathbf{ss}^\ell\pi_3$, the singular determiners *every* and *each* should have type $\mathbf{ss}^\ell\pi_3\mathbf{n}_1^\ell$, but the plural determiner *all* requires the type $\mathbf{ss}^\ell\pi_2\mathbf{n}_2^\ell$. If the subject quantifier *somebody* is assigned the type $\mathbf{ss}^\ell\pi_3$, the determiner *some* may have type $\mathbf{ss}^\ell\pi_k\mathbf{n}_i^\ell$ with $(k,i) = (3,1)$ or $(2,2)$ to account for

some person snores,
some people snore.

Now *some* in the first of these sentences may be replaced by the indefinite article *a*, to which we have hitherto assigned the type $\overline{\mathbf{n}}_1\mathbf{n}_1^\ell$. I will refrain from changing its type to $\mathbf{ss}^\ell\pi_3\mathbf{n}_1^\ell$ at this stage of our discussion. According to Bertrand Russell's theory of description, even the definite article *the* may require the quantifier $\exists\,!$ (= there is a unique) in its logical translation, as in Russell's favourite example

Scott is the author of Waverley .

(I was amused to find that this example, after a free translation into German, was translated back into English as *Goethe is the poet-author of Faust.*)

23. Anaphora.

Although pregroup grammars have proved themselves useful for investigating a number of problems of interest to linguists, some problems can easily be approached without their help. On the boundary of syntax and semantics is the question: when can a pronoun represent a noun phrase in the same sentence? In other words, when can a pronoun denote the same person or object as is denoted by the noun phrase? Consider, for example, the sentence

(i) $$\textit{although John}_1 \textit{ did not like her}_2, \textit{ he}_1 \textit{ kissed } (\textit{the girl})_2.$$

Here *he* may refer to John and *her* may refer to the girl mentioned. (We have adopted the usual practice of using numerical subscripts to indicate a common reference.)

In answer to this question: first of all, the pronoun must agree with the represented noun phrase in the features person, number and gender; in our example (i), the third person singular and masculine (1) versus feminine (2). In English, gender is a purely semantic feature, e.g. we know that John is male and that the girl is female.

This is not necessarily so in other languages. In some, the morphology offers a clue to the gender, but this may be misleading when taken literally. For example, in Latin the ending $+a$ usually indicates feminity, but *agricola* is one of a small number of exceptions. In Hebrew, the ending $+im$ usually indicates the masculine plural, but I recall being quite embarrassed after opening a door displaying the word *nashim*. In French, the moon is denoted by a feminine noun and in German by a masculine one; in both languages a pronoun referring to the moon must display the corresponding gender. In both French and German, the word translating English *person* is feminine, confusingly even when referring to a male. (French *personne* may be masculine, but then it translates the English *nobody*.)

In addition to the semantic criterion of feature agreement, English also requires certain purely syntactic criteria that must be satisfied for a pronoun to represent a noun phrase in the same sentence. For example, in (i) the pronoun *her* is allowed to represent *the girl* because it is properly contained in a subordinate clause that does not contain the noun phrase. Similarly, *he* can represent *John*, because the former is properly contained in a sentential constituent not containing the latter. A competing explanation would be that *John* is properly contained in a subordinate clause, but that *he* occurs after this clause.

The syntactic criteria have to be formulated rather carefully. For, if we reverse the order in (i) and consider instead

(ii) $$\textit{he kissed } (\textit{the girl})_2, \textit{ although John did not like her}_2,$$

he can no longer represent *John*, although *her* can still represent *the girl*. Surprisingly, the syntactic criteria can be formulated without a thorough grammatical analysis of the sentence, such as the one we have advocated here using pregroup grammars. It suffices to recognize certain key constituents of the sentence.

By a *key constituent* we will understand one of the following:

(0) a sentential constituent,

(1) a noun phrase (possibly an indirect sentence),

(2) a subordinate clause,

(3) a quasi-sentence (like the phrase *him to see her*),

(4) a spatio-temporal propositional phrase.

Here is how I would formulate the hopefully necessary and sufficient condition for allowing a pronoun to represent a noun phrase in the same sentence (avoiding the currently favoured technical notion of "C-command").

(I) The pronoun agrees with the noun phrase in the features person, number and gender, and either

(II) the pronoun is properly contained in a key constituent (1) to (4), from which the noun phrase is excluded,

or

(III) the noun phrase is properly contained in a key constituent (0) to (3) and the pronoun occurs after this constituent.

We will make the blanket assumption that these criteria also apply to any key constituent of the sentence in question, in particular to any sentential constituent containing both the pronoun and the noun phrase. Consequently we can state

(IV) if the pronoun and the noun phrase are properly contained in the same minimal key constituent, then the former cannot represent the latter.

To make the above conditions work, some minor editing of the key constituents may be required. For the following illustrations, I have used the notation (IIk) to indicate an application of criterion (II) with key constituent (k).

(II1) $John_1$ *kissed* (*the girl* he_1 *liked*);
(his_1 *girlfriend*) *kissed* $John_1$;
$John_1$ *said* ([*that*] he_1 *liked Jane*);
(*that* he_1 *liked Jane*) *was clear to* $John_1$.

(II2) $John_1$ *returned* (*when* he_1 *was hungry*);
although he_1 *liked girls*, $John_1$ *never married.*

(II3) $John_1$ *let* $Jane_2$ ($\#_2$ *kiss* him_1).

Here some editing was required to eliminate

(iii) $$John_1 \ let \ {}^*him_1 \ (\#_1 \ kiss \ Jane)$$

and

(iv) $$John_1 \ let \ Jane_2 \ (\#_2 \ kiss \ {}^*her_2),$$

by inserting an invisible copy # of the object of *let* into the infinitival phrase, turning it into what I will call a "quasi-sentence". Both (iii) and (iv) now violate (IV). Similarly we justify

$$John_1 \ wanted \ [for] \ Jane_2 \ (\#_2 \ to \ kiss \ him_1).$$

Here the complementizer *for* is optional, but it becomes obligatory if *wanted* is replaced by *wished*. When the complementizer is present, an alternative interpretation is possible: we may regard *for Jane to kiss him* as an abstract noun phrase. The complementizer becomes obligatory when this noun phrase appears in subject position. In principle, we could also admit

$$John_1\ wanted\ ([for\ him_1]\ \#_1\ to\ kiss\ Jane),$$

but then the expression *for him* is usually omitted. The remaining invisible noun phrase $\#_1$ resembles Chomsky's empty category PRO and guards against

$$John_1\ wanted\ (\#_1\ to\ kiss\ {}^*him_1),$$

which would violate (IV).

$$\begin{array}{ll}(\text{II4}) & John_1\ saw\ (an\ eagle)_2\ (\#_2\ above\ him_1), \\ & (\#_2\ above\ him_1),\ John_1\ saw\ (an\ eagle)_2.\end{array}$$

Again, the invisible noun phrase $\#_2$ helps to eliminate

$$John_1\ sat\ (\#_1\ beside\ {}^*him_1).$$

Next consider

$$(\text{III0}) \qquad (John_1\ saw\ Jane_2)\ and\ [he_1]\ kissed\ her_2.$$

The editorial insertion of he_1 will help to eliminate

$$(John_1\ saw\ Jane)\ and\ ([he_1]\ kissed\ {}^*him_1),$$

which would violate (IV).

(III0) $(John_1\ saw\ Jane_2)\ and\ she_2\ kissed\ him_1$,

(III1) $(John's_1\ girlfriend)\ kissed\ him_1$,

$([for\ her_2]\ \#_2\ to\ see\ John_1)\ Jane_2\ visited\ him_1$,

(III2) $(when\ John_1\ sleeps)\ he_1\ snores$.

Note that (III3) won't come up, since a quasi-sentence (without complementizer) cannot occur as a subject.

The chapter on anaphora has been presented as independent of the pregroup grammar developed earlier. This is not to say that it cannot be reformulated in pregroup terminology or that the special symbol # introduced here ought not to be incorporated into our earlier pregroup description.

24. Kinship grammar.

Although semantic notions have occasionally cropped up in the preceding pages, I don't feel too comfortable with the formal treatments of semantics currently available. A case has been made in Chapter 21 for linking the fashionable Montague semantics to pregroups, but my personal preference is for what I would like to call "Capulet semantics".

My favorite semantic field is that of kinship relations, traditionally beloved by older women of the tribe and by anthropologists. Its proper description requires a surprising amount of mathematical insight, even if not an acquaintance with pregroups. As Chomsky once said:

"[kinship systems] may be the kind of mathematics you can create if you don't have formal mathematics. The Greeks made up number theory, others made up kinship systems".

Consanguineous kinship relations in English and other languages can be described with the help of the basic binary relations

P = parent,
C = child,
S = sibling,
M = equality between males,
F = equality between females,

and with the help of what mathematicians call the "relative product" of two binary relations. For example, given persons x and y,

$$\mathrm{xPSMy}$$

means:

there are persons u and v such that

$$\mathrm{xPu \text{ and } uSv \text{ and } vMy},$$

that is to say:

[*one of*] x'*s parents is* u *and* [*one of*] u'*s siblings is* v *and* v *is male and equal to* y.

In other words, xPSMy means

y *is a male sibling* (*brother*) *of a parent* (*mother or father*) *of* x.

Putting this more concisely, we say

[*one of*] x'*s uncles is* y

or

y *is an uncle of* x.

Note that speakers of modern English do not care whether the intermediary u is male or female. Old English did distinguish between the mother's brother and the father's brother; and many modern languages still do, for example Hindi, which also distinguishes between the father's older and younger brother. In fact, our present word *uncle* is derived from Latin

avunculus = *little grandfather*,

originally referring to the mother's brother, who became her legal guardian after the death of her father. The present meaning of the word *uncle* can be expressed by the rewrite rule

PSM → *uncle*.

Our word *nephew* is similarly derived from an old Indo-European word for *grandson*. In fact, in modern Italian, *nepote* can still mean either nephew or grandson. For modern English, we have the rewrite rule

SCM → *nephew*.

Looking at an article by Goodenough on "Yankee kinship terminology", I was puzzled by the observation that, in our present terminology,

PPSM → *great uncle*,

but

SCCM → *grandnephew*.

Although this distinction is not shared by all English speakers, its local occurrence still requires an explanation. Of course, the prefixes *grand+* and *great* were originally used to describe direct ancestors or descendents. Here are a few rewrite rules:

$$\begin{array}{lll} \text{PM} \rightarrow \textit{father}, & \text{PF} \rightarrow \textit{mother}, \\ \text{CM} \rightarrow \textit{son}, & \text{CF} \rightarrow \textit{daughter}, \\ \text{PP} \rightarrow \begin{cases} \textit{grand}+\ \text{P} & \text{before M or F} \\ \textit{great}\ \text{P} & \text{otherwise.} \end{cases} \\ \text{CC} \rightarrow \begin{cases} \textit{grand}+\ \text{C} & \text{before M or F} \\ \textit{great}\ \text{C} & \text{otherwise.} \end{cases} \end{array}$$

(Note that French would translate *grand* or *great* by the French word for *little* in the last rule.)

We look at some sample calculations:

PPPM → *great* PPM → *great grand+* PM → *great grandfather*,

CCCF → *great* CCF → *great grand+* CF → *great granddaughter*,

PPSM → *great* PSM → *great uncle*.

When we turn to SCCM, none of the above rules apply. We therefore ought not be surprised that Lounsbury's Yankees have picked the rule

$$\begin{array}{lll} \text{SCC} & \rightarrow & \textit{grand}+\ \text{SC before M or F} \\ & \rightarrow & \textit{great}\ \text{SC otherwise.} \end{array}$$

It follows that

$$\begin{array}{lll} \text{SCCCM} & \rightarrow & \textit{great}\ \text{SCCM} \\ & \rightarrow & \textit{great grand}+\ \text{SCM} \\ & \rightarrow & \textit{great grandnephew}. \end{array}$$

Compared to many other languages that have been studied by anthropologists, modern English is quite poor in such rewrite rules and even in kinship descriptions. The only relevant consanguineous kinship descriptions in modern English are

$$\text{P}^m\text{SC}^n\text{G},\ \text{P}^{m+1}\text{G},\ \text{C}^{n+1}\text{G},$$

where G=M or F and $m, n \geq 0$.

In many languages it is important to express intermediate genders. For example, in the language spoken on the Trobriand islands in the Western Pacific, our kinship description for first cousins by

$$\text{PSCG (G = M or F)}$$

should be replaced by

$$G_0PG_1SG_2CG_3\ (G_i = \text{ M or F}).$$

In spite of the theoretical possibility of describing $2^4 = 16$ different kinds of cousins, the Trobriand language has only 8 different consanguineous kinship terms, six of which can describe certain cousins. Here are the primary meanings of these terms, as observed by Malinowski [1932]:

tabu = grandparent or grandchild,

ina = mother,

tama = father,

kada = mother's brother or a man's nephew or niece,

luta = sibling of opposite sex,

tuwa = older sibling of same sex

bwada = younger sibling of same sex,

latu = child.

Malinowski claimed that these same terms had numerous secondary meanings, but his analysis was challenged by Leach [1958], who said:

"when terms are projected onto a genealogical diagram, the underlying logic is utterly incomprehensible".

Somewhat later, Lounsbury [1965] came to Malinowski's defence and showed that a small number of mathematical postulates, in the form of rewrite rules, would suffice to explain the derivation of the secondary meanings of the basic kinship terms. Following Bhargava and Lambek [1995], his postulates may be stated as follows, where G=M or F:

PFC → S,

GSG → G after P or C,

MSF → MPF after P,

FSM → FCM before C,

PGP → P,

CGC → C,

FSM → M when no other rule applies,

MSF → M likewise.

Note the symmetry between P and its converse C.

Now let us calculate the Trobriand terms for what we would call "cousins":

$$
\begin{aligned}
G_0P\underline{FSF}CG_3 &\rightarrow G_0\underline{PFC}G_3 \rightarrow G_0SG_3 \\
&\rightarrow \begin{cases} \textit{tuwa} \text{ or } \textit{bwada} & \text{if } G_0 = G_3 \\ \textit{luta} & \text{if } G_0 \neq G_3 \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
G_0P\underline{FSM}CG_3 &\rightarrow G_0\underline{PFC}MG_3 \rightarrow G_0SMCG_3 \\
&\rightarrow \begin{cases} MCG_3 \quad (\text{if } G_0 = M) \rightarrow \textit{latu} \\ \textit{tabu} \quad \text{if } G_0 = F \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
G_0P\underline{MSF}CG_3 &\rightarrow G_0PM\underline{PFC}G_3 \rightarrow G_0PMCG_3 \\
&\rightarrow \begin{cases} G_0PM \rightarrow \textit{tama} & \text{if } G_3 = M \\ G_0PMPF \rightarrow \textit{tabu} & \text{if } G_3 = F \end{cases}
\end{aligned}
$$

$$
G_0P\underline{MSM}CG_3 \rightarrow G_0\underline{PMC}G_3 \rightarrow ?
$$

There is no Trobriand translation for the last kind of cousin. It appears that PMC is not viewed as a relative, and anthropologists have suspected that, at one time Trobrianders did not understand the rôle of the father in procreation. However, a more probable explanation emerges from the observation by Therroux [1992] that promiscuous behaviour by women is licensed during the month-long annual yam-festival, which would make it difficult to determine the father of any resulting offspring.

25. Capulet versus Montague.

Semantics, like the city of Verona, is the battleground of two rival factions. In connection with quantifiers (Chapter 22), we saw how Montague semantics, also known as *functional* semantics, fits into the framework of pregroup grammar. I like to contrast this with Capulet semantics, better known as *lexical* semantics, to which we will return presently.

At its best, Montague semantics assigns to each linguistic expression its logical form, as Chomsky likes to call it, translating it into the language of higher order intensional logic. Unfortunately, this is not the kind of logic that mathematicians are most familiar with, as it does not satisfy the substitution rule:

$$\text{if } x = y \text{ then } \varphi(x) \Rightarrow \varphi(y),$$

for any formula $\varphi(x)$. This rule would allow us to infer that

$$(\varphi(x) \wedge \neg\varphi(y)) \Rightarrow \neg(x = y),$$

hence to accept the following alibi:

the killer is known to have been here;
John is not known to have been here;
therefore, John is not the killer.

(Recent treatments of intensional logic are due to Gilmore [2005] and Muskens [2007].)

Unfortunately, the logical form does not help to reveal the meaning of words. Translating *John loves Mary* into logical form does not tell us anything about the nature of love, only that *loves* represents a binary relation, as does *hates*. In connection with kinship terminology (Chapter 24), we got a glimpse at Capulet semantics, which defines certain concepts in terms of more primitive ones, hopefully avoiding vicious circles. I am quite impressed by traditional dictionaries, which often succeed in capturing the essential meaning of a word in an amazingly compact form. However, sometimes they get it wrong. For example, several dictionaries I have consulted define *moot* to mean "debatable", although it currently seems to mean "no longer debatable".

Capulet semantics is most successful in certain limited fields, for instance with the part of language dealing with kinship or colour. But even the most primitive concepts may vary from one society to another. Thus Homer refers to the "wine-red sea", and a Japanese friend once told me to look at the "beautiful green sky".

What is the difference between English *eat* and *drink*? Most people would agree that one eats solids and one drinks liquids. Yet, in America one eats soup and drinks tea, whereas in Britain one drinks soup and eats tea (when the latter consists of cucumber sandwiches). Chemists insist that glass is a liquid, yet people don't drink glass, although they may drink a glass (of wine).

Concepts may undergo a change in meaning in our society, even while I am writing this. Must spouses be of different sex? Can a person have more than one spouse at a time? Can a male turn into a female?

Sometimes the social or ethnic basis of semantics tends to be exaggerated. I recall listening to a talk on how the Aymara of Peru differ from neighbouring speakers of Quechua or Spanish in thinking of the future as lying *behind* and the past *in front*. However, this metaphorical interpretation of temporal in terms of spatial position seems to be quite common. Thus, English *before* means both "in front" and "in the past", while Latin *posterior* means both "behind" (in English even as a noun) and "later".

Although Capulet semantics has not been seriously addressed in the present text, the meaning of words was occasionally invoked to motivate their syntactic types. Thus, the following observations turned out to be useful:

Participles of transitive verbs may be treated as adjectives when these verbs express the causation of mental or emotional states.

Doubly transitive verbs usually denote the causation of a state expressed by a simply transitive verb.

Verbs of causation and perception may be complemented by bare infinitives without *to*.

Prepositions expressing a spatial or temporal relationship frequently behave differently than other prepositions.

VII. Background and limitation of pregroup grammars.

26. Remarks on the history of categorial grammar.

Categorial grammar attempts to place most of the grammar of a natural language into the dictionary. It has antecedants in ideas of Charles Sanders Peirce, who conceived of a chemical analogy, viewing a word like *loves* as having two unsaturated bonds: one on the left, looking for a subject, and one on the right, looking for an object. (A modern analogy might point out that a protein is a chain of amino acids.) A systematic attempt to present the grammar of a natural language in similar terms was made by Zellig Harris [1951, 1968].

Officially, categorial grammar, as it is now conceived, was introduced by Ajdukiewicz [1935], who traced the idea back to Lesniewski and Husserl. His pioneering "Syntactic connexion" (sic in English translation) influenced Bar-Hillel [1953], who re-introduced the distinction between left-looking and right-looking categories. I arrived at a similar idea while studying homological algebra, in collaboration with George Findlay, and in preparing my book "Lectures on rings and modules" [1966]. We had noticed that the tensor product of bimodules was accompanied by two binary operations, *over* and *under*, establishing bi-unique correspondences between bilinear mappings

$$A \otimes B \to C, \quad A \to C/B, \quad B \to A\backslash C.$$

Having come across the type theory of Alonzo Church [1940], I realized that a similar technique could be applied to natural languages, and I introduced what I called the "syntactic calculus", a logical system with three binary operations $\otimes$, / (over) and \ (under) satisfying

$$\begin{aligned} x \otimes y \to z \quad & \text{if and only if } x \to z/y \\ & \text{if and only if } y \to x\backslash z. \end{aligned}$$

Here the arrow stands for logical deduction, but we may also take it to be a partial order, when talking about an ordered algebraic system, namely a "residuated semigroup". Later, an identity element was added to the syntactic calculus and the ordered algebraic system became a "residuated monoid".

Anyway, after first realizing the possible application to natural languages, I rushed to the library and found the article by Yehoshua Bar-Hillel in the journal "Language" [1953]. He had essentially developed the same arithmetical notation, except that he only adopted my symbol for "under" in a later publication [1964]. As might have been expected, he turned out to be the referee of my [1958] paper, which I had submitted to the American Mathematical Monthly. He had two objections:

1. my typist had written "categorical" instead of categorial,
2. the system I advocated had no obvious decision procedure.

I did not confess that it was I, not the typist, who was responsible for (1), but I was able to rectify (2) by turning to Gentzen's sequent calculus, which I had learned while teaching a logic course using Kleene's [1952] book. It turned out that, for my "syntactic calculus", Gentzen's three "structural rules", interchange, contraction and weakening, were not needed. In fact, the proof of the crucial cut elimination theorem was easier in the absence of these rules.

A couple of years later, I was invited to a symposium of the American Mathematical Society on the applications of mathematics to linguistics, which was organized by Roman Jacobson [1961]. Hard pressed to find anything new to say, I discussed a non-associative version of my syntactic calculus and pointed out that Bourbaki's [1948] introduction of the tensor product was essentially a categorical version of a Gentzen-style introduction rule. This idea was later to give rise to my notion of a "multicategory" [1989].

At the symposium I learned of a parallel development by Haskell Curry [1961] (see also Curry, Feys and Craig [1959]). He had similarly explored positive intuitionistic propositional logic to arrive at types in linguistics, which I now prefer to call "semantic types". In fact, they later gave rise to "Montague semantics", in view of the so-called "Curry-Howard isomorphism" (see Howard [1980]) between proofs in intuitionistic logic and the lambda terms of Alonzo Church [1941]. Curry himself preferred "combinators", essentially equivalent to lambda expressions, but avoiding bound variables. As I later realized [1980], these combinators are essentially arrows in a "cartesian closed category", an important notion introduced by Bill Lawvere [1969].

The ideas contained in the syntactic calculus did not catch on with the linguistic community of the time (even evoking some antagonism by several graduate students), although two linguists, Noam Chomsky and Robert Lees, tried to encourage me to continue. Still, I could see the new wave of Chomsky's generative-transformational grammar sweep everything else aside and I turned my attention to other things.

Only many years later was there a revival of interest in the syntactic calculus. Theoretical questions were answered by Buszkowski [1982, 1999], Andréka [1994], Pentus [1993] and others. The connection to Montague semantics was pointed out by van Benthem [1982, 1988] and exploited in textbooks by Moortgart [1988], Morrill [1994] and Carpenter [1988]. Moortgat and Oehrle introduced certain unary operations, called "modalities", into the non-associative syntactic calculus to license associativity and commutativity when needed (see Moortgat [1997]).

Another stream of ideas flooded the scene with Girard's "linear logic" (see Troelstra [1992]). This differed from the syntactic calculus in generalizing classical rather than intuitionist logic, in retaining Gentzen's interchange rule and introducing some new "modalities" to license contraction and weakening.

It led Abrusci [1991] and me [1993] to study a classical version of the syntactic calculus, a non-commutative version of Girard's Linear logic, though without his modalities, which I came to call "classical bilinear logic". Neither of us envisaged a linguistic application.

I believe it was in 1997, when Claudia Casadio gave a lecture in Montreal and showed how classical bilinear logic ought to be applied to linguistics. While listening to her talk, it occurred to me that the distinction between the tensor product and its De Morgan dual introduced an unwanted complication. Identifying these two binary operations one obtains a simpler system, which I now call "compact bilinear logic". The word "compact" had been introduced by Max Kelly in a categorical context and was used thus by Michael Barr in his "star-autonomous categories" [1979, 1998], essentially a categorical version of linear logic, but preceding it. Inasmuch as Boolean algebras are algebraic versions of the classical propositional calculus, our pregroups are those of compact bilinear logic.

The new algebraic approach to grammar via free pregroups was first presented at a meeting in Nancy organized by François Lamarche in 1997 and appeared in [Lambek 1999c]. Claudia and I worked on a joint article, to be entitled "A tale of two grammars", but "two" ultimately grew to "four" [2002]. Since then, much theoretical work has been done by Buszkowski [2001, 2002], who proved cut-elimination for compact bilinear logic and showed that pregroup grammars are context-free.

Since 1998, pregroup grammars have been applied to fragments of a number of languages: English, French, German, Italian, Turkish, Arabic, Polish, Latin and Japanese, and there is work in progress on Mandarin and Persian.

27. The mathematical background.

This chapter is intended for readers who want to get their teeth into the underlying mathematics. To start with, here is a review of some elementary algebraic notions.

A *semigroup* is a set with a binary operation, henceforth denoted by juxtaposition, satisfying the *associative law*

$$(xy)z = x(yz).$$

A *monoid* is a semigroup with a *unity element* 1 such that

$$x1 = x = 1x.$$

A set is *partially ordered* if it is equipped with a binary operation $\to$ such that

$x \to x$ (reflexive law),

if $x \to y$ and $y \to z$ then $x \to z$ (transitive law),

if $x \to y$ and $y \to x$ then $x = y$ (antisymmetry law).

For a semigroup or monoid to be partially ordered we also require

if $x \to y$ and $x' \to y'$ then $xx' \to yy'$ (compatibility),

or equivalently

if $x \to y$ then $uxv \to uyv$ (substitutivity).

An element x in a partially ordered monoid is said to have a *left adjoint* x^ℓ if

$$x^\ell x \to 1 \to xx^\ell.$$

The following property of left adjoints is easily proved:

(1) left adjoints are unique,

Proof. Suppose x^λ is another left adjoint of x, then

$$x^\lambda = x^\lambda 1 \to x^\lambda(xx^\ell) = (x^\lambda x)x^\ell \to 1x^\ell = x^\ell.$$

Similarly, $x^\ell \to x^\lambda$ and so, by antisymmetry $x^\lambda = x^\ell$.

A *left pregroup* is a partially ordered monoid in which each element has a left adjoint. One can easily prove

(2) $1^\ell = 1,\ (xy)^\ell = y^\ell x^\ell$.

For example,

$$(y^\ell x^\ell)(xy) = y^\ell(x^\ell x)y \to y^\ell 1 y = y^\ell y \to 1$$

and similarly $1 \to (xy)(y^\ell x^\ell)$, hence $y^\ell x^\ell = (xy)^\ell$, by uniqueness (1).

(3) if $x \to y$ then $y^\ell \to x^\ell$ (contravariance).

Proof. Assume that $x \to y$, then

$$y^\ell = y^\ell 1 \to y^\ell x x^\ell \to y^\ell y x^\ell \to 1x^\ell = x^\ell,$$

making use of the substitution rule.

Applying contravariance once more, we obtain the covariance of double adjunction:

(3′) if $x \to y$ then $x^{\ell\ell} \to y^{\ell\ell}$.

A *group* is a monoid in which each element x has a left *inverse* x^{-1} such that $x^{-1}x = 1$. Surprisingly, the left inverse is also a right inverse, since

$$\begin{aligned} xx^{-1} = 1xx^{-1} &= (x^{-1})^{-1}x^{-1}xx^{-1} = (x^{-1})^{-1}1x^{-1} \\ &= (x^{-1})^{-1}x^{-1} = 1. \end{aligned}$$

This is the main reason why the familiar notion of a group turned out to be less useful in linguistic applications than that of a pregroup. (For example, had we assigned to the attributive adjective *poor* the type $n_i n_i^{-1} = 1$ instead of $n_i n_i^{\ell}$ $(i = 1, 2, 3)$, its insertion anywhere in a sentence would be justified.)

A partially ordered monoid in which each element has an inverse is commonly called a *partially ordered group* (even though inversion of elements is contravariant). Evidently, a group is a left pregroup in which the inverse serves as the left adjoint.

Even readers with some training in mathematics will be hard put to provide an example of a left pregroup which is *not* a partially ordered group. Here is my favorite example: take the ordered set of natural numbers

$$\mathbb{N} = \{0, 1, 2, \cdots\}$$

and consider the set of all order preserving unbounded mappings f from $\mathbb{N}$ to $\mathbb{N}$. By this we mean:

if $m \leq n$ then $f(m) \leq f(n)$

and

if $n \to \infty$ then $f(n) \to \infty$.

This set is a monoid if we define the binary operation and the unity element thus:

$$\begin{aligned} (gf)(n) &= \ g(f(n)), \\ 1(n) &= \ n. \end{aligned}$$

The set is partially ordered if we define $f \to g$ to mean $f(n) \leq g(n)$ for all $n \in \mathbb{N}$.

It is a left pregroup, because every element f has a left adjoint f^{ℓ} defined thus:

$f^{\ell}(m) =$ minimum of all n such that $m \leq f(n)$.

f^{ℓ} is the left adjoint of f because

$$(f^{\ell}f)(k) = f^{\ell}(f(k)) = \min\{n \in \mathbb{N} | f(k) \leq f(n)\} \leq k = 1(k),$$

hence $f^{\ell}f \to 1$, and

$$1(k) = k \leq f(\min\{n \in \mathbb{N} | k \leq f(n)\} = f(f^{\ell}(k) = (ff^{\ell})(k),$$

hence $1 \to ff^{\ell}$.

A *right pregroup* is a partially ordered monoid in which each element x has a *right adjoint* x^r such that

$$xx^r \to 1 \to x^r x.$$

A *pregroup* is both a left and a right pregroup. It is not difficult to show that, in a pregroup,

$$x^{r\ell} = x = x^{\ell r}. \tag{4}$$

Proof. Since $xx^r \to 1 \to x^r x$, it follows that x is *a* left adjoint of x^r; but left adjoints are unique, hence x is *the* left adjoint of x^r, otherwise denoted by $x^{r\ell}$.

Evidently a partially ordered group is a pregroup. A pregroup other than a partially ordered group is the monoid of all order preserving mappings from $\mathbb{Z}$ to $\mathbb{Z}$, where $\mathbb{Z}'$ is the set of integers

$$\mathbb{Z} = \{......,\ -2,\ -1,\ 0,\ 1,\ 2,\ ....\}.$$

The system of types set up in Chapter 4 describes a very special pregroup, namely the pregroup $\mathcal{F}(\mathcal{B})$ *freely generated* by the partially ordered set of basic types. For the trained mathematician, this has the following *universal* property: any order preserving mapping from $\mathcal{B}$ into a pregroup $\mathcal{P}$ can be extended uniquely into an order preserving homomorphism from $\mathcal{F}(\mathcal{B})$ to $\mathcal{P}$ (which preserves not only the given binary operation, but also the two unary operations of left and right adjunction).

Since calculations in a pregroup may involve both contractions and expansions, one may wonder why, in our linguistic application, contractions alone were sufficient. The reason for this is the following.

Switching Lemma. When proving that $x \to y$ in the free pregroup $\mathcal{F}(\mathcal{B})$, one may assume, without loss of generality, that all (generalized) contractions precede all (generalized) expansions. Hence, if y is a simple type, no expansions are needed.

Proof (sketched). To show that $x \to y$ one must proceed step by step:

$$x = x_0 \to x_1 \to x_2 \to ...... \to x_{n-1} \to x_n = y,$$

where each step $x_i \to x_{i+1}$ has one of the following forms:

$$\begin{array}{ll} ua^{(z)}v \to ub^{(z)}v & \text{(simple step)}, \\ ua^{(z)}b^{(z+1)}v \to uv & \text{(generalized contraction)}, \\ uv \to ub^{(z+1)}a^{(z)}v & \text{(generalized expansion)}, \end{array}$$

where z is even and $a \to b$ in $\mathcal{B}$ or z is odd and $b \to a$ in $\mathcal{B}$.

We will look at a couple of cases only. For a complete proof, the reader may wish to consult [Lambek 1999]. Assume $a \to b \to c$ and suppose we have calculated

$$uac^r v \ \to\ ubc^r v \ \to\ uv$$

with a simple step followed by a generalized contraction, then we may replace these two steps by a single generalized contraction

$$uac^r v \ \to\ uv.$$

Next, suppose we have calculated

$$uc^r v \ \to\ ua^r bc^r v \ \to\ ua^r v$$

with a generalized contraction followed by an adjacent generalized expansion, then we may replace these two steps by the single step

$$uc^r v \;\to\; ua^r v.$$

On the other hand, if a generalized expansion is followed by a non-adjacent generalized contraction, then the order of the two steps may be reversed. For example, if $a \to b$ and $c \to d$, then the calculation

$$uab^r vw \;\to\; uab^r vcd^r w \;\to\; uvcd^r w$$

may as well be replaced by

$$uab^r vw \;\to\; uvw \;\to\; uvcd^r w.$$

The reader will have noticed that the mathematical system advocated here is associative and non-commutative. Occasional departure from associativity in the linguistic application is indicated by retaining some (left) brackets. Licensing of apparent commutativity is achieved by metarules which permute the simple components of the type of a word in the dictionary. Both devices could have been achieved by Moortgat's modalities.

28. Limitation of free pregroup grammars.

Grammars based on free pregroups have been applied to fragments of a number of languages; although, so far, iterated adjoints have only shown up in modern European languages, but not, for example, in Arabic or Latin. We may ask: can such pregroup grammars adequately describe natural languages?

Consider, for example, so-called "parasitic gaps" in English, as in

$$\begin{array}{l} \textit{whom did she tell } - \textit{ that she loves } - \, , \\ (\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\,\mathbf{q}^{\ell})(\mathbf{q}_2\,\mathbf{i}^{\ell}\,\pi^{\ell})\pi_3(\mathbf{i}[\hat{\mathbf{o}}^{\ell}\,\sigma^{\ell})(\overline{\mathbf{s}}\,\mathbf{s}^{\ell})\,\pi_3(\pi_3^r\,\mathbf{s}_1\,\mathbf{o}^{\ell}) \end{array} \rightarrow \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}[\hat{\mathbf{o}}^{\ell}\mathbf{o} \overset{?}{\rightarrow} \overline{\mathbf{q}}$$

which seems to require the rule

$$\hat{\mathbf{o}}^{\ell\ell}[\hat{\mathbf{o}}^{\ell}\mathbf{o}^{\ell} \rightarrow 1.$$

This could be derived from a putative postulate

$$\hat{\mathbf{o}} \rightarrow \mathbf{o}\hat{\mathbf{o}},$$

since then

$$\hat{\mathbf{o}}^{\ell}\mathbf{o}^{\ell} = (\mathbf{o}\hat{\mathbf{o}})^{\ell} \rightarrow \hat{\mathbf{o}}^{\ell}.$$

However, the only postulates we have admitted into our pregroup grammars are of the form $x \rightarrow y$, where x and y are simple types.

It is not even clear that English can be described by a *context-free* grammar, one in which all grammatical rules have the form

$$A_1 \ldots A_n \;\rightarrow\; A_{n+1},$$

where the A_i are elements of a vocabulary which includes a number of auxiliary grammatical terms in addition to the words of the object language. In fact, Buszkowski [2001] has shown that pregroup grammars are context-free.

While the jury is still out about English, it is known that context-free grammars cannot describe certain languages, e.g. Swiss German and Dutch. (See the discussion in Savitch et al. [1987].) The following comparison of subordinate clauses (with complementizer) in English, German and Dutch will throw some light on this phenomenon. I have taken the advice of Mark Twain [1910] and replaced German and Dutch words by English ones. In order to clearly distinguish the types of the three objects *her, John* and *Jane*, I have labeled them $\mathbf{o}_1, \mathbf{o}_2, \mathbf{o}_3$ respectively.

English : *that he will let her see John kiss Jane*;

$$(\overline{\mathbf{s}}\,\mathbf{s}^{\ell})\,\pi_3(\pi^r\,\mathbf{s}_1\,\mathbf{j}^{\ell})(\mathbf{i}\,\mathbf{i}^{\ell}\,\mathbf{o}_1^{\ell})\mathbf{o}_1(\mathbf{i}^{\ell}\,\mathbf{o}_2^{\ell})\mathbf{o}_2(\mathbf{i}\,\mathbf{o}_3)\mathbf{o}_3 \;\rightarrow \overline{\mathbf{s}}$$

(The discussion of Chapter 16 has been simplified, replacing $\mathbf{i}'^{\ell}$ by $\mathbf{i}^{\ell}$.)

German : *that he her Peter Mary kiss see let will* :

$$(\overline{\mathbf{s}}\,\mathbf{s}'^{\ell})\,\pi_3\,\mathbf{o}_1\,\mathbf{o}_2\,\mathbf{o}_3(\mathbf{o}_3^r\,\mathbf{i})(\mathbf{i}^r\,\mathbf{o}_2^r\,\mathbf{i})(\mathbf{i}^r\,\mathbf{o}_1^r\,\mathbf{i})(\mathbf{i}^r\,\pi_3^r\,\mathbf{s}') \;\rightarrow\; \overline{\mathbf{s}}$$

($\mathbf{s}'$ = indirect sentence without complementizer)

Dutch : *that he her Peter Mary will let see kiss.*

$$(\overline{\mathbf{s}}\mathbf{s}'^{\ell})\pi_3\mathbf{o}_1\mathbf{o}_2\mathbf{o}_3\;(\mathbf{s}'\pi_3^r\mathbf{i}^{\ell})(\mathbf{i}\mathbf{o}_1^r\mathbf{i}^{\ell})(\mathbf{i}\mathbf{o}_2^r\mathbf{i}^{\ell})(\mathbf{i}\mathbf{o}_3^r) \;\rightarrow\; \overline{\mathbf{s}}$$

We have omitted the underlinks which would illustrate the contractions $\mathbf{o}_1\mathbf{o}_1^r \to 1$, $\mathbf{o}_2\mathbf{o}_2^r \to 1$ and $\mathbf{o}_3\mathbf{o}_3^r \to 1$. Such "crossed dependencies" are not permitted in a free pregroup grammar. What is required here is a postulate allowing

$$\pi_3\mathbf{o}_1\mathbf{o}_2\mathbf{o}_3\mathbf{s}^\ell \to \mathbf{s}^\ell\mathbf{o}_3\mathbf{o}_2\mathbf{o}_1\pi_3,$$

which is not even allowed in a context-free grammar. Note, however, that, before contraction, this calculation involves nine simple types, which we have identified with Miller's chunks of information, thus reaching his proposed upper limit for people's capacity to process information.

It has been argued that the Dutch example will admit a context-free analysis after all if we ignore the subscripts 1, 2 and 3 (see Savitch et al.). However, this excuse won't work for Swiss German, which distinguishes direct and indirect objects by case endings.

It is known from formal language theory that the intersection of two context-free languages need not be context-free. The typical example is the language whose set of sentences is given by

$$\mathcal{L} = \{a^n b^n c^n | n \geq 1\}.$$

Clearly, this is the intersection of two context-free languages

$$\mathcal{L}_1 = \{a^m b^n c^n | m, n \geq 1\}, \qquad \mathcal{L}_2 = \{a^m b^m c^n | m, n \geq 1\}.$$

It is easy to obtain a pregroup grammar for $\mathcal{L}_1$ with sentences of type $\mathbf{s}_1$ and for $\mathcal{L}_2$ with sentences of type $\mathbf{s}_2$. Hence we can obtain a grammar for $\mathcal{L} = \mathcal{L}_1 \cap \mathcal{L}_2$ with sentences of type $\mathbf{s} = \mathbf{s}_1 \wedge \mathbf{s}_2$. This suggests defining a *lattice pregroup* as a pregroup with an additional operation $\wedge$ such that

$$c \to a \wedge b \text{ if and only if } c \to a \text{ and } c \to b.$$

One can then define another operation $\vee$ such that

$$a \vee b \to c \text{ if and only if } a \to c \text{ and } b \to c,$$

by

$$a \vee b = (a^\ell \wedge b^\ell)^r = (a^r \wedge b^r)^\ell,$$

the last equation being provable.

At the time of writing this, I am not aware of a cut-elimination theorem for lattice pregroups, comparable to that Buszkowski showed to be implied by the Switching Lemma of Chapter 27. However, a decision procedure is readily available: when looking at the intersection of languages $\mathcal{L}_1$ and $\mathcal{L}_2$, we may verify $x \to \mathbf{s}$ by calculating $x \to \mathbf{s}_1$ and $x \to \mathbf{s}_2$ in parallel. At this point in time, I don't see how this will help with Dutch or Swiss German.

Since writing the above, I had the privilege of conducting a course-seminar jointly with my colleague Brendan Gillon from the Linguistics Department. He suggested that working in the product of two or more free pregroups, carrying out several computations in parallel, might help not only to deal with the intersection of context-free languages, but also to simplify type assignments in other respects. Four of the students (Meaghan Fowley, Telyn Kusalik, Walter Pedersen and Jozina Vander Klok) followed his proposal by investigating how features could be encoded separately from our basic types. While most of them showed how Gillon's proposal would help with feature agreement in French, Kusalik pointed out that it might even be useful for English. A somewhat related suggestion had been made by Ed Stabler [2004] to facilitate comparison with Chomsky's program.

Let me illustrate the new approach with one example. Assigning to **m**ass nouns, **c**ount nouns and **p**lurals the new types

$$\begin{pmatrix} \mathbf{n} \\ \mathbf{m} \end{pmatrix}, \begin{pmatrix} \mathbf{n} \\ \mathbf{c} \end{pmatrix}, \begin{pmatrix} \mathbf{n} \\ \mathbf{p} \end{pmatrix}$$

in place of the original types $\mathbf{n}_0, \mathbf{n}_1$ and $\mathbf{n}_2$ respectively, we require only the single type $\begin{pmatrix} \mathbf{n}\mathbf{n}^\ell \\ 1 \end{pmatrix}$ for attributive adjectives. Assigning the type $\begin{pmatrix} \overline{\mathbf{n}}\,\mathbf{n}^\ell \\ \mathbf{p}\,\mathbf{p}^\ell \end{pmatrix}$ to the determiner *many* ($\overline{\mathbf{n}}$ standing for 'noun phrase'), we compute the type of

$$\begin{matrix} \textit{many} & \textit{old} & \textit{books} & & \\ \begin{pmatrix} \overline{\mathbf{n}}\,\mathbf{n}^\ell \\ \mathbf{p}\,\mathbf{p}^\ell \end{pmatrix} & \begin{pmatrix} \mathbf{n}\mathbf{n}^\ell \\ 1 \end{pmatrix} & \begin{pmatrix} \mathbf{n} \\ \mathbf{p} \end{pmatrix} & \to & \begin{pmatrix} \overline{\mathbf{n}} \\ \mathbf{p} \end{pmatrix} \end{matrix}$$

to be that of a plural noun phrase, the calculations resulting in $\overline{\mathbf{n}}$ and $\mathbf{p}$ being carried out in parallel.

The above tentative type assignment raises a new problem: how to account for the fact that *old books* can also have type $\begin{pmatrix} \overline{\mathbf{n}} \\ \mathbf{p} \end{pmatrix}$? The rule $\mathbf{n} \to \overline{\mathbf{n}}$ would wrongly predict that *old book* is also a complete noun phrase. We might wish to postulate instead that

$$\begin{pmatrix} \mathbf{n} \\ \mathbf{p} \end{pmatrix} \to \begin{pmatrix} \overline{\mathbf{n}} \\ \mathbf{p} \end{pmatrix},$$

but then we would no longer be operating in the product of two *free* pregroups, for which a decision procedure is guaranteed.

Kusalik's solution is more sophisticated: while retaining our basic type $\mathbf{n}_1$ (as his $\mathbf{n}_s$), he combines our types $\mathbf{n}_0$ and $\mathbf{n}_2$ into one type $\mathbf{n}_m$ and postulates

$$\mathbf{n}_s, \mathbf{n}_m \to \mathbf{n},\ \mathbf{n}_m \to \overline{\mathbf{n}}.$$

This allows him to analyze *old books* as a complete noun phrase, but the single type $\begin{pmatrix} \mathbf{n}\mathbf{n}^\ell \\ 1 \end{pmatrix}$ would then no longer serve for adjectives modifying both singular and plural nouns. I am sure that this problem can be overcome as well, but it shows that there are pitfalls to be avoided.

After writing the above, I came across two papers by Francez and Kaminski, who consider augmenting free pregroup grammars by a finite set of additional "inequations" (= arrows) between types, leading to a class of "mildly context-sensitive languages", which admit reduplication, crossed dependencies and multiple agreement, yet still possess a decision procedure. At first sight, it would seem that these additional rules should also allow us to handle parasitic gaps and may even absorb some of our metarules. Alternatively, one might view the new rules as metarules no longer restricted to the dictionary.

29. Conclusion.

We have investigated a representative fragment of English grammar, at a rough guess perhaps five percent, aiming to see how well it can be elucidated with the help of free pregroups. Many aspects of English grammar have been ignored, and so have many modern linguistic discoveries.

A free pregroup grammar assigns to each English word a type, namely an element of a certain (freely generated) algebraic or logical system. As an algebraic system, we are dealing with a generalization of the ubiquitous concept of a group, namely a partially ordered monoid in which each element has both a "left" and a "right adjoint", these terms having been borrowed from category theory.

A pregroup reduces to a group if the partial order is just equality. More generally, it reduces to a partially ordered group if the left and right adjoints coincide. As a logical system, we are dealing with a non-commutative version of Girard's linear logic (a classical substructural "logic lacking Gentzen's three structural rules of contraction, weakening and interchange), in which the tensor product and its DeMorgan dual coincide.

Miraculously, the types occupy approximately the same space as the words under which they are written. More significantly, a calculation on types may proceed in real time as a speaker produces an output or as a hearer absorbs the input. Miller's discovery that we can store at most seven (plus or minus two) chunks of information in our short term memory is consistent with the assumption that each simple type counts as one chunk of information.

The types assigned to words are assumed to be stored permanently in our "mental" dictionary. To prevent overloading the dictionary, we have employed two devices. We have introduced "inflectors", operations which transform a dictionary entry into the inflected form occurring in actual discourse. (These are absent in certain "analytic" or "isolating" languages, but predominate in others, such as Latin, Arabic or Turkish.) As our second device we have invoked certain "metarules", which assert that, if the dictionary assigns a certain type to a word, then this word may also have certain other types.

Various so-called "constraints of transformations" can be expressed with the help of certain left brackets pertaining to the "deep structure" of compound types, such as occur in $[a^\ell \cdots a]$ or $[a \cdots a^r]$ (and have been replaced by underlinks in most of our exposition). We discard all the right brackets and retain only some of the left ones. This will allow us to impose the "block constraint", which says that a type calculation is to be aborted as soon as one reaches two consecutive simple types $[a[b$ preceded by left brackets that have been retained.

The left brackets to be retained are:

(a) those that preclude premature contraction as in $x[y^r$ or $y^\ell[x$ when $x \to y$;

(b) those in front of double adjoints, such as $[a^{\ell\ell}$;

(c) those in the type of a relative pronoun, such as $\mathbf{n}_i^r \mathbf{n}_i [\mathbf{s}^\ell$;

(d) those in the type $x^r y [x^\ell$ of a coordinate conjunction;

(e) those in the type $\mathbf{i}^r \mathbf{i} [\mathbf{s}^\ell$ of a subordinate conjunction;

(f) that in the type $\mathbf{t}[\mathbf{s}^\ell$ of *whether* (if my critics are right).

Undoubtedly, I have overlooked other cases where left brackets ought to be retained as well.

In my opinion, such constraints should be viewed not as grammatical rules, but as expressing processing difficulties. At this time, it is not clear to me how and why processing difficulties arise, especially in comparison with Miller's observed limitation of the short term memory.

Problems with anaphora have been handled at a more elementary level than that of type assignment by pregroups, relying only on a rudimentary constituent analysis of the sentence. When can a pronoun represent a noun phrase in the same sentence? We have arrived at the following necessary conditions which together should also be sufficient:

(I) the pronoun agrees with the noun phrase in person, number and gender;

(II) the pronoun is properly contained in a key constituent (1) to (4), from which the noun phrase is excluded;

(III) the noun phrase is properly contained in a key constituent (0) to (3) and the pronoun occurs after this constituent.

By a *key constituent* is meant one of the following:

(0) a sentence (of type $\mathbf{s}_1 ...$);

(1) a noun phrase (of type $\overline{\mathbf{n}}, ...$);

(2) a subordinate clause (of type $\mathbf{i}^r \mathbf{i}\mathbf{s}^\ell, ...$);

(3) a quasi-sentence (of type $\mathbf{o}^r \overline{\mathbf{j}}, ...$);

(4) a spatio-temporal prepositional phrase (of type $\mathbf{i}^r \mathbf{i}, ...$).

To make these criteria work, we had to insert certain invisible copies of occurring noun phrases, which resemble Chomsky's "empty categories". More work should be done to express these criteria in pregroup terminology.

In our discussion of quantifiers, we have indicated how pregroup typing may be called upon to represent a word as a function, in the spirit of Curry-Montague semantics. Unfortunately such representations are not single-valued, unless the original square brackets are taken into account. This is in contrast to lexical semantics, which we have ascribed to the apocryphal Capulet. We have illustrated this by a brief introduction to kinship grammar, based on the theory of binary relations. This is an example of a semantic field that can be approached mathematically, even if not in terms of pregroup grammars.

Buszkowski [2003] had proved that our Switching Lemma (so named by him) shows that pregroup grammars are context-free. While the jury is still out about English, certain languages, e.g. Dutch, are known to be non-context-free. A fortiori, they cannot be completely described by a free pregroup grammar. The known counter-examples resemble those in a formal language, when a sentence may have type $\mathbf{s} = \mathbf{s}_1 \wedge \mathbf{s}_2$, $\mathbf{s}_1$ and $\mathbf{s}_2$ being sentence types of two other languages. It may be helpful to calculate $x \to \mathbf{s}$ by calculating $x \to \mathbf{s}_1$ and $x \to \mathbf{s}_2$ in parallel, thus working in the product of two free pregroups.

Let me conclude by suggesting that free pregroup grammars help to describe the syntax of natural languages to a good first approximation, but that some additional devices are required to achieve a closer fit. Here I have attempted to carry out such a program for one language only, not my mother tongue, and some of my grammatical judgements may be disputed by native speakers.

VIII. Summary.

30. Mathematical Rules.

(1)	$(xy)z = x(yz)$	(associativity)
(2)	$x1 = x = 1x$	(identity)
(3)	$x \to x$	(reflexivity)
(4)	if $x \to y$ and $y \to z$ then $x \to z$	(transitivity)
(5)	if $x \to y$ and $y \to x$ then $x = y$	(antisymmetry)
(6)	if $x \to y$ and $x' \to y'$ then $xx' \to yy'$	(compatibility)
(7)	if $x \to y$ then $uxv \to uyv$	(substitutivity)
(8)	$x^\ell x \to 1$ and $xx^r \to 1$	(contractions)
(9)	$1 \to xx^\ell$ and $1 \to x^r x$	(expansions)
(10)	if $x \to y$ then $y^\ell \to x^\ell$ and $y^r \to x^r$	(contravariance)
(11)	$(xy)^\ell = y^\ell x^\ell$ and $(xy)^r = y^r x^r$	(adjoints of products)
(12)	$1^\ell = 1 = 1^r$	(adjoints of identity)
(13)	$x^{\ell r} = x = x^{r\ell}$	(adjoints of adjoints)

(1) holds in any *semigroup*;
(2) makes the semigroup into a *monoid*;
(3) to (5) define a partially ordered set, a.k.a. *poset*;
(6) and (7) are equivalent ways of imposing a partial order on a monoid;
(1) to (7) define a *partially ordered monoid*;
(8) and (9) introduce *left* and *right adjoints*;
(1) to (9) define a *pregroup*;
(10) to (13) are easily derived consequences of the above, as is the following:

(14) left and right adjoints are unique.

See Chapter 27 for proofs.

31. The poset of basic types.

The pregroups that play a rôle in the study of natural languages are *freely generated* by the partially ordered sets of basic types. These differ from one language to another, although those for the languages studied so far bear a remarkable similarity to each other. For English, we find it convenient to introduce the following basic types, which represent traditional grammatical terms.

π = subject,

π_1 = first person singular subject,

π_2 = second person singular and any plural personal subject,

π_3 = third person singular subject,

$\hat{\pi}_3$ = pseudo-subject,

$\mathbf{s}$ = declarative sentence, a.k.a. statement,

$\mathbf{s}_1$ = statement in present tense,

$\mathbf{s}_2$ = statement in past tense,

$\overline{\mathbf{q}}$ = question,

$\mathbf{q}$ = yes-or-no question,

$\mathbf{q}_1$ = yes-or-no question in present tense,

$\mathbf{q}_2$ = yes-or-no question in past tense,

$\overline{\mathbf{s}}$ = indirect statement,

$\mathbf{t}$ = indirect question,

σ = direct or indirect sentence,

$\mathbf{i}$ = infinitive of intransitive verb,

$\mathbf{j}$ = infinitive of complete verb phrase,

$\overline{\mathbf{j}}$ = complete infinitive with *to*,

φ = quasi-sentence with infinitive,

ψ = quasi-sentence with participle,

$\mathbf{o}$ = direct object,

$\mathbf{o}'$ = indirect object,

$\hat{\mathbf{o}}$ = pseudo-object,

$\mathbf{n}$ = name,

$\mathbf{n}_0$ = mass noun,

$\mathbf{n}_1$ = count noun,

$\mathbf{n}_2$ = plural noun,

$\overline{\mathbf{n}}$ = complete noun phrase,

$\overline{\mathbf{n}}_i$ = complete noun phrase with $i = 0, 1, 2$,

$\mathbf{a}$ = predicative adjective,

$\overline{\mathbf{a}}$ = predicative adjectival phrase,

$\mathbf{p}_1$ = present participle,

$\mathbf{p}_2$ = past participle.

To impose a partial order on the set of basic types, we introduce the following postulates, which allow other order relations to be derived by transitivity.

$\pi_k \to \pi \qquad (k = 1, 2, 3);$

$\hat{\pi}_3 \to \pi_3;$

$\mathbf{s}_j \to \mathbf{s} \qquad (j = 1, 2);$

$\mathbf{q}_j \to \mathbf{q} \to \overline{\mathbf{q}};$

$\mathbf{s} \to \sigma,\ \overline{\mathbf{s}} \to \sigma,\ \mathbf{t} \to \sigma,\ \overline{\mathbf{q}} \to \sigma;$

$\overline{\mathbf{s}} \to \pi_3;$

$\mathbf{i} \to \mathbf{i}' \to \mathbf{j}' \to \mathbf{j};$

$\varphi, \psi, \overline{\mathbf{j}}, \mathbf{p}_1 \to \pi_3;$

$\mathbf{o}' \to \mathbf{o};$

$\hat{\mathbf{o}} \to \mathbf{o};$

$\mathbf{n}_0 \to \overline{\mathbf{n}}_0 \to \pi_3, \mathbf{o};$

$\mathbf{n}_2 \to \overline{\mathbf{n}}_2 \to \pi_2, \mathbf{o};$

$\overline{\mathbf{n}}_1 \to \pi_3, \mathbf{o};$

$\overline{\mathbf{n}}_i \to \overline{\mathbf{n}} \to \mathbf{o} \qquad (i = 0, 1, 2);$

$\mathbf{n} \to \overline{\mathbf{n}};$

$\mathbf{a} \to \overline{\mathbf{a}}.$

32. Type assignments.

Sentences built with compound tenses are assigned the type $\mathbf{s}_j$ where j denotes the tense of the auxiliary verb; thus $j = 1$ in (a) and $j = 2$ in (b):

(a) *she will come, she has come, she is coming*;

(b) *she would come, she had come, she was coming.*

We have found it convenient to regard *would* as the past of *will*, even if this is not justified semantically.

Words in the (mental) dictionary are assigned compound types, that is, elements of the pregroup freely generated from the poset of basic types.

Finite forms of modal verbs

$$\textit{will, can, ..., won't, can't, ...}: \pi^r\mathbf{s}_1\mathbf{j}^\ell, \mathbf{q}_1\mathbf{j}^\ell\pi^r,$$

the former type in statements, the latter in questions.

Finite forms of the emphatic auxiliary:

do, don't: $\pi_k^r\mathbf{s}_1\mathbf{i}^\ell$, $\mathbf{q}_1\mathbf{i}^\ell\pi_k^\ell$ $(k = 1, 2)$;

does, doesn't: $\pi_3^r\mathbf{s}_1\mathbf{i}^\ell$, $\mathbf{q}_3\mathbf{i}^\ell\pi_3^\ell$;

did, didn't: $\pi^r\mathbf{s}_2\mathbf{i}^\ell$, $\mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell$.

Infinitive of the auxiliary verb *be*:

$\mathbf{j}'\mathbf{p}_1^\ell$ for expressing the progressive tense,

$\mathbf{i}'\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell$ for constructing the passive voice,

$\mathbf{i}'\overline{\mathbf{a}}^\ell$ for the copula.

However,

get: $\mathbf{i}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^\ell$ for an alternative construction of the passive voice.

Infinitive of the auxiliary verb *have*:

$\mathbf{j}\mathbf{p}_2^\ell$ for expressing the perfect tense.

The following main verbs are listed in the infinitive.

Intransitive verbs: *come, go*, ...: $\mathbf{i}$;

transitive verbs: *like, smash*, ...: $\mathbf{i}\mathbf{o}^\ell$;

verbs which may be transitive or intransitive:

see, drink, ...: $\mathbf{i}(\mathbf{o}^\ell)$;

verbs with a sentential complement:

believe, know, say, think, ...: $\mathbf{i}\overline{\mathbf{s}}^\ell$;
know, wonder, ask, ...: $\mathbf{i}\mathbf{t}^\ell$;

verbs with a quasi-sentential complement:

want: $\mathbf{i}\phi^\ell$, $\mathbf{i}\overline{\mathbf{j}}^\ell\mathbf{o}^\ell$, $\mathbf{i}\overline{\mathbf{j}}^\ell$;
hope: $\mathbf{i}\phi^\ell$, $\mathbf{i}\overline{\mathbf{j}}^\ell$;
recall: $\mathbf{i}\psi^\ell$, $\mathbf{i}\mathbf{p}_1^\ell\mathbf{o}^\ell$, $\mathbf{i}\mathbf{p}_1^\ell$;

verbs which are usually doubly transitive:

give, show, tell, ...: $\mathbf{i}\overline{\mathbf{n}}^{\ell}\mathbf{o}'^{\ell}$, $\mathbf{i}\mathbf{o}'^{\ell}$;

verbs with one direct object and a sentential or quasi-sentential complement:

tell, show, ...: $\mathbf{i}\sigma^{\ell}\mathbf{o}^{\ell}$, $\mathbf{i}\overline{\mathbf{j}}^{\ell}\mathbf{o}^{\ell}$;
persuade, ...: $\mathbf{i}\overline{\mathbf{s}}^{\ell}\mathbf{o}^{\ell}$, $\mathbf{i}\overline{\mathbf{j}}^{\ell}\mathbf{o}^{\ell}$;

verbs of causation and perception:

have, help, let, make; see, hear, ...: $\mathbf{i}\mathbf{i}'^{\ell}\mathbf{o}^{\ell}$;

verbs with detachable particles:

bring[-]*in, turn*[-]*off,...*: $(\mathbf{i}\delta^{\ell}\mathbf{o}^{\ell}) - \delta^{\ell}, (\mathbf{i}\overline{\mathbf{n}}^{\ell}\delta^{\ell})\delta^{\ell}$;

Adjectives:

good, old, ...: $\mathbf{a}, \mathbf{a}\mathbf{a}^{r}\mathbf{n}_i\mathbf{n}_i^{r} \rightarrow \mathbf{n}_i\mathbf{n}_i^{r}$ $(i = 0, 1, 2)$;
eager: $\mathbf{a}$, $\mathbf{a}\mathbf{a}^{r}\mathbf{n}_i\mathbf{n}_i^{r}$, $\mathbf{a}\overline{\mathbf{j}}^{\ell}$,
easy: $\mathbf{a}$, $\mathbf{a}\mathbf{a}^{r}\mathbf{n}_i\mathbf{n}_i^{\ell}$, $\mathbf{a}\hat{\mathbf{o}}^{\ell\ell}\overline{\mathbf{j}}^{\ell}$.

Participles usually just replace $\mathbf{i}$ by $\mathbf{p}_j$. But participles representing a mental or emotional state may be genuine adjectives:

amusing, interesting,...: $\mathbf{p}_1, \mathbf{a}, \mathbf{a}\mathbf{a}^{r}\mathbf{n}_i\mathbf{n}_i^{\ell} \rightarrow \mathbf{n}_i\mathbf{n}_i^{\ell}$;
amused, interested, ...: $\mathbf{p}_2, \mathbf{a}_2, \mathbf{a}\mathbf{a}^{r}\mathbf{n}_i\mathbf{n}_i^{\ell} \rightarrow \mathbf{n}_i\mathbf{n}_i^{\ell}$.

Adverbs may have one of the following types:

$$xx^{\ell}(x = \mathbf{s}, \mathbf{s}\pi^{r}, \mathbf{j}, \mathbf{p}_2, \mathbf{i}^{r}, \mathbf{s}^{r})$$

very: $\mathbf{a}\mathbf{a}^{\ell}$;
too: $\overline{\mathbf{a}}\overline{\mathbf{a}}^{\ell}$;
not: $xx^{\ell}(x = \mathbf{i}, \mathbf{j}, \mathbf{p}_1, \mathbf{p}_2, \overline{\mathbf{a}})$.

Prepositions (also called transitive adverbs):

$$x^{r}x\mathbf{o}^{\ell}\ (x = \mathbf{n}_i, \mathbf{s}, \mathbf{j}, \cdots);$$
$$\overline{\mathbf{a}}\mathbf{o}^{\ell},\ \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell\ell}\mathbf{q}^{\ell}, \mathbf{n}_i^{r}\mathbf{n}_i\hat{\mathbf{o}}^{\ell\ell\ell}\mathbf{n}_i^{\ell}\mathbf{n}_i \cdots\ .$$

Conjunctions:

coordinate: $x^{r}yx^{\ell}$, $\pi_i^{r}\pi_2\mathbf{n}_i^{\ell}$;
subordinate: $\mathbf{j}^{r}\mathbf{j}\mathbf{s}^{\ell}$.

Complementizers:

to: $\overline{\mathbf{j}}\mathbf{i}^{\ell}$;
that: $\overline{\mathbf{s}}\mathbf{s}^{\ell}$;
whether, if: $\mathbf{t}\mathbf{s}^{\ell}, \mathbf{t}\overline{\mathbf{j}}^{\ell}$;
for: $\varphi\overline{\mathbf{j}}^{\ell}\mathbf{o}^{\ell}$;
Gen: $\psi\mathbf{p}_1^{\ell}\mathbf{o}^{\ell}$.

Determiners modifying nouns:

modifying mass nouns, e.g. *much*: $\overline{\mathbf{n}}_0\mathbf{n}_0^{\ell}$;

modifying count nouns, e.g. *a(n)*: $\overline{\mathbf{n}}_1\mathbf{n}_1^\ell$;

modifying plurals, e.g. *many*: $\overline{\mathbf{n}}_2\mathbf{n}_2^\ell$;

modifying any nouns, e.g. *the, my*: $\overline{\mathbf{n}}_i\mathbf{n}_i^\ell$ $(i = 0, 1, 2)$.

Quantifiers may be treated provisionally as pronouns, but more accurately as follows:

somebody, something: π_3, $\mathbf{os}^r\mathbf{s}$, $\mathbf{oi}^r\mathbf{i}$, $\mathbf{ss}^\ell\pi_3$;
everybody, everything: π_3, $\mathbf{oi}^r\mathbf{i}$, $\mathbf{ss}^\ell\pi_3$, $\pi_3\mathbf{ii}^\ell$;
anybody, anything: π_3, $\mathbf{os}^r\mathbf{s}$, $\mathbf{oq}^r\mathbf{q}$, $\pi_3\mathbf{iq}^r\mathbf{qi}^\ell$.

Adjectival quantifiers may be treated provisionally as determiners, but more accurately as follows:

every, each: $\overline{\mathbf{n}}_1\mathbf{n}_1^\ell$, $\mathbf{ss}^\ell\pi_3\mathbf{n}_1^\ell$;
all: $\overline{\mathbf{n}}_2\mathbf{n}_2^\ell$, $\mathbf{ss}^\ell\pi_2\mathbf{n}_2^\ell$;
any, some: $\overline{\mathbf{n}}_i\mathbf{n}_i^\ell$, $\mathbf{ss}^\ell\pi_k\mathbf{n}_i^\ell$ where $(k, i) = (3, 0), (3, 1)$ or (2.2).

Direct question words:

adverbial, *where, when, why, how*: $\mathbf{s}^r\overline{\mathbf{q}}, \overline{\mathbf{q}}\mathbf{q}^\ell, \overline{\mathbf{q}}\mathbf{i}^\ell\mathbf{i}^{\ell\ell}\mathbf{q}^\ell$;

asking for an object, *whom, what*: $\mathbf{os}^r\overline{\mathbf{q}}, \overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell$;

asking for a subject, *who, what*: $\overline{\mathbf{q}}\mathbf{s}^\ell\pi_3, \overline{\mathbf{q}}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell$;

asking for an object modifier, *whose, which*: $\overline{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell\mathbf{n}_i^\ell$;

asking for a subject modifier, *whose, which*: $\overline{\mathbf{q}}\hat{\pi}_k^{\ell\ell}\mathbf{q}^\ell\mathbf{n}_i, \overline{\mathbf{q}}\mathbf{s}^\ell\pi_k\mathbf{n}_i^\ell$,
when $(k, i) = (3, 0), (3, 1)$ or $(2, 2)$.

Indirect question words, many of which may be derived from Metarule 4 (see Chapter 31):

asking for affirmation: *if, whether*: $\mathbf{ts}^\ell, \mathbf{t}\overline{\mathbf{j}}^\ell$;

asking for an object: *whom, what*: $\mathbf{t}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell, \mathbf{t}\hat{\mathbf{o}}^{\ell\ell}\overline{\mathbf{j}}^\ell$; and so on.

Restrictive relative pronouns describing an object:
whom, which, [*that*]: $\mathbf{n}_i^r\mathbf{n}_i\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$, $\mathbf{n}_i^r\mathbf{n}_i\hat{\mathbf{o}}^{\ell\ell}\overline{\mathbf{j}}^\ell$;

restrictive relative pronoun describing a subject:
who, which, that: $\mathbf{n}_i^r\mathbf{n}_i\mathbf{s}^\ell\pi_k$, where $(k, i) = (3, 0)(3, 1)$ or $(2, 2)$;

restrictive relative pronoun modifying another noun on the right:
whose: $\mathbf{n}_i^r\mathbf{n}_i\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$, $\mathbf{n}_{i'}^\ell, \mathbf{n}_i^r\mathbf{n}_i\mathbf{s}^\ell\pi_k\mathbf{n}_{i'}^\ell$;
where $(k, i') = (3, 0), (3, 1)$ or $(2, 2)$.

Non-restrictive relative clauses are introduced by the same relative pronouns, except [*that*].

33. Dictionary rewriting rules.

In order not to overload the dictionary we employ two devices which allow us to calculate the type of a word form: inflectors and metarules.

English is almost an analytic (a.k.a. isolating) language like Chinese. Yet it still has retained some remnants of the inflections, which predominate in most of its Indo-European relatives. We have found it convenient to introduce certain "inflectors", which are supposed to operate on dictionary entries to produce inflected forms occurring in actual discourse.

For example, the infinitive *go* is listed in the dictionary, together with its irregular past *went*, and the inflector C_{jk} converts these entries into the finite forms

$$\begin{array}{lcl} C_{jk}\ go & \to & \begin{pmatrix} go\ go\ goes \\ went\ went\ went \end{pmatrix} \\ (\pi_k^r \mathbf{s}_j \mathbf{j}^\ell)\mathbf{i} & \to & \pi_k^r \mathbf{s}_j \end{array}$$

where the upper arrow describes an ordinary rewrite rule, as in a generative grammar, and the lower arrow represents the partial order of the pregroup.

We summarize a few remaining inflectors.

C_{jk} : $\pi_k^r \mathbf{s}_j \mathbf{j}^\ell \quad (j = 1, 2;\ k = 1, 2, 3)$;
P_1 : $\mathbf{p}_1 \mathbf{i}'^\ell$;
P_2 : $\mathbf{p}_2 \mathbf{j}'^\ell$;
Acc : $\mathbf{o}\pi^\ell$;
Plur : $\mathbf{n}_2 \mathbf{n}_1^\ell$;
Gen : $\overline{\mathbf{n}}_i \mathbf{n}_i^\ell \mathbf{o}^\ell$.

Metarules take the form:

Something of type x also has type y. We will abbreviate this here by $x \Rightarrow y$, not to be confused with a postulate of the form $x \to y$, where x and y are restricted to be basic types. In particular, we do not wish to imply that $y^\ell \Rightarrow x^\ell$.

Metarule 1. $\pi_k^r \mathbf{s}_i x^\ell \Rightarrow \mathbf{q}_k x^\ell \pi_k^\ell$ $(x = \mathbf{i}, \mathbf{j}, \mathbf{p}_1, \mathbf{p}_2, \hat{\mathbf{o}}^\ell, \overline{\mathbf{a}})$.

Metarule 2. $\mathbf{o}^\ell \Rightarrow \hat{\mathbf{o}}^\ell \mathbf{i}^\ell \mathbf{i}$.

Metarule 3. $\pi_k^\ell \mathbf{s}_i \mathbf{j}^\ell \Rightarrow \mathbf{q}_j \pi_k^\ell \mathbf{j}^\ell$.

Metarule 4. $\overline{\mathbf{q}} \cdots \mathbf{q}^\ell \cdots \Rightarrow \mathbf{t} \cdots \mathbf{s}^\ell \cdots$.

Metarule 5. $\mathbf{i} x^\ell \mathbf{o}^\ell \Rightarrow \mathbf{i}\hat{\mathbf{o}}^\ell x^\ell$ $(x = \overline{\mathbf{s}}, \mathbf{t}, \bar{\mathbf{j}}, \overline{\mathbf{n}})$.

Metarule 6. $\mathbf{p}_2 \overline{\mathbf{n}}^\ell x^\ell \Rightarrow \overline{\mathbf{a}} x^\ell$ $(x = 1, \mathbf{o}, \cdots)$.

Metarule 7. $\mathbf{i} x^\ell \mathbf{o}^\ell \Rightarrow \mathbf{i}\overline{\mathbf{n}}^\ell x^\ell$ $(x = \mathbf{i}', \overline{\mathbf{a}}, \delta, \cdots)$.

These metarules are applied in the following contexts:

1. to the finite form of a modal or auxiliary verb, to convert a declarative sentence into an interrogative one, and similarly when the subscripts j or k are omitted.

2. in the type of a transitive verb or preposition;

3. in the type of the inflector C_{jk} to convert a pseudo-question into a direct question;

4. in the type of a question pronoun to convert a direct question into an indirect one;

5. in the type of a verb with two complements;

6. for past participles of verbs with two complements;

7. in the type of a verb with two complements.

The reader is invited to formulate other metarules which will help to simplify the dictionary even further and to combine some of those we have listed, e.g. **5** and **7**.

IX. Appendices.

34. Related mathematical recreations.

While relevant elementary mathematics has been introduced gingerly in chapters 4, 27 and 30, some peripheral recreational mathematics, involving more tricky arguments, has been banished to this appendix, lest readers with a phobia for mathematics be frightened off.

Let us return to the problem raised in chapter 13: in how many ways can a string of $2m+1$ occurrences of the word *police* be interpreted as a declarative sentence? (I must confess that originally I thought the answer was $m!$, but Telyn Kusalik persuaded me that I had counted some of the interpretations more than once and that the correct answer was the Catalan number $f(m+1)$.)

Such an interpretation would employ the plural noun *police* $m+1$ times and the verb *police*, meaning *control*, m times, with $m \geq 1$. Moving the main verb to the end of the string, we obtain the equivalent problem: in how many ways can such a string be interpreted as a plural noun phrase? In this formulation, we may even allow $m \geq 0$, since the noun *police* is a noun phrase. Replacing the verb *police* by a right parenthesis, replacing the noun *police* by the letter p and inserting left parentheses judiciously, we transform the problem into a familiar one: in how many ways can a string of $m+1$ letters be made into a bracketed string?

Given a set $\sum$ of one or more letters, we define a *bracketed string* in $\sum$ as follows:

(1) all elements of $\sum$ are bracketed strings,
(2) if A and B are bracketed strings then so is (AB),
(3) nothing else is a bracketed string in $\sum$.

We may as well take $\sum = \{p\}$, and we denote the number of bracketed strings of length n by $f(x)$.

n	bracketed strings	$f(n)$
1	p	1
2	(pp)	1
3	$((pp)p),\ (p(pp))$	2
4	$(((pp)p)p),\ ((pp)(pp)),\ (p(p(pp))),$ $(p((pp)p)),\ ((p(pp)p))$	5

We note that $f(5)$ is easily calculated:

$$\begin{aligned} f(5) &= f(1)f(4) + f(2)f(3) + f(3)f(2) + f(4)f(1) \\ &= 5 + 2 + 2 + 5 = 14. \end{aligned}$$

In general, we have the recurrence formula

$$(*) \qquad \begin{aligned} f(n) &= f(1)f(n-1) + f(2)f(n-2) + \cdots + f(n-1)f(1) \\ &= \sum_{k=1}^{n-1} f(k)f(n-k). \end{aligned}$$

This formula does allow us to calculate $f(n)$ inductively, but we shall do better than that. Consider the formal infinite power series in the indeterminate x:

$$F(x) = \sum_{n=1}^{\infty} f(n)x^n.$$

Then

$$\begin{aligned}
F(x)^2 &= \sum_{k=1}^{\infty} f(k)x^k \sum_{\ell=1}^{\infty} f(\ell)x^\ell \\
&= \sum_{n=2}^{\infty} \left(\sum_{n=k+\ell} f(k)f(\ell) \right) x^n \\
&= \sum_{n=2}^{\infty} \left(\sum_{k=1}^{n-1} f(k)f(n-k) \right) x^n \\
&= \sum_{n=2}^{\infty} f(n)x^n \\
&= F(x) - x.
\end{aligned}$$

Thus

$$F(x)^2 - F(x) + x = 0,$$

and so

$$\begin{aligned}
F(x) &= \frac{1}{2}(1 \pm (1-4x)^{\frac{1}{2}}) \\
&= \frac{1}{2} \pm \frac{1}{2} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4x)^n,
\end{aligned}$$

by the generalized binomial theorem.

Since $f(1)$ is positive, we must choose the minus sign, hence

$$F(x) = -\frac{1}{2} \sum_{n=1}^{\infty} \binom{\frac{1}{2}}{n} (-4x)^n,$$

and therefore the coefficient $f(n)$ of x^n is given by

$$(**) \qquad f(n) = \frac{1}{2} \binom{\frac{1}{2}}{n} (-1)^{n-1} 4^n.$$

We do not wish to leave the answer in this form, but we shall use the fact that the binomial coefficient

$$\begin{aligned}
\binom{\frac{1}{2}}{n} &= \frac{1/2(-1/2)(-3/2)\cdots(1/2-n+1)}{1 \cdot 2 \cdot 3 \cdot \ldots \cdot n} \\
&= \frac{(-1)(-3)\ldots(-(2n-3))}{2^n n!}.
\end{aligned}$$

Multiplying top and bottom by $2 \cdot 4 \cdot 6 \cdot \ldots \cdot 2n-2$, we obtain

$$\binom{\frac{1}{2}}{n} = \frac{(-1)^{n-1}(2n-2)!}{2^n n! 2^{n-1}(n-1)!}.$$

Substituting this into formula $(**)$, we get

$$\begin{aligned}
f(n) &= \frac{1}{2} \frac{(-1)^{n-1}(2n-2)!}{2^n n! 2^{n-1}(n-1)!} (-1)^{n-1} 4^n \\
&= \frac{(2n-2)!}{n!(n-1)!} \\
&= \frac{1}{n} \binom{2n-2}{n-1}.
\end{aligned}$$

Writing $n-1=m$, we end up with the neater formula

$$(***) \qquad f(m+1) = \frac{1}{m+1}\begin{pmatrix} 2m \\ m \end{pmatrix}, \text{ for all } m \geq 0.$$

For example, taking $m=4$, we calculate

$$f(5) = \frac{1}{5}\begin{pmatrix} 8 \\ 4 \end{pmatrix} = \frac{1}{5}\frac{8\cdot 7\cdot 6\cdot 5}{1\cdot 2\cdot 3\cdot 4} = 14.$$

The formula $(***)$ was discovered by Catalan in 1838.

Pregroups have previously occurred in recreational mathematics and appeared in exercises for elementary number theory, even if not under this name. Recall that a pregroup degenerates into a partially ordered group when $x^\ell = x^r$ for all elements x. An example of a non-degenerate pregroup is the partially ordered monoid of all unbounded order preserving mappings $\mathbb{Z} \to \mathbb{Z}$ [L1994]. Even earlier, a one-sided analogue appeared in a paper written jointly with Leo Moser [1954].

By a *left pregroup* we mean a partially ordered monoid in which each element has a left, but not necessarily a right, adjoint. A left pregroup degenerates into a partially ordered group if $x^{\ell\ell} = x$ for all elements x. To obtain a non-degenerate left pregroup, consider the partially ordered monoid of all unbounded order preserving mappings $f : \mathbb{N} \to \mathbb{N}$, where we define

$$\begin{aligned} &(gf)(x) = g(f(x)), \\ &f \to g \text{ iff } \forall_{x\in\mathbb{N}}(f(x) \leq g(x)), \\ &f^\ell(x) = \min\{y \in \mathbb{N} | x \leq f(y)\}, \end{aligned}$$

hence

$$f^\ell(n) \leq y \text{ iff } x \leq f(y).$$

If $f(x) = [x/2]$, then

$$f^\ell(x) = 2x, \;\; f^{\ell\ell}(x) = [(x+1)/2],$$

so this left pregroup is not a partially ordered group.

The early interest this left pregroup attracted (though not under this name) arose from a remarkable property:

Theorem 34.1. The sets

$$\{f^\ell(x) + x | x \in \mathbb{N}\}, \;\; \{f(y) + y + 1 | y \in \mathbb{N}\}$$

are complementary subsets of $\mathbb{N}$, that is, every natural number occurs in one and only one of these two sets.

Before proving this, let us look at an example. Let

$\pi(x) =$ the number of primes $\leq x$,

then

$\pi^\ell(x) = 0$ if $x = 0$, p_x if $x \geq 1$, where p_x is the x-th prime number. Now look at the following table:

x	$\pi^\ell(x)$	$\pi^\ell(x) + x$	$\pi(x)$	$\pi(x) + x + 1$
0	0	0	0	1
1	2	3	0	2
2	3	5	1	4
3	5	8	2	6
4	7	11	2	7
5	11	16	3	9
6	13	19	3	10
...	...	...	...	...

We observe that the third and fifth columns are complementary. (The missing numbers 12, 13, 14, 15, 17 and 18 would appear further down.)

Proof. Since $f^\ell(x) \leq y$ iff $x \leq f(y)$,

$$\begin{aligned} f^\ell(x) + x \leq x + y \text{ iff } & x + y \leq f(y) + y \\ \text{iff } & x + y < f(y) + y + 1, \end{aligned}$$

hence no number of the form $f(x) + x$ can be equal to a number of the form $f(y) + y + 1$.

Next, let $0 \leq x \leq n$ and $y = n - x$, n being a given natural number. Then

$$\begin{aligned} f(x) + x \leq n \text{ iff } & n < f(y) + y + 1 \\ \text{iff } & \text{not } f(y) + y + 1 \leq n. \end{aligned}$$

Therefore, exactly one of the numbers $f^\ell(x) + x$ and $f(y) + y + 1$ lies in the interval $[0, n]$. Since n is arbitrary, it follows that the two sets in Theorem 34.1 are complementary.

35. Selected etymologies.

algebra derived from the Arabic title of a mathematical treatise. *Algebra* means "restoring", and the word *algebrista* in old Spanish also meant "bonesetter". The full title of the book was something like "The science of restoring what is missing and equating like with like".

algorithm coined after the ninth century Persian mathematician Al-Khowārizmi, the author of the above text. He wrote in Arabic, just like Europeans writing in Latin. The letters "th" arose from a mistaken analogy to "logarithm".

calculate/calculus from the Latin word for "pebble". Primitive calculations were performed by moving pebbles on a table. Dentists still use the word "calculus" to describe the tartar deposited on teeth.

case from Latin *casus* meaning "fall", referring to the six ways a die can fall. Latin substantives had six grammatical cases.

category from Greek *katégorein*, originally meaning "to accuse", a compound word made up from the prefix *kata* meaning "against" and *agora* meaning "forum" or "public assembly". The word was used by Aristotle to mean "type", and is still used in that sense in philosophy and linguistics. In mathematics the word has acquired a completely different meaning, adopted by Eilenberg and Mac Lane in their pioneering article on "category theory".

compute/count from Latin *computare* and French *compter*. The related French *comte* means "tale". The analogy between *count* and *recount* seems to be common to many European languages, e.g. German *zählen* and *erzählen*.

element is said to be related to *elephant*, by at least one etymologist Klein [1971]), who claims that the Etruscans thought that the letters of the alphabet were as precious as ivory. On the other hand, it has been suggested that *eleph* is related to the first letter of the Phoenician alphabet, where it symbolizes an ox. Perhaps it is this rather than ivory which explains the connection (if any) between *elephant* and *element*?

glamour/grammar are derived from the Greek *graphein*, meaning "to write". In the Middle Ages, a knowledge of Latin grammar was supposed to endow a scholar with magical power, hence *glamour* means "the power to enchant".

language/linguist from the Latin word for *tongue*, actually cognate to the latter, the original initial dental consonant having undergone a phonological transformation in Latin.

logic≠logistics, the latter being derived from French *loger* (English *lodge*) and refers to the quartermaster's science, while the former comes from Greek *logos*, meaning "word". Philo of Alexandria interpreted Plato's "ideas" as "words", hence the New Testament statement "in the beginning was the Word". I was once invited to Paris by a French mathematician who had obtained a grant from the military to apply category theory to logistics, evidently exploiting the confusion between the two words.

+ly<like. The morpheme +*ly* which converts an adjective into an adverb is derived from an obsolete word meaning "body", still surviving in British English *lychgate* and German *Leiche*. The French find themselves on the opposite side of the mind-body divide and use the morpheme +*ment*, related to English *mind*, for the same purpose.

mathematics comes from a Greek word meaning "learning", which is also cognate to English *mind* and *mean*. Originally, mathematics was stipulated by the Pythagoreans to consist of four subjects: Arithmetic, Geometry, Astronomy and Music. Much later, Boethius, at the court of Theodoric in Ravenna, adopted the same four subjects as a prerequisite for the more elementary trivium, giving rise to the "quadrivium" of the medieval seven liberal arts.

person in grammar refers to the first, second or third person, singular or plural. The word *person* is derived from Latin *persona*, meaning an actor's mask or character in a play. German *Person* and French *personne* are supposed to be represented by a feminine pronoun, even if the person in question is male as in "this person shaved her beard" in a literal translation of a German sentence. In the sixties of the twentieth century, political correctness required "man" to be replaced by "person" when the sex was considered to be irrelevant. People still speak of "a chairperson". In some contexts, this substitution has resulted in "person" now referring to a woman, e.g. in "the person Friday". Curiously, French *personne* becomes masculine when it means "nobody".

syntax is derived from a Greek word meaning "putting together", like Arabic *algebra*. (It does not refer to a penalty for immoral behaviour.)

tale/talk/tell are all related to German *Zahl*, meaning "number". Like French *compte/conte*, reminding us of the popular view that linguistic communication (= recounting) involves computing or counting.

trivial originally refers to the *trivium*, the three subjects, logic, grammar and rhetoric, that were to serve as prerequisites to the quadrivium (see "mathematics") at medieval universities.

uncle nowadays means "male (spouse of) sibling of parent". It is derived from Latin *avunculus*, little grandfather, referring to the mother's brother, who became her legal guardian after her father's death. Reciprocally, the word *nephew* originally meant "grandson", and Italian *nepote* can still be translated as either "nephew" or "grandson".

verb from Latin *verbum* originally meant "word", and like *word* is derived from Indo-European **werdh*. (In the present text, *word* sometimes refers to a dictionary entry and sometimes to is inflected form; the reader will easily decide from the context which meaning is intended.)

36. Postscript on Minimalism.

Readers familiar with current linguistic theory will have noticed that many ideas underlying the present monograph have been influenced by the profound insights gained by Chomsky and his school in the last half century. However, there are technical differences in how these ideas are displayed. Mainstream (Chomskyan) grammarians assume that words are arranged in labeled bracketed strings, or equivalently as leaves on a plane tree with labeled nodes, and that sentences are generated and analyzed by manipulating these structures.

In categorial grammars, such calculations are performed not on labeled bracketed strings of words, but on strings of types, and this is assumed to be done in real time, even as the words are uttered or perceived during actual communication. In particular, this approach rules out any movement from the end of a sentence to the beginning.

Chomsky's ideas have undergone a number of significant changes over the years, which have usually led his many disciples to revise their methodology as well. (I am reminded of the French comic book "Asterix and Cleopatra", in which Obelix accidentally breaks off the nose of the Sphinx, whereupon the souvenir vendors hasten to hack off the noses of their replicas.)

For the last dozen years or so, the official version of the Chomskyan enterprise has been the so-called "minimalist program" (described as "maximalist" by some). It is not clear to me which of the earlier notions have been discarded and which have been absorbed in the new program. The two texts I have looked into rely heavily on a discussion of the earlier theories. For example, Chomsky [1995] discusses government theory, binding theory, case theory, X-bar theory, phrase structure theory and theta-theory, to judge by the section headings only, not counting trace theory and optimality theory, which are mentioned elsewhere (see Lasnik et al.[2005]). Is the reader required to study all these theories before engaging on the minimalist program?

I must confess that I haven't mastered all these theories, although I suspect that what some of them are about can be incorporated into the pregroup framework. For example, Lasnik and Uriagereka [2005], under the topic of case theory, discuss the assertion that English nouns and adjectives, as opposed to verbs and prepositions, never require a direct object. In the language of pregroup grammar, this fact is expressed by noting that nouns and adjectives never have types $\mathbf{n}_i\mathbf{o}^\ell$ $(i = 0, 1, 2)$ and $\mathbf{ao}^\ell$ respectively, whereas transitive verbs have type $\mathbf{io}^\ell$ and prepositions have type $\mathbf{j}^r\mathbf{jo}^\ell$ among others, thus tempting me to call them "transitive adverbs".

The minimalist program ultimately asserts that there are only two basic computational steps: *merge* and *move*. Pregroup grammars match *merge* by the contractions $a^\ell a \to 1$ and $aa^r \to 1$ of simple types. Some elementary illustrations of *move* are implicit in our dictionary rewriting rules. Take, for example, "affix hopping", which implies

ing work $\to$ *work* + *ing.*

This is treated here as an action of the inflector P_1, standing for "present participle", in the contraction

$$(\mathbf{p}_i\mathbf{i}^\ell)\mathbf{i} \to \mathbf{p}_i.$$

Also the movement

she did $\to$ *did she*

in questions is handled by our Metarule 1, which allows us to rewrite the type of *did* as follows in certain contexts:

$$\pi^r\mathbf{s}_2\mathbf{i}^\ell \;\Rightarrow\; \overline{\mathbf{q}}_2\mathbf{i}^\ell\pi^\ell.$$

I wish to remind the reader that such dictionary rewriting rules are strictly speaking unnecessary; their only purpose is to prevent overloading the dictionary with a large (though still finite) number of additional entries.

The operation *move* is frequently assumed to proceed from the end of a sentence to its beginning, leaving a trace behind. However, I do not believe that humans engaged in linguistic interaction can move items from the future to the past. In pregroup grammars such movements are replaced by double left adjoints. Even so, some of the old "island constraints" survive as "block constraints" in our terminology.

Much of the recent debate among mainstream grammarians is couched in the technical language of parsing trees. There have been a number of papers analyzing the minimalist program from an algebraic point of view, for example by Stabler [1997], and by several authors in de Groote et al. [2001], showing that minimalist grammars are weakly equivalent to multiply context-free ones. As far as I can judge, they still rely on an arboreal representation. While I cannot say that I have completely understood the minimalist program, I believe that in spirit, if not in technical details, pregroup grammars may be viewed as an attempt to carry out such a program.

36. Acknowledgements.

I am indebted to the Social Sciences and Humanities Research Council of Canada for its support and to the following:

Claudia Casadio for being the first to propose that non-commutative linear logic be used for grammatical analysis of natural languages;

François Lamarche for encouraging my burgeoning ideas and inviting me to present them at a conference in Nancy in 1997;

Michael Moortgat for keeping me on my toes with repeated critical challenges;

Avarind Joshi for pointing out the relevant pioneering work by Zellig Harris;

Danièle Bargelli for helping me to apply pregroup grammars to French, Turkish and Arabic;

Anne Preller for collaborating on applications to German and for investigating the categorical generalizations;

Wojciech Buszkowski for settling a number of theoretical problems concerning pregroups, in particular for showing context-freeness and cut-elimination;

Brendon Gillon for his active interest in this work and for encouraging his students to do so as well;

Telyn Kusalik for catching me on a combinatorial miscalculation and for realizing the relevance of Catalan numbers.

Larry Lambek for his critical reading of the early chapters of this book with a view of making them accessible to non-specialists;

Inspector Morse for insisting that "whom" should replace "who" when asking for the direct object;

Marcia Rodriguez and Phil Scott for helping with the bibliography;

Elaine Swan for patiently transforming my handwritten manuscript into LaTeX.

Finally, I wish to thank the McGill students who actively participated in three course-seminars on the application of pregroup grammars to English.

Index of names