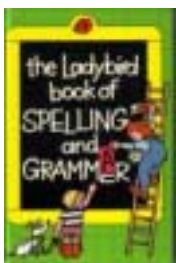


Modelling language

I have written this guide for students who are following GCE Advanced level (AS and A2) syllabuses in English Language. This resource may also be of general interest to language students on university degree courses, trainee teachers and anyone with a general interest in language science.

Models are everywhere

Almost any book that teaches about language will use some kind of model to represent the teaching to the reader. Here is a simple example – Dorothy Paul's Ladybird Book of Spelling and Grammar. But it's not that simple – this book contains things from which even quite sophisticated students and teachers can learn.



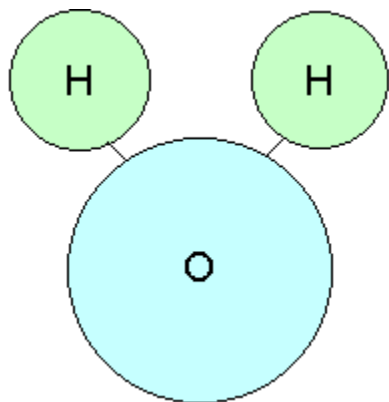
The book is organized into:

Abbreviations (front and back), the alphabet, alphabetical order, a sentence, nouns, pronouns, prepositions, conjunctions, adjectives, verbs, participles, auxiliary verbs, singular and plural, noun and verb agreement, apostrophes, sentence structure, spelling rules, silent letters, ways with words, prefixes and suffixes, punctuation marks, direct and reported speech, derivations, letter writing, synonyms and antonyms, proverbs, metaphors and the calendar.

Behind this set of contents lies an idea of what are the important things for people learning about language – to know how it works (grammar) and to be able to use standard forms (spelling and punctuation). You might like to think about where, if anywhere, the subjects in the list above would fit into the course you may be studying. Sometimes the names would be different – so where the Ladybird book refers to derivations you might know these as loan words, within the subject of etymology, which is a part of the study of the lexicon. But you would see a balance between knowing things (so that you can understand the underlying theory) and being able to use language in appropriate ways, including standard written forms.

What are models?

Models are representations of things that are complex. We use models to help us make sense of the complex reality, or attend only to those features of it that are of interest to our present enquiry. You are probably familiar with many models from all sorts of academic subjects that you have studied at school. For example, the following diagram shows a simple model of the molecular structure of water.

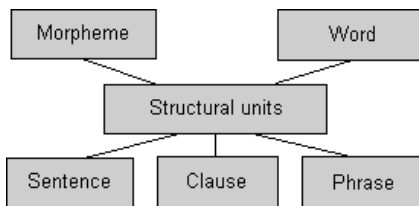


Other symbols, for example H_2O , can express the same model. Some people will find the picture easier to understand, while others can cope with the model expressed in letters and numbers. This becomes less easy, though, for more complex molecules, such as that for sucrose, $C_{12}H_{22}O_{11}$. And models for long protein molecules are almost impossible for many of us to understand. Note that these models for water concentrate on a few things (the proportion of hydrogen to oxygen atoms and the atomic links) and ignore everything else – such as the physical properties of water, its states of matter at different temperatures and pressures, its value to organic life-forms and so on.

Modelling language

As a student of language, you cannot make much progress without using models. In fact, it is almost impossible for your teacher to get by without models. How appropriate these are is another matter. And they may not correspond very accurately to the reality they seek to represent. In fact, they can be positively misleading. For example, you may sometimes be asked to analyse sentences in terms of their clause structure, then phrases within clauses, then the sequence of words in phrases, and finally the morphology of individual words. This is fine, so long as you do not reverse the process as a means of generating your own sentences (except as an experiment in finding out its impracticality). You do not speak or write by selecting a root, adding prefixes or suffixes, to build up whole words, then sequence them into phrases, then clauses and finally sentences. The way in which real people (children and adults) generate sentences is the subject of a special branch of study, and our knowledge is not at all complete – but we do know that some past theories are mistaken.

Models need not be one or two-dimensional – I might create an animated sequence as a model of something to do with language. And for some things (like language change) this would make sense, as the model would share with its subject the quality of change over time – except that I would massively reduce the timescale and the content. Here are a few models for you to look at.



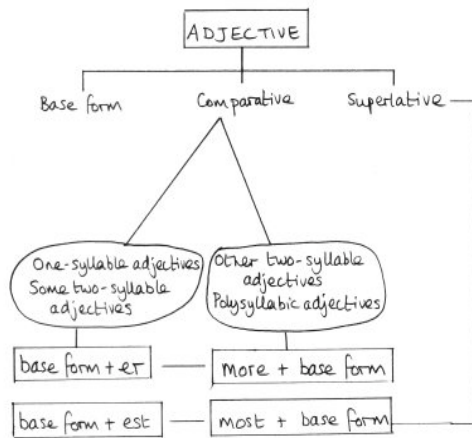
The first model is a familiar way of looking at the core structure of language in a series of hierarchies or steps. It is helpful because it is easy to remember, but it is not very accurate. A more accurate representation might be a Venn diagram in which units at each level would be enclosed by the next unit in the hierarchy.

The model is also dangerous because one might believe that morpheme, word, phrase, clause and sentence are fundamental and universal features of all natural languages, and that all languages organize them in the ways we know from English or other European languages. But these things are not true, or not wholly true. (Morphemes are fundamental and universal, but beyond this there are great variations in the ways in which we may combine them.)



This model is not very different from the contents of a reference book. The author has been economical and used a small typeface in order to present a description of English grammar on two sides of a printed card. Apart from a small diagram to explain personal pronouns, the text is essentially linear, though the use of headers and columns enables the reader to skip things and go straight to the subject he or she wants. In this much-reduced thumbnail version, you cannot easily read anything beyond the title (*The Fundamentals of English Grammar*). The red headings correspond to word categories or parts of speech.

Modelling language



This model is a two-dimensional diagram produced by a teacher (not me) before he began to teach GCE English language to sixth-form students. It is one of a series of more than 30 such diagrams, and is very much a draft. Unlike the hierarchical model of structure, this does correspond quite closely to what really happens – we do (in our speaking as in our writing) often adapt a root adjective form with an “-er” ending or by prefacing it with “more”. (In speaking we would normally be more tolerant of using either form, whereas in writing we might take time to select the more appropriate one, if this were not obvious to us.)

Are models the same as diagrams?

No, they are not – but for convenience we sometimes confuse the two. It is tedious to repeat that this diagram **represents** a particular model, rather than **is** it – but as the example of water shows, we may have different diagrams to depict the same model.

Diagrams and mental models

In trying to understand a subject, you will perhaps wish to learn models in the sense that you can recall them when you need to organize your ideas. Sometimes you may wish to learn a sequence of words or statements, while at other times you may wish to learn a particular diagram or other pictorial representation. You can, of course, attempt to do this in a time of revision, but it is much more easy to do, if you have been looking at the model or diagram for a longer time. How can you do this? By making posters for your classroom or smaller diagrams to display in the room where you normally study at home. A clever author may help you to learn a model by including it in the title of a book, as Shirley Russell has done with **Grammar, Structure and Style**. (For some kinds of assessed task, this three-part structure will help you to have a basic sense of what to do – but there is a limit to how much you can teach by making it the title of a book. For an example outside of language study, think of John Gray’s **Men Are from Mars, Women Are from Venus**.)

It makes sense for you to use diagrams that fit the way you think – they can be more verbal and linear or more pictorial and even colourful, while still representing the same basic model. As a teacher, I would often give a rudimentary diagram from a textbook to a student whom I knew to be a gifted artist, and ask him or her to turn it into a bigger and more visually appealing version for display in the teaching room. I would place the resulting poster behind my teaching area, so that over time students would come to know it – and perhaps make use of it in exams.

Modelling language

A very familiar way to recall models is to reduce them to mnemonics, such as acronyms. John Shuttleworth commends a model for analysing unfamiliar texts:

- subject
- purpose
- audience
- genre
- identification
- example
- comment

Mr. Shuttleworth reduces this to the acronym SPAG ICE. It is of no use, of course, unless the student who learns it can expand the seven headings into a more detailed procedure. It is a wholly appropriate model for GCE students because it lists things that the examiners have already (in writing the syllabus) identified as being important. It would be unsuitable for an academic linguist, perhaps, because it is essentially reductive – if we stick to this list we will never consider other approaches to analysis, but will become better and better at ploughing these particular furrows. (That is, the list is a model of a conventional method, rather than of universal truths about what texts are. An intelligent Martian, with a different value system, might ask wholly different questions about a text.)

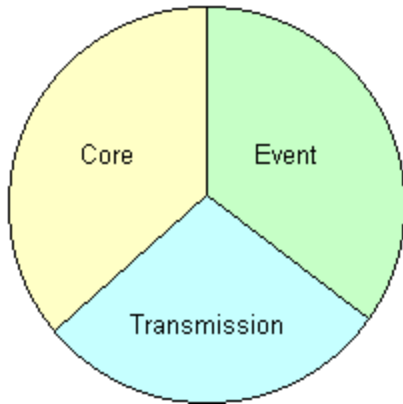
Models for academic linguists

I am not an academic linguist (one of the roads not taken earlier in life), and I may begin to look foolish if I pretend to know in detail the models that would underlie, say, a higher degree in general linguistics or in some specialized field, such as Sociolinguistics.

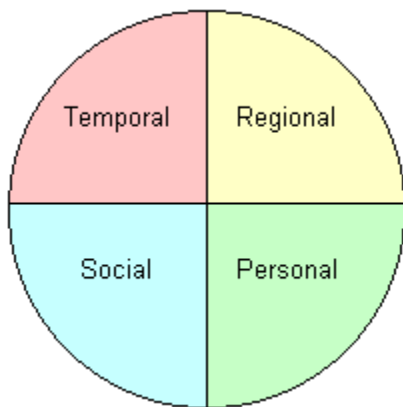
One way to find out, is to look at the syllabus or course-components for first and higher degrees – you may see these in university prospectuses in print or on a Web site. You quite possibly already have some prospectuses – see what they tell you. You may not know what all the names of different parts of a course mean – but at least you will see some things about which you could find out more.

A better way might be to look at the content of an encyclopaedic work – and fortunately there are some good examples available to you. Professor David Crystal is probably the most distinguished living linguist who has a general command of the subject (he does also specialize in particular areas of study). The contents of his reference books provide a kind of model – “kind of”, because print publication may impose some constraints on his abstract understanding of theory. Books are linear and something has to come first. Professor Crystal alerts the reader to this, in his [Cambridge Encyclopaedia of the English Language](#). The first pages of this book show us two models. One is a model of **structure**, the other of **use**. The models are presented as diagrams, containing headings, which are amplified in further text. That is, without the explanatory gloss, the diagram does not give the complete model.

Modelling language



Structure



Use

Core

- Grammar - rules governing construction of sentences
- Lexicon - vocabulary of a language

Event

- Text - a coherent, self-contained unit of discourse

Transmission

- Graphology - writing system of language
- Phonology - pronunciation system
- Sign - visual language used by the deaf

Temporal variation

- Long and short term language change

Regional variation

- How geography affects language

Social variation

- How society affects language

Personal variation

- How individual preferences and choices affect language

I have not copied Professor Crystal's copyrighted diagrams very closely – they are heavily stylized and appear to be projected from the twin mouths of Janus on a Roman coin (a rather odd distraction from the models of use and structure).

Other models can be inferred from the contents of language textbooks. Note that some writers will deliberately restrict themselves to a given area of study, such as history and language change (Robert McCrum: *The Story of English*) or stylistics (G. W. Turner: *Stylistics*). Here are some examples. David Crystal's *Encyclopaedia of Language* has these section headings:

- Popular ideas about language
- Language and identity
- The structure of language
- The medium of language: speaking and listening
- The medium of language: writing and reading
- The medium of language: signing and seeing
- Child language acquisition
- Language, brain and handicap
- The languages of the world
- Language in the world
- Language and communication

Modelling language

In the same author's [Encyclopaedia of the English Language](#) we find:

- Preface: Modelling English
- Part I: The history of English
- Part II: English vocabulary
- Part III: English Grammar
- Part IV: Spoken and written English
- Part V: Using English
- Part VI: Learning about English

In both cases, these are further broken down into chapters. Part III of [The Cambridge Encyclopaedia of the English Language](#) has:

- Grammatical mythology
- The structure of words
- Word classes
- The structure of sentences

In turn, Chapter 14 (The Structure of Words) divides into these sections:

- Morphology
- Suffixation
- Adjectives
- Nouns
- The apostrophe
- Pronouns
- Verbs

Professor Whitney F. Bolton's [The English Language \(Sphere History of Literature in the English Language, Vol. 10\)](#) has these section headings:

- Language and languages
- Phonology
- Morphology
- Syntax
- Lexis
- Style
- The early history of English
- The later history of English
- The study of English

You can see many other models by looking at textbooks for yourself.

Historical and contemporary models

If we go back in time to look at the models for studying language we may find that these are very different from what linguists study today – or rather, we may find

- that some things seem to have remained as important concerns,
- that some new areas of study have emerged, and
- that others have become less important.

This is not the same as noting that ideas **within** an area of study have changed. So grammar remains an essential concept – but our ideas of what it is, where it comes from and how it works have altered quite radically.

We can also say that models have become more scientific – that is, that language study is no longer dominated by those who advocate it as a means of social control or point of etiquette. Instead, we find that linguists methodically observe and record what happens and interpret it logically. This observation is often quantitative and objective. The interpretation uses agreed methods that other linguists can replicate in their own studies. And linguists share much common ground with other kinds of scientist who study human mental processes – psychologists, anthropologists, philosophers, paediatricians and sociologists.

If we take an area of theory within general linguistics (say, pragmatics) or a particular development within that field (say, speech-act theory), then we may find that it has appeared only in the last half-century. This does not mean that people suddenly started to use speech acts some time in the 20th century. But it does mean that this particular way of explaining what people have done for thousands of years is a recent development.

Or take Geoff Leech's Politeness Principle. We know that for courtiers in the 17th century there were rules both implicit and explicit about how to give and receive politeness. George Herbert (1593-1633) makes this clear in his poem, *The Pearl*:

"I know the wayes of honour, what maintains
The quick returns of courtesie and wit"

What Geoff Leech has done is to give a structured and coherent account, proposing names for some kinds of interaction. I could suggest that Leech's theory is useful but perhaps flawed, as there is no sincerity maxim – so I could practise his six maxims without meaning a word I say. But this is exactly why Geoff Leech's account is a good scientific model – because it is offered to all language students, who can, over time, add to, or qualify, the maxims.

Models for students

Students in secondary schools may have an interest in seeing language theory whole – which you can easily do by having a copy of one of Professor Crystal's encyclopaedic works. You will realize almost at once that perhaps the whole of one's life, let alone a few months of study at school, would be too short to learn about every part of linguistics.

Language teachers and examiners devise courses, both for schools and universities, which are very selective. They make decisions about what parts of this huge subject are both important and suitable for you to study. At the same time, they must find ways to measure your knowledge, abilities and understanding – which is not something that the authors of reference books usually set out to do. (The same people may, however, publish the reference books on Monday, write an exam paper on Tuesday, teach you on Wednesday, examine you on Thursday and mark your work on Friday, or something similar with a different time scale.) If your interest in modelling language is wholly about success in exams, then things become simpler and quite straightforward. You can work out models that more or less mirror the exam course.

Modelling language

At the highest level, the model IS the course structure – and it may be that discrete modules keep wholly in one area of study, so that the course as it appears to you sorts out language study into clear categories.

But at the next level down, there is the model of what you do within a particular part of the course, and the assessment activity (say an exam or written coursework) that is associated with it. Examiners will usually suggest this to you in the guidance that accompanies particular tasks, while teachers should give it for coursework. (Most exam boards back this up by publishing guides and notes for teachers.)

These models may be slightly artificial – for example creating a module for study called language development, which embraces child language acquisition and historic language change, which have little in common, and which would not be grouped in this way in a reference work (except, of course, for a reference work written explicitly to help students taking this exam).

Below is a simple three part model represented as a list of bullet points. It comes from the three assessed modules of a GCE course that concern areas of language theory. (The other three modules are presented in terms of tasks without specific content.)

- Introduction to the study of language – an overview of language theory and methods of study, as well as a clearing away of popular errors
- Language and social contexts – broken down into three representative contexts (language and gender, language and occupation, language and power)
- Language development – a handy way to combine two quite different areas of theory (language change and language acquisition)

The AQA syllabus A guidance for the module called Introduction to the Study of Language also provides two models.

One is in terms of variation (it omits temporal variation, which appears in Professor Crystal's model).

Three related concepts form the basis of study for this module.

- **idiolect**: the language style acquired by individuals as a result of their personal characteristics, systems of belief (ideologies) and social experience
- **dialect**: the variations in language produced as a result of local community and regional diversity
- **sociolect**: language variations produced by the effects of education, socio-economic class, systems of belief (ideologies), occupation and membership of any other social groups.

The other is a model of some of the ways in which we can explore and describe language data. They are not really comparable groups (like apples, oranges and pears) but different concepts that apply to the same data. If this seems hard to understand think how easily you can distinguish, say, the workings of a car's engine (grammar) from the different actions a driver can make (lexis) and how its movement is affected by the presence of other vehicles (pragmatics). This is not an exact analogy – but a loose way of showing that we see quite different things in the same data, according to our purpose. Examiners want to know that you can see them all, but in real-world linguistics, you might focus on one only.

Modelling language

The following descriptions of language will enable students to identify and discuss both the language system and its diverse uses and expressions.

- **lexis:** vocabulary choices available to users of English
- **grammar:** the fundamental structures and functions of English: the written word, spoken utterances, sentences and texts
- **phonology:** the contribution made by the voice to the phonetics communication of meaning, including intonation, rhythm, pace, volume, word stress and pauses in spoken English
- **semantics:** the ways in which meanings, implicit and explicit, are constructed and understood in English both spoken and written
- **pragmatics:** the ways in which social conventions and implied meanings are encoded in spoken and written language.

Investigating language – using models

Using models is helpful when you are investigating language. You can use them to illustrate some ideas or theory about which you want to find out more – that is, in explaining a specific area of theory, the task you hope to undertake or your research methods.

And when you have collected evidence (language data of some kind), you should use models to present your analysis. This short guide is not the place to expand this – but you may, for example, want to learn more about how you can use charts to show things that you have recorded in tables, using spreadsheet software.

Finding examples

This is easy but is a very useful task. Look at one or more reference books that set out to teach language. Start by reading the contents.

- What do these tell you?
- Which sections are about things you know (well, so-so, a little)?
- Which sections appear to be about things that you do not know?

You ought to have seen (ideally to have your own copy – and you can find one online) the syllabus for your course of study. Can you match the parts of the reference book to your course – on your own, or with help from a teacher?

Next look within a section or chapter. Try to skim or speed read, while noting what appear to be significant headings. A good author (or an author with a good editor) should use headings consistently to indicate the importance of a subject. You may not, however, have noticed this. Look again and work out what kinds of heading style correspond to what level of importance – now try to get an outline of the current section or chapter.

You can take this further by looking at the same subject in different text books – say language and gender or language change.

- Do the authors use the same, similar or dissimilar organization?
- Do they have a similar or different underlying structure?
- Can you account for these different approaches?

Modelling language

In creating your own model you may wish to take things from more than one source, or to keep to one that makes sense for you, with perhaps a few additions from others. For example, in drafting guides for this site on pragmatics or language and gender, my first task was to look at other writers' and teachers' approaches and to list their headings. Next I threw these into the air and watched how they settled – well not quite: I rearranged them, keeping some things and putting others into a quite different order. My version has the order I think most satisfactory. But because it is an electronic text with hyperlinked headings, the reader can move through it in more or less any order. (You can do this with a print text, but it is not as easy – and it seems to encourage you to take it in sequence. With a novel this makes sense. With a reference work it is rarely a sensible thing to do. You need to get used to changing the order of things, as you will certainly do this in your own coursework and exams.)

Move down a further step – and look for diagrams or other kinds of models on particular pages. The diagrams are easy to spot – they tend to be labelled **Figure 2: classes of nouns** or **Table 3: consonant combinations**. But sometimes models are lurking in the text of a sentence.

You may wonder why these authors' models are so different from each other. There are various reasons for this:

- There is no universally agreed model.
- Ideas of how to model language are constantly changing (as language scientists learn from other academic disciplines, such as anthropologists or psychologists).
- The models have different purposes and different audiences.
- Most textbooks (in order to be readable) will have linear or two-dimensional explanations of language - so the authors have to choose just one among many possible sequences.

A different vision of the topology of language science appears in Jean Aitchison's **The Language Web** (originally broadcast on BBC Radio 4 as the 1996 Reith Lectures). The metaphor of the web suggests multiple strands and interconnections. It may have been borrowed from the World Wide Web - the collection of linked sites on the Internet. This recognises that different theoretical explanations are interdependent. If you set out to discuss etymology, you are likely to find yourself also discussing semantics.

Adapting examples

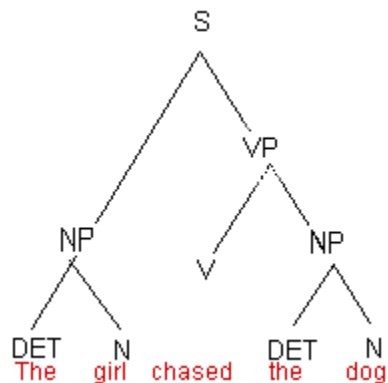
Not only is this more enjoyable than simply staring at the English language according to Professor Crystal – but he (and your teacher) would want you to do this. That is, take as a starting point any model in a textbook or other guide – and adapt it to produce your own.

In doing this, you may ask which kinds of models are the most accurate or helpful in the way they represent a complex reality. At times you may want the model to be complex (as this may include more things). At other times you will want it to be more simple, to show one thing more clearly. If you are in a school, then you may be able to get help from teachers of other subjects – such as maths or science.

If you wish to use charts or graphs, make sure you use an appropriate model. A pie chart is good to show how different things make up varying proportions of some other thing – the whole of which they are parts. But it isn't so good if there are hundreds of parts in the whole – you could not see them. The best place to find out more is to use the Help files in your spreadsheet program or to search online for guides.

Immediate constituent diagrams

Among many different kinds of diagram, some are of very great value. Constituent diagrams are very helpful for analysing the structure of short texts. (Beyond the level of a sentence, you will need a series of diagrams.) Here is an example.



Immediate constituent diagram

We see that the sentence (S) is divided into a noun phrase (NP), which is its subject, and a verb phrase (VP). The subject noun phrase splits into the determiner (DET), in this case "the" and the noun (N), "girl". The main verb phrase splits into the verb (V), "chased" and another noun phrase, which contains the determiner "the" and the noun "dog".

We could include further information within this same style of representation – for example, noting that the verb ("chased") is finite, transitive and in the past tense, or that the first noun phrase is the subject of the main clause, while the second is the direct object. We could replace S with C, since this sentence has a single clause. For compound or complex sentences we would have two or more nodes marked by C below the S at the apex of the diagram.

Lego and Meccano

Professor Crystal introduces his models of language use and structure by noting that his book

"...does not devote its space to techniques of moulding the English language in Play-Doh, Meccano or Lego. To model the English language is, rather, to provide an abstract representation of its central characteristics, so that it becomes easier to see how it is structured and used."

George Keith, a long-time chief examiner for English language, and author of many guides, writes (in e-magazine, April 2000) about a diagram that

"contrasts how language is put together (structures, the Meccano bits) and the uses (functions) to which language is put".

Modelling toys, such as those Professor Crystal lists, rely on our using the basic materials or parts to create complex structures. This is a loose similarity with some things that we do with language. However, it might be perfectly possible for you, if you are imaginative, to make 3-dimensional models to support your learning. Adapting Lego bricks (or simply writing on them) might be one way to learn about morphology – especially inflections or affixes.

Why do this, when you can perhaps show the same information in a sketch that takes a few seconds? No reason at all, unless you are spending more time, in order to become more familiar with the thing represented. Professor Crystal's diagrams do not need to use colour, nor do his colour choices (which differ from mine) reflect anything about real language – the core of grammar and lexis is not really blue, and regional language variation is not really yellow. But the colours may help some students remember the abstract idea in an exam. In this sense, if you want to use your own colour conventions, you may be wise in doing so – making a very neat diagram for revision may help you create the swift one even more swiftly under pressure.

Some dangers of models

Some models are so unlike the thing they represent that there is little danger of confusion. For example, in David Crystal's diagram showing different kinds of language use, we can have a sense of what change over time or regional variation mean – because we have observed the reality.

In the case of grammar, however, models for analysis seem so like the thing they represent that teachers have in the past commended the model as a way to make speech or writing conform to an idea of correctness. And some teachers of the very young have produced learning materials that reflect the belief that natural language comes from assembling smaller units, from morphemes up to sentences.

It may yet turn out that something very like this does happen – but as yet we cannot say this, and there are alternative ideas about how we generate utterances apparently spontaneously – but in reality by using familiar forms and adapting or transforming them, more or less.

In particle physics, scientists have long accepted that they cannot know by experience most of the subjects of the theory. Some elementary particles have properties that can be expressed only in complex mathematical models. For convenience, the physicists have given some of these properties names – such as “strangeness” and “charm”. (If you really want to know, quarks and hadrons have charm, while the charm of leptons and gauge bosons is zero.) But the words are not at all synonymous with charm and strangeness in their everyday senses.

While we are speaking or writing, our utterances do not organize themselves into immediate constituent diagrams – these are useful only to analyse already uttered or written data. (We could, if we were clever, produce them within seconds of hearing another person speak – but we could only be sure that the diagram was an accurate representation, when the utterance was complete.