Hirst, D. & A. Di Cristo. ms. *Proposals for an international transcription system for intonation.* Institut de Phonétique d'Aix.

Kingdon, R. 1958. *Groundwork of English intonation.* London: Longmans

Pike, K.L. 1945. *The intonation of American English.* Ann Arbor: University of Michigan Press.

Rischel, J. 1987. 'Some reflections on levels of prosodic representation and prosodic categories'. *Nordic Prosody* V, eds. K. Gregersen & H. Basbøll, 3-30. Odense University Press.

The principles of the International Phonetic Association. 1949. London.

A Systemic Model for Ellipsis

Barbara Cairns

Abstract
Brief background information on systemic theory is given first, followed by a model incorporating the theory and a discussion of the problems and drawbacks as well as the advantages of using the model. Some examples from the London-Lund corpus of conversational English are also included to illustrate how the model might function.

INTRODUCTION

The following is a preliminary attempt to describe ellipsis by means of a systemic approach. In order to help the reader who may not be familiar with systemic theory there follows a very brief outline of the theory (with the emphasis on brief) as far as it is relevant to the model for ellipsis.

The basic idea of a systemic theory is one of choice. Language is treated as an activity rather than an abstract entity and a systemic theory is based on choice between mutually exclusive possibilities. Whenever a speaker makes an utterance he makes a choice from a finite set of possibilities depending on the situation he finds himself in, what it is that he wants to communicate, relationship to the listener etc. The choices are not random, they are, rather, part of a system with entry conditions which then lead on to further choices. There are many systems within language and they operate at all levels of linguistic use. The idea of language as a set of systems was introduced first by Firth in the 1950's:

> Various systems are to be found in speech activity and when stated must adequately account for such activity. Science should not impose systems on languages, it should look for systems in speech activity, and, having found them, state the facts in a suitable language. (Firth 1957:144)

and has since been further developed by others, mainly Halliday (See e.g *Categories of the theory of grammar* 1961, and more recently *An Introduction to Functional Grammar* 1985, which is a basic introduction to the theory behind systemic linguistics as a whole.)

Systemic theory is, essentially, socio-orientated. Language is seen as a social activity in a specific social context and therefore most of the work

which has been done within it has been text-based or with the emphasis on spoken dialogue. It seems therefore a suitable model to try to adapt for the investigation of ellipsis which is also a feature occurring predominantly in less formal language situations such as dialogue.

The networks produced within systemic theory can be very complex but all work in the same basic direction – from left to right, increasing in *delicacy* i.e depth of detail. Individual systems may be on different scales. "The endpoint set to grammar on the exponence scale is where abstraction ceases: one has to move from abstract category to exponential item" (Halliday 1961:272). In other words, at the end of the scale an item of language is the exponent of an abstract grammatical category. The finest point in delicacy is where no further breakdown by grammatical categories is possible. This is where the lexis and syntagmatic lexical relations such as collocation take over.

A system is a representation of paradigmatic sets in language in which the choice is closed, to a greater or lesser extent, where paradigm is taken to mean "groupings of items on the basis of likeness in their own structure". Halliday describes grammar as deterministic, lexis as probabilistic. The choice of a specific lexical item is open, unlike the choice of grammatical category, but is nevertheless subject to collocational probability.

ELLIPSIS

Ellipsis has long been recognised as a normal feature of casual speech but attempts to analyse it along the same lines as regular speech have not been very successful.

> Ellipsis is a minus feature and is very stable historically. Slang is a plus feature and is absolutely unstable. Yet both signify the same: that the addressee, an insider, will understand what not everybody would be able to decipher. (Joos 1962:20)

There are different types of ellipsis, classified according to whether the information omitted is retrievable from the situation only or from within the text/dialogue itself. In Chomsky's early generative theory of language deletion is only permitted according to this 'recoverability principle' and it was not given much attention since it was regarded as a feature working on the surface structure only and the focus of interest at that time was mainly on the deep structure. Since then, however, ellipsis seems to have started attracting a lot more interest also within the generative school.

If it is only retrievable from the extralinguistic situation of the utterance an omission is often referred to as situational ellipsis (see Quirk et al. 1985). The model being presented here however is intended as a description of ellipsis within the text of a dialogue where the information omitted by the speaker is recoverable from elsewhere in the text. For instance in a co-ordinated sentence with *and* in some instances the second occurrence of a verb may be said to trigger omission, the necessary information being retrievable from the previous use of the verb within the same utterance. Information is carried across sentence boundaries too and need not necessarily appear in two adjacent utterances for ellipsis to occur, although it is less likely to do so than when the utterances are consecutive.

Ellipsis can be both interpersonal and individual, i.e. the information omitted from one person's utterance may only be retrievable from an utterance made by the previous speaker, so that the text as a whole needs to be examined rather than the speech of one individual.

NETWORKS DESCRIBING ELLIPSIS

The traditional approach to ellipsis, and the approach taken here, is that the speaker already has in mind a full utterance from which an element is elided. This approach entails the assumption that the topic (current topics) and some representation of the preceding utterance are accessible. I have developed the model on the basis of a large body of empirical material which includes the London-Lund Corpus of conversational English.

In the model for ellipsis what we are really describing is the set of conditions under which a speaker might choose to use elliptical constructions, omit items etc. Of course the ultimate choice of whether to omit or not will always remain open but the probability of omission increases greatly in certain situations. The reason for ellipsis in natural language appears mainly to be that of economy, but there are other conceivable reasons why a speaker might omit certain items of language. Some of the reasons are social, for instance a way of showing solidarity (or vice versa) with the listener by using a familiar, informal style or a way of indicating a less serious attitude to a situation. There are of course some speakers who habitually use an elliptical form of speech although they can usually provide the full form if requested to do so.

In fact we really need more than one network to describe ellipsis – one to illustrate the situational factors involved and a further one to illustrate the actual choices available at a grammatical level, since not all constructions lend themselves to ellipsis. An examination of the lexical choice pattern would require its own network, if indeed it would be possible to devise such a complex system

An entry condition leads into the network which is of a hierarchical nature so that a choice made to the right entails the choices already made to the left but not vice versa. The choices increase in *delicacy* from left to right – in other words they become finer in terms of the distinction between the choices until they reach a point, theoretically, where it is no longer possible to make a distinction. In the network below, for example (see Figure 1), the selection of the interrogative implies selection of *given* which implies the selection of the term *informal* to the left of it. The terminology used in the network is taken from the field of discourse analysis as this is the most suitable for the purpose in hand. It would of course be possible to devise a new set of terms but this would entail defining the new terminology explicitly before it could be used. It seems reasonable to use terminology which has already gained acceptance and some degree of familiarity. This is also desirable for the purpose of comparing the model with others of a similar kind.

At the end of the network a point should be reached where no further choices are possible other than the ultimate decision of whether to apply omission or not. This must be a final decision rather than an initial one since it cannot be made until the situation has been assessed by the speaker and the utterance has been, at least partially, planned. It must be emphasised that what the model is describing is an increase in likelihood for ellipsis – there is never any question of omission being obligatory.

The concept of paradigm is important to systemic theory since it is the paradigmatic relation of elements which is being described rather than the syntagmatic relations they enter into. In 1957 Firth referred to syntagmatic relations as 'relations of structure' and paradigmatic relations as 'relations of system'. A system might thus be seen as a representation of relations on the paradigmatic axis, a set of features which contrast in a given environment, whereas the structural description is describing syntagmatic relations. The two can then be said to complement rather than replace each other. "A system is (thus) a representation of relations on the paradigmatic axis, a set of features contrastive in a given environment" (Halliday 1961:264).

The question of choice in the network is also an important one, because there is a crucial difference between the options available at the beginning of the network and those at a more delicate level. On the left side of the network the choices are not directly under the control of the speaker since they are part of the extralinguistic situation. Moving to the right the choices become gradually more speaker controlled. One way of describing it would be to say that the network goes from external to internal. Also that the units gradually
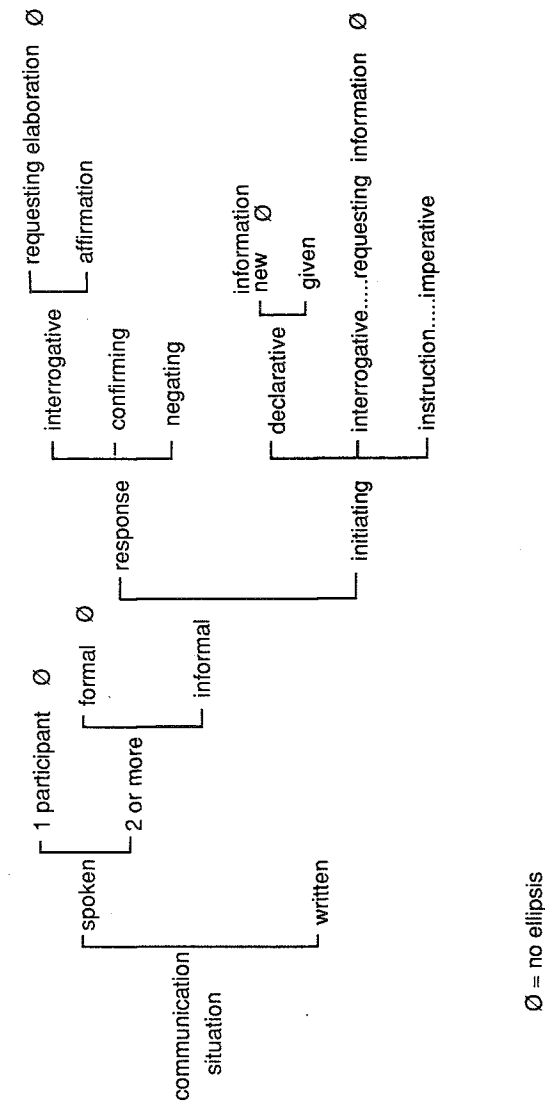
Figure 1. Model for ellipsis in dialogue, where ellipsis is more likely at endpoints not marked by Ø.

decrease in size from left to right. It should however be pointed out that this type of model in no way implies that the speaker is consciously aware of making these choices nor does the model claim to be a psychological representation of how a speaker plans an utterance, although further research may reveal that it does have some element of 'psychological reality'.

APPLICATION OF THE MODEL

A Ah....there's an excruciating shortage of seats in the library
B well there isn't actually
D well there isn't at the moment
B there isn't except in the summer term at the moment   (s3.3)

As an example of how the model might work we can apply it to the above extract from a conversation taken from the London-Lund corpus (Quirk & Svartvik 1980).

There are four speakers involved in a discussion about the university library generally (the fourth speaker makes no contribution here, however). Speaker A initiates a sub-topic discussion about seating problems in the library. Since this is a new topic the speaker cannot omit anything or the others would not be able to understand what the new topic for discussion is. Speaker B, however, now that the subject for comment has been introduced, can be sure that everyone present knows what is the subject under discussion and therefore omits the whole of the phrase *an excruciating shortage of seats in the library* in order to avoid a very laborious and unnecessary repetition. Speaker D then follows suit and omits exactly the same element, again to avoid tedious repetition. Speaker B makes a further comment but assumes that the others still have the same topic in mind since nobody has initiated a new topic, and therefore omits the old information about the shortage of seats but adds the new information that there is only an acute shortage in the summer term. Thus ellipsis is carried over three turns in the conversation. It would be inte-resting to see how many more turns it would be able to tolerate without the need to repeat the subject again but that is not the feature to be examined here.

If we refer to the model at this point we will see that it does in fact describe exactly the pattern we have observed in the extract above. The situational factors which are described at the beginning of the network hold true, which in itself is not so surprising, but even at a more delicate point of the network the model is applicable. Speaker A is initiating a topic for dicussion. We can

call this "new information" although it is only new in the sense of not having appeared in the conversation before – the other speakers obviously know about the overcrowding in the library but now it has been brought up as the topic for discussion. Since it is 'new' therefore, the speaker cannot omit anything, as the model predicts. Speaker B takes up the topic and adds a negative comment. What is new here is the negation, not the shortage of seats, so the subject may be omitted with no loss of understanding. Speaker D adds a proviso to B's comment, but again the *shortage of seats* is taken as given. The new information is that there is no shortage *at the moment* and this obviously cannot be omitted. Finally speaker B makes a further comment that there is only really a problem with seating in the summer term – still assuming the same topic but building on to what D has just said by inserting *except in the summer term* into D's comment, whilst retaining the rest of what D said.

A further example from the same source shows even more implicit elliptical speech and would in fact be impossible for an outsider to interpret accurately:

A is it this year that Nightingale goes
B no next year
A ehm... sixty (f)...........four sixty five
B sixty five   yeah
A I thought it was before sixty five so it's not until next year that the job
   will be advertised
B January   (s1.1)

This extract is from a conversation between two people who are obviously very well known to each other as work colleagues and therefore have a lot of common background knowledge. It is not until the end of the extract that it becomes clear to an outside listener that they are talking about another colleague who is due to retire and that the advertisement for his replacement will appear in January. The figure sixty five which they mention is in fact ambiguous since it could be the age of the person in question (sixty five is the usual retirement age) or it could be the year, nor is it ever really resolved during the course of the conversation, although it obviously causes no problem to the speakers. The economy in speech here, especially B's contributions, is about as tight as it could be without rendering the conversation totally incomprehensible.

The model correctly describes A's initiating question where nothing is omitted because it is an interrogative requesting information of the listener. B, on the other hand, is making a response to the request, in this case negating A's question, and also adding new information. The negation and new information may not be omitted but the given information about Nightingale's retirement does not require repetition and is therefore omitted, (For whatever reason, B seems to want to be as economical as possible in his speech.) A not only does not repeat his initial reference to Nightingale's retirement but simply refers to the figure sixty five, assuming that B knows what he is referring to and that he is approaching the age for retirement. This is really a comment on B's remark, so he is allowed to omit here. B confirms A's assumptions about their colleague's age by repeating it – he does not omit the age because A made a hesitation when he was speaking so B has to confirm which of his comments was correct. Therefore he does not simply say *yes*, as he might otherwise have done.

Now A replies with a fuller utterance because he wants to have some information confirmed (or negated) by B. However he still does not repeat the basic given information about N's retirement. B's reply is minimal, containing only the essential information that the job will be advertised in January by using one word only, since this is the only new information required by A.

Again, the model is able to describe quite well what is happening in the conversation, although it would never be possible to predict omission with 100% certainty since it is always, rather, a question of likelihood rather than an absolute.

ADVANTAGES OF THE MODEL
One major advantage of a systemic perspective on language is that it takes its material for investigation directly from actual text, spoken dialogue etc. and is thus a way of examining real utterances rather than studying isolated sentences which then have to be empirically checked for verification. The model allows the inclusion of pragmatic as well as grammatical features and this is relevant for ellipsis since there is a strong pragmatic element involved in the choices.

The model is dynamic rather than static and this seems intuitively right for one which endeavours to be applicable to spoken language, particularly dialogue. It describes a continuing process which may lead to ellipsis if the speaker chooses to apply it. Since ellipsis is never obligatory there will always

be the final decision whether to use it or not which remains open. If it is used the speaker must expect his listener to be aware that it has been used, otherwise misunderstanding can arise and Grice's maxim of cooperation is not fulfilled.

Most ellipsis in spoken language occurs in dialogue, often inter-sententially (see Carberry 1989) and the systemic model is able to capture this aspect in a way that a more conventional syntactic or semantic model could not. The model tries to approach ellipsis from the speaker's point of view and this also seems to be a profitable way of seeing it. Other researchers, including Carberry, are of the same opinion: "We contend that recognition of the speaker's discourse or communicative goal must be an integral part of any strategy that really understands ellipsis." (Carberry 1989:80).

Syntactic based approaches to ellipsis have generally been within the generative theory and have examined individual sentences on an abstract level without any consideration of the utterance as it might appear in context (for example Lobeck 1987), but such an approach must surely be of limited interest because it fails to look beyond the sentence to the situation of use and, in the case of ellipsis, which is a most obvious feature of spontaneous speech, it is essential to do so if we want to get the whole picture. Further, a syntactic based approach is not really sufficient even at an abstract level, because it is the repetition of the meaning of a word which determines whether it may be omitted or not and therefore semantics must invariably be taken into account.

Another advantage of the model is its hierarchical aspect whereby the choices available to the speaker are reduced to finer and finer distinctions and the likelihood of omission is greatly increased as one progresses along the network.

The model is primarily intended to represent the speaker's part in a dialogue but it may also be useful as an indication of the strategies used by the listener in decoding elliptical utterances. A shared knowledge of what is likely to be omitted together with a shared background knowledge of the subject under discussion are important for the understanding of elliptical speech by the participants in a dialogue.

In her article, referred to above, Carberry places emphasis on the importance of discourse goals for the interpretation of ellipsis. Examples of discourse goals which she identifies are *obtain information* and *seek clarify question*. Although not necessarily using the same terminology these kinds of goals are also able to be incorporated into the systemic model in the form of a

paradigm on the 'initiation – response' level, and because of the hierarchical system thus affect all subsequent choices.

As indicated by the discussion above a positive feature of the model is its flexibility – items can be added or removed without destroying the whole model, unlike other, more structural approaches to ellipsis. The entry conditions to the model may also be altered according to the requirements of the situation so that it is potentially adaptable for differing situations. The network can be made finer and more intricate or left at a more general level, depending on our requirements. It is also possible that the order of choices is sometimes interchangeable.

DISADVANTAGES

As can be seen from the model above, some items appear in more than one selection possibility. This may or may not be a disadvantage of the model – it implies an element of redundancy which could be seen as anomalous, especially in a model for ellipsis. It may in fact be necessary to adjust the model in a way such that this kind of duplication is eliminated. On the other hand the two occurrences of 'interrogative', for example, are arrived at by different choice routes and also lead to different options so it is difficult to see how they can be prevented from overlapping. In fact, with increasing *delicacy* in the network there arise more and more duplications which are dependent on different routes being taken. This is a less satisfying feature of the model and may imply that there should also be links within the system, between sets of choices. Superficially the problem could be eliminated by changing the terminology but that is not really a proper solution. On the other hand it may be so that it really is a question of two different types of interrogative and these do not necessarily conflict with one another in the system.

The placing of Ø in the model is perhaps a little misleading because it implies that ellipsis is impossible in these situations whereas that is obviously not the case. What it actually indicates in this model is that ellipsis is *unlikely* at such stages of the network. It should be possible in some more elaborate version of the model to indicate *degrees* of likelihood based on probability (see later comments).

Problems could arise with the interpretation of the model as a psychological reality i.e. that a speaker actually runs through such a system in his mind before arriving at the final decision to omit an element. The model does not make such a claim yet it is tempting to place such an interpretation on it. In actual fact the model is intended as a description of the choices available

to a speaker and the system is a neat way of representing them and also gives an indication of the interrelationship and interdependance which holds between them. The arrangement is not arbitrary but neither is it totally inflexible and it is possible that certain elements could be ordered in a different way.

It should not be assumed that because, from the speakers' point of view the system applies from left to right, the listener processes the information from the opposite direction. This would give completely the wrong results. The listener uses the same direction as the speaker in interpreting the message and what has been omitted from it.

A more serious theoretical problem with the model is one which was briefly referred to above i.e. that the 'choices' at the beginning, left hand side of the network are not really choices which are directly under the control of the speaker and it would arguably be better to have two separate systems such that one system would describe the relevant extralinguistic communication situation and then lead into a further system which elaborates on the actual linguistic choices available to a speaker. It seems in some ways unsuitable to have all the options in one model when they are obviously of a different nature.

One system would describe situation possibilities and the other would describe speaker options. On the other hand, since the dividing line between the two is fairly clear (in this model at least) it may be a question of appearance only as to which solution would look better.

Another feature which ought to be built into the model but has not yet been developed is what Lyons 1968:74 refers to as 'potentiality of occurrence' or acceptability. This would seem to play an important role in ellipsis, particularly for the listener who has to decode and understand the message and must therefore somehow fill in the omitted item from a limited number of possibilities. In many cases there is only one possible item, as, for example, when there is anaphoric reference in a text. But there are cases where there is more than one way of interpreting an incomplete utterance and it is there that potentiality or acceptability is important for the correct interpretation of the message by the listener.

It is clear from the second short conversation extract above that ellipsis often requires a combination of linguistic and extra-linguistic clues for its complete resolution. The extra-linguistic element obviously cannot be included in a model, but its contribution should nevertheless not be completely overlooked.

There are unquestionably a number of improvements which still need to be made to the model but these may not become obvious until it has been applied to sufficient examples of elliptical speech. It works for the instances tried above but it remains to be seen how wide an application it has. The model was described at the beginning of this paper as a preliminary one – this in anticipation of the many adjustments which will no doubt be made to it in later stages.

REFERENCES

Carberry, S. 1989. 'A pragmatics-based approach to ellipsis resolution'. *Computational Linguistics* 15, 75-96.

Firth, J.R. 1957. *Papers in linguistics 1934-51*. London: Oxford University Press.

Halliday, M.A.K. 1961. 'Categories of the theory of grammar'. *Word* 17, 241-292.

Halliday, M.A.K. 1985. *An introduction to functional grammar*. London: Edward Arnold.

Joos, M. 1962. 'The five clocks'. *International Journal of American Linguistics* 28, 9-62.

Lobeck, A.C. 1987. *Syntactic constraints on VP ellipsis*. Bloomington: Indiana University Linguistics Club.

Lyons, J. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Quirk, R. & J. Svartvik, eds. 1980. *A corpus of English conversation*. Lund: Gleerup.

Algorithms for Automatic Segmentation of Speech

Lars Eriksson

INTRODUCTION

A problem in speech recognition and also in automatic phonetic transcription from read speech is accurate segmentation of the incoming speech signal into syllable-sized segments. Several methods for automatic speech segmentation using computers have been developed. One common and also quite simple algorithm is to use the intensity from the original signal and the intensity from one or more bandpass filtered versions of the signal. These are compared using different criteria to determine the syllabic boundaries in the speech signal.

The syllable is linguistically defined as "a sequence of speech sounds having a maximum or peak of inherent sonority (that is apart from factors such as stress and voice pitch) between two minima of sonority" according to Robins 1966. The physical parameter that measures sonority would be intensity, and intensity could thus be used for segmentation of speech into syllables. The segments produced by a segmentation based on intensity are, however, not always the same as the syllables defined on a phonological level. The intensity segmentation often locates boundaries that cannot be identified as phone boundaries or even word boundaries.

Mermelstein shows in his work on the convex hull speech segmentation method that it is more convenient to talk about 'syllabic units' rather than syllables when working with automatic segmentation. A syllabic unit consists of the syllable nucleus, that is, the vowel, but may also contain consonants from other neighbouring syllables, especially from intervocalic consonant clusters. For example the words and syllables, *piles of* [pajlz ɔf] would form the syllabic units [pajl] - [zɔf].

One method for speech segmentation using unfiltered speech intensity and bandpass filtered speech intensity has been developed at Lund. This method is used within the project 'Prosodic Parsing for Swedish Speech Recognition',