

## The Frame-for-Term Approach to Terminology in Eurotra

Archibald Michiels

Volume 39, numéro 1, mars 1994

La traduction et l'interprétation dans la Belgique multilingue

URI : <https://id.erudit.org/iderudit/003648ar>

DOI : <https://doi.org/10.7202/003648ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Michiels, A. (1994). The Frame-for-Term Approach to Terminology in Eurotra. *Meta*, 39(1), 184–193. <https://doi.org/10.7202/003648ar>

Résumé de l'article

Le projet de traduction automatique Eurotra a choisi le champ des télécommunications pour illustrer la typologie des textes à traduire. Une telle décision est lourde de conséquences. En particulier, un sous-langage montre des restrictions langagières qu'un système devrait utiliser pour assigner les bonnes structures aux textes et y choisir les informations pertinentes. Il s'agit en fait d'essayer de passer des termes aux concepts, c'est-à-dire de parties textuelles à la connaissance qu'elles renferment. L'article analyse les diverses procédures et montre les cheminements à suivre.

# THE FRAME-FOR-TERM APPROACH TO TERMINOLOGY IN EUROTRA<sup>1</sup>

ARCHIBALD MICHIELS

*University of Liège, Liège and ISTI, Brussels, Belgium*

## **Résumé**

*Le projet de traduction automatique Eurotra a choisi le champ des télécommunications pour illustrer la typologie des textes à traduire. Une telle décision est lourde de conséquences. En particulier, un sous-langage montre des restrictions langagières qu'un système devrait utiliser pour assigner les bonnes structures aux textes et y choisir les informations pertinentes. Il s'agit en fait d'essayer de passer des termes aux concepts, c'est-à-dire de parties textuelles à la connaissance qu'elles renferment. L'article analyse les diverses procédures et montre les cheminements à suivre.*

## **BACKGROUND ASSUMPTIONS**

The EUROTRA machine translation project has chosen the field of telecommunications to provide the type of texts that it intends to show it is able to translate. Such a decision has far-reaching consequences. In particular, a sublanguage exhibits restrictions that the system should exploit in order to assign the right structure to pieces of text and select the relevant readings of the lexical items therein.

Some of these restrictions relate to text type rather than subject matter: for instance, the total absence of second person personal pronouns. Sentence 1 is no more likely to occur in the texts the Eurotra project has chosen to concentrate on than 2, although 1 is appropriate from the point of view of the knowledge it expresses:

1. *You should take into account the partial loss of signal due to rain attenuation.*
2. *Would you care for another cup of cocoa, Miss Snappet?*

We will not deal with such text-type-related restrictions here, but instead will deal with those that relate to the knowledge that domain-specific texts express or — more importantly — assume. The frames for terms proposal, sketched in the remainder of this paper, is a modest approach to tapping domain-related knowledge for the purposes of structure assignment and reading assignment, the two main tasks of source text analysis. The hypothesis is that the best way to integrate domain-related knowledge into a basically linguistic framework is to anchor it to the terms pertaining to the domain, and to the environment of these terms, itself often made up of other terms.

It should be clear from the very nature of the task the Eurotra project is tackling that we move from terms to concepts, *i.e.* from pieces of text to the knowledge they embody. Yet this should not hide the fact that logical primacy ought to be given to the concept, which exists prior to the term that gives it expression. In a given domain, the relation of concept to term is a one to one relation and does not admit of either indeterminacy (vagueness) or ambiguity (polysemy). A concept entertains well-defined links with the other terms in the field. The term's definition is a crucial source in establishing the number and nature of such links, though it is not the only one: the study of a corpus by a linguist assisted by a domain expert is also crucial to determining the knowledge-based frame of a given term.

Faced with a text from the telecommunications domain, such as those in the Eurotra corpus (the CCIR collection — see References), the linguist is, understandably, somewhat at a loss. He has fewer intuitions than in the case of general language. He will find, for instance, that the assignment of syntactic frames to terms is a highly difficult task. He may feel comfortable dealing with subject and object relations for the terms that are deverbals, but will be tempted to relegate all prepositional phrases (pps) to the *modifier* class, *i.e.* regard them as *non-arguments*, *i.e.* outside the *valence* of the arg-bearer. General semantic relations for modifiers will similarly prove difficult to assign, let alone compute. A purely linguistic analysis will end up assigning a less dense structure to a domain-specific technical text than to a general language text. This means that knowledge, although available through concertation with domain experts, is left untapped.

But why should SYNTACTIC frames enjoy a special status? Because the linguist knows them better, and is on surer ground assigning them? Because they are deemed to be more reliable, *i.e.* more consistently assignable?

We need to go briefly into these points to sketch the background to the frames for terms proposal.

#### WHY ARE TERMS SPECIAL, OR ARE THEY?

It may be appropriate to start with Harris' contention, discussed in Sager 1986, that 3 is not a sentence of biochemistry, although it is a sentence of English (it would not be starred on purely linguistic grounds):

3. *Hydrochloric acid was washed in polypeptides.*

The first thing to agree on is that we have to account for the sentences in our domain, *i.e.* telecommunications or one or several subdomains thereof, and not the sentences of English. In other words, we are ready to exploit the further constraints that the sublanguage of the domain places on acceptability. Our main purpose is, as usual, disambiguation, or, to express it in more positive terms, reading assignment.

If we agree on what we are trying to achieve, then it becomes easier to define the means for achieving it. In 3, we can capture the violation in terms of the frame assigned to WASH (the domain-specific reading), *i.e.* specifying the proper slot fillers for argument 2 (the deep object) and the head of the IN-phrase. Frames for terms are not in any way special: they reflect the LEXICALLY DETERMINED environment of a given item under a given reading, in casu WASH as used in biochemistry.

What may look special at first sight is the restriction we can impose on the slot fillers. We can use domain knowledge (in our framework simply embodied in features).

It is not so much that general English is unconstrained while the technical sublanguages are highly constrained; it is rather that we cannot formalize the constraints general English obeys (the knowledge is too common sense, too diffuse), whereas we have some ideas on how to formulate the constraints that obtain for a given domain.

These constraints can best be captured if we centre our attention on the domain's terms and their frames, captured with the full apparatus of features that domain knowledge enables us to specify.

But before we go any further we must be quite clear about what we understand by FRAME.

#### WHAT ARE FRAMES?

Economical, elegant distributional statements reflecting LEXICALLY GOVERNED distributional behaviour of classes with more than a single member.

### Limiting cases

When the class boils down to a single member, we are better served regarding the governing and the governed items as forming a single multiword unit. For instance, *HAVOC* in *PLAY HAVOC WITH* is not part of the frame of *PLAY*, but part of the multiword unit *PLAY HAVOC*, which itself is assigned a frame, so that the *WITH*-phrase may be catered for.

When the distributional statements cannot be anchored to lexis, we have no use for frames, because there is no way of stating what the governors are. For instance in 4:

4. *In Paris he decided he had better recant.*

in order to account for the distribution of the *IN*-phrase, we do not specifically refer to *DECIDE* or to a class that includes *DECIDE*; we simply say that *EVENTS* are locatable in space and time. The act of deciding is an event; so is the occurrence of accidents, and we can have:

5. *In Paris there are at least a hundred road accidents a day.*

Of course, there are non-event verbs; hence the strangeness of

6. *In Paris my latest book contains six chapters.*

(reinterpreted as: *they think that..., I pretend that...*)

However, we cannot readily pin events down to a lexical class, via a *+EVENT* feature that we could assign to lexical items, and we therefore link the *IN*-phrase to the whole *S*, as a modifier.

Contrast this treatment with that of the *THAT*-clause (with optional complementizer) in 4. We can link it to lexically storable classes of verbs and nouns, and we therefore say that the *THAT*-clause is governed by *DECIDE*, *i.e.* belongs to its frame.

### HOW DO WE CHARACTERIZE FRAMES?

In some cases, syntax alone suffices. We say that *DECIDE* governs a *THAT*-clause, without having to further determine the make up of the *THAT*-clause. Although the reading of *DECIDE* is not the same in 7 and 8, we do not generally try to reflect the difference in the frame:

7. *He decided that he would go there.*  
8. *He decided that he had been there.*

However, in many cases we do not rely on syntax alone to provide a characterization of frames; we very often use a mixture of syntax and lexis, as when we state that *DECIDE* governs a prepositional phrase (syntax) whose preposition is *ON* (lexis). We stop there in the case of *DECIDE*, since it is not possible to further restrict the class of NPs that can function as heads of the *ON*-phrase:

9. *He decided on Mary.*  
10. *He decided on the boat.* (with arg reading of “*on the boat*” — shades of *Aspects*)  
11. *He decided on what the others had already decided on.*

Often, though, it is possible to further restrict the syntactic (or lexico-syntactic) class that will be assigned as frame element to a given governor. The nature of these further restrictions can broadly be called semantic, with various degrees of fine-grainedness, from *+ANIMATE* or *+ABSTRACT* to assignment to a domain-determined thesauric class.

The point to remember is that if an element (however we characterize it) is lexically governed by another, the frame of the governor is the place to capture that government. Syntactic frames have no claim to a special status, which would make them more easily or more consistently assignable than knowledge-based frames. In particular, the linguist is unable to assign knowledge-based frames, but should not for that very reason regard them as suspicious (more idiosyncratic, less reliable).

#### WHAT ARE FRAMES FOR?

They contribute to enabling a parser to adhere to the *densest match first* principle, lexically determined ties having priority (an *arg* reading is denser than a *modifier* reading, *ceteris paribus*).

The densest match first principle is itself a way of cashing in on the redundancy of natural language. If we find X and Y together, and X has a reading where it requires a Y-like element in its frame, and Y has a reading on which it fulfils the properties required to meet the conditions for being a frame element of X, and X and Y are syntactically related the way the appropriate frame for X requires, then we (the parser) note the density of the X-Y link if the appropriate readings of X and Y are selected (the redundancy is greater, information content is less), and we therefore select them. Thereby we determine what goes with what, accomplish lexical disambiguation, choose the right preposition reading, etc.

#### HOW DO WE ESTABLISH FRAMES FOR TERMS?

Knowledge-based frames can be determined only through cooperative work by a linguist and a domain expert. On the basis of a representative corpus, the linguist can explore the grammatical-lexical environment of the terms, and put forward tentative distributional generalizations. The domain expert can determine which pieces express relations that properly belong to the concept represented by the term.

Take for instance the concept of COMMUNICATION as used in information theory (a concept linguists are certainly familiar with). There is no cogent syntactic reason to say that the relation OVER + CHANNEL belongs to the term more than the relation expressed, for example, by IN + YEAR: in a syntactic treatment both would be assigned to the modifier class.

From a knowledge engineering point of view, however, the difference between the two modifiers is striking: OVER + CHANNEL is a constitutive part of the concept of communication in information theory, whereas IN + YEAR is extrinsic to the concept.

In the frames for terms proposal, the OVER + CHANNEL would be regarded as a PLOT, *i.e.* a *paradigmatic slot*, a slot opened up by the term COMMUNICATION and best capturable in terms of a paradigmatic description of the slot fillers, *in casu* the class of CHANNELS. The IN + YEAR would receive modifier treatment, just as in the syntactic frame assignment exercise.

If we detach the OVER + CHANNEL slot from the term COMMUNICATION (as in the modifier treatment), we have to rely on the conjunction of preposition and type of np in the pp to establish the proper interpretation and translation of the prepositional phrase. We deprive ourselves of a precise link (that between an act of communication and its channel) and rely on the putative monosemy of the OVER + CHANNEL group.

It should be clear that the knowledge-based approach applies to the terms of a given well-defined domain. To go back to the lexeme COMMUNICATION, it is also used as a term in a totally different domain, where it would be assigned a totally different frame, along the lines of:

**BY author ON/ABOUT topic AT conference**

exemplified by: *a communication on terminology in Eurotra by S. Perschke at the IBM 1989 European Institute*

#### HOW DO WE INDEX FRAME ELEMENTS?

Given that the domain expert and the linguist working together can assign a knowledge-based frame to a term, we must still determine how indexing the frame elements is to be achieved. Plots are indexed by three means:

- 1) lu's (lexical units) of prepositions
- 2) deep syntactic information
- 3) subject field code

It can easily be shown that subject field codes cannot be replaced by domain-independent general semantic features. The reader is referred to the case study of ILLUMINATE, annexed to this section.

Pforms (the language-specific forms of prepositions) are wholly language-dependent, and it would be nice to find a way of interlingually coding frames for terms. Such an endeavour is not wholly devoid of a-priori plausibility, since the world described is the same, the relations are domain-bound and the various languages are not left free to view reality in strikingly different ways, along Whorfian lines, as this would lead to scientific and technological chaos.

*Semantic relations* would seem to be the obvious candidate. However, the usual sets (with SOURCE, GOAL, PATH, and the like) may not be very helpful. It is not at all clear how we can avoid working with relations that are domain-specific, or even subdomain-specific, or even sub-subdomain-specific,... perhaps sometimes down to the level of the lexical item itself.

The third way of indexing frame elements is via a classification of the possible *fillers* for a given frame element. Here domain knowledge is crucial.

A problem with any classification scheme is the degree of delicacy it must exhibit. This degree depends on the purposes of the classification and is very hard to decide on a priori. We have proposed a three-tier classification scheme, where the third tier can be seen as somewhere between category assignment and domain-related, but nevertheless general, semantic features.

A problem with the coding scheme proposed here is that a scientifically revolutionary text in the selected field could well prove unanalysable, if we stipulate too narrowly the classes that serve to index the args. In other words, the indexing scheme should leave room for things that have not yet been said; at the same time, it should be precise enough to be able to filter out non-term uses of strings that look like terms but are used in a (lexically, not syntactically) different environment. The task is difficult but not impossible.

#### CASE STUDY I: ILLUMINATE / ILLUMINATION

This short case study aims at demonstrating that we need domain-dependent knowledge (not just semantic features of the common-or-garden variety) to properly disambiguate the term ILLUMINATE and ILLUMINATION as used in our field, *i.e.* telecommunications. Consider 10-15, excerpted from our present corpus, to wit CCIR Vol.4, Fixed Services:

10. *the Westar satellites which illuminate the continental United States, Hawaii and Puerto Rico...*
11. *the SBS satellites illuminate the continental United States*
12. *the attenuation of further side lobes, which illuminate the Earth...*

13. *SATCOM illuminates the continental United States, Alaska and Hawaii, while COM-STAR illuminates the continental States, Hawaii and Puerto Rico*
14. *if all satellites utilizing the geostationary-satellite orbit, illuminating the same geographical area....*
15. *received by a satellite antenna which is illuminating the earth...*

Even the casual reader will have come to the conclusion that we are dealing with a domain-dependent reading of ILLUMINATE here, one which has got nothing to do (or not that much, anyway) with the shedding of light. I may be looking for a lost key on a pitch dark night, and the fact that the area I am looking for my key in is illuminated by satellite so-and-so won't help me much...

Unfortunately the other ILLUMINATE (the one we all know about, or at least a reading nearer to that one...) is also used in telecommunications, namely in fibre optics, as in 16 and 17:

16. *To illustrate how a fiber operates, see Figure 4.5. A light source is illuminating the end of a short section of pipe. (Telecommunications Networks, p.77)*
17. *In such applications, an entire bundle of fibers is illuminated, rather than a single fiber. (Digital Communications, p.6)*

Needless to say, ILLUMINATION, a deverbal, follows the tracks of the ILLUMINATE it nominalizes:

18. *Fig. 2.2. Maximum illumination of the earth by a geo-synchronous satellite (Digital Communications, p.56).*

And since the merits of fibre optics and communications satellites are often debated in telecommunications, the two readings cannot be kept separate on the basis of determining, however finely, the text's topic.

We are left with two solutions:

- 1) the non-solution: hope and pray — and even check — that in all Eurotra languages — and in all those that are ever likely to be added to the nine we must deal with at present — the distinction between these two readings of ILLUMINATE is translationally irrelevant;
- 2) work out a way of distinguishing between the two — it must (epistemic must) be on the basis of arg1 (deep subject) and arg2 (deep object) fillers, and it must go beyond what we can capture via semantic features of the usual kind. For instance, the first reading of ILLUMINATE is broadly characterizable as such:

- arg 1: SATELLITE (the generic term), names of satellites, relevant part of satellites (expert knowledge is necessary here — I'm afraid I can't provide it)
- arg 2: a geographical area

It would seem that the most elegant answer to the disambiguation task here is to allow ourselves to tap domain-dependent knowledge as embodied in features on slot-fillers in frames.

#### CASE STUDY II: TRANSMIT AND TRANSMISSION

It is of course extremely difficult to report on theory-free observations, since they do not exist. However, analysis of our corpus (CCIR Vol. 4) enables us to state certain requirements that must be met by our treatment of the environment of the items TRANSMIT and TRANSMISSION.

**TRANSMIT**

Deep subject (arg1)

Fillers (all given fillers are corpus-attested; they refer to lemmas, not wordforms, unless otherwise stated): *satellite, network, station, antenna, channel, system*

Deep object (arg2)

Fillers: *telephony, television, program, indication, frame, sample, channel, stream, information, message, alarm, burst, phase*

FROM-phrase

Fillers: *station, link, area, space*

TO-phrase (TOWARD also attested in addition to TO)

Fillers: *antenna, satellite, exchange, repeater*

IN-phrase

Fillers: *band, bit, channel, frame*

ON-phrase

Fillers: *frequency, link*

BY-phrase

Fillers: *reflection, FM, time-division multiplex, CPSK, radio channel*

OVER-phrase (perhaps THROUGH and VIA are other exponents of the same relation, expert knowledge is necessary to decide)

Fillers: *system, satellite, channel, line, link*

Fillers in THROUGH-phrase: *link, system, channel*

Fillers in VIA-phrase: *satellite, DSI* (digital speech interpolation) (VIA refers to the same relation as BY in this last case)

AT-phrase

Fillers: *rate, frequency, speed*

BETWEEN-phrase

Filler: *station*

**TRANSMISSION**

Arg1: realizations:

1) prepositional phrase with pform = OF  
fillers: *station*

2) prepended n  
fillers: *satellite, station*

Arg2 realizations:

1) prepositional phrase with pform = OF  
fillers: *traffic, data, signalling, television, programme, radio wave, channel, picture, service*

2) prepended n fillers: *picture, voice, alarm, information, data, sound, message, telephony, telegraphy, signalling, signal*

IN-phrase:

fillers: *service, system, network*

OVER-phrase:

fillers: *hop, system, link*

(BY / VIA / WITH / THROUGH: fillers: *satellite, link*)

BY-phrase:

1. METHOD

fillers: *colour-separation process*

(METHOD relation also expressed via prepended nominal group, as in:

“*time-division multiple access transmission*”

“*FM-SCPC telephony transmission*”

(note arg ordering: *method* precedes arg2)

2. MEANS

fillers: *telephone, videophone, telegraph*

(MEANS relation also expressed via prepended n fillers: *telephone, videophone, telegraph*)

AT-phrase:

fillers: *interval, rate, power*

(relation also expressible through prepended nominal group, as in:

“*high speed data transmission*”

(note arg ordering again: *rate* precedes arg2)

FROM-phrase:

fillers: *antenna, network, section, station*

TO-phrase

fillers: *area, end, destination, station*

I wish to concentrate on the verb TRANSMIT first. As is clear from the fillers, if we think of frames as a way of accounting for lexically determined environments, we have to regard all 10 relations as belonging to the frame of the verb TRANSMIT. In case the reader is wondering whether the present approach has any use for modifiers, let him rest assured that it does; the criterion is lexical determination, and for instance the capitalized phrases in 19 are modifiers:

19. *UNDER CERTAIN CIRCUMSTANCES, the same satellites may also transmit separate information on the same frequencies TWICE OVER OR EVEN A GREATER NUMBER OF TIMES.*)

Arg1 and arg2 can be indexed on the basis of the syntactic position they fill, namely deep subject and deep object. What about the other 8?

It is obvious that pform alone is not sufficient; as mentioned in the case of the BY-VIA pair, they can point to the same relation, one we could perhaps refer to as METHOD: *by CPSSK (coherent phase shift keying) via DSI (digital speech interpolation)*

If pforms are not adequate, or not enough, what can we then turn to?

#### a) semantic relations

The usual sets (with SOURCE, GOAL, PATH, and the like) may not be very helpful. In the case of TRANSMIT, we would presumably be able to account for the FROM (source), TO (goal), BY/VIA1 (method), OVER/THROUGH/VIA2 (path) phrases, but the others (IN, ON, AT, BETWEEN phrases) are not so obvious (a relation such as location is clearly inadequate to operate the necessary distinctions). It is not clear at all how we can avoid working with relations that are domain-specific, or even refer to the lexical item itself (in casu TRANSMIT and its nominalization TRANSMISSION).

#### b) classification of the possible fillers for a given frame element.

The situation is not too promising to start with: we find the same lexemes (*channel, frame, frequency, link, satellite, system*) filling different slots, so that a reference to a class, even conceived as a list of possible lexemes, has not enough discriminatory power. Again, we must probably go down to the level of reading, where it might appear (again, expert domain knowledge is necessary) that in 20 and 21 we have to do with 2 different readings of the lexical item CHANNEL:

20. TO TRANSMIT A CHANNEL (as in: "The station uses an adaptive delta-modulation/4-phase PSK/SCPC system and can transmit one two-way telephone channel.")
21. TO TRANSMIT IN A CHANNEL (as in: "It may be useful to transmit the dispersal waveform in a subsidiary channel as an alternative to regenerating the waveform locally.")

It seems that we might well have to use a mixture of all three methods:

- 1) syntax cum lexis: subject, object, prepended nominal group, prepositional phrase with pform X in language Z;
- 2) domain-relevant semantic relations: source, goal, method, rate, ....;
- 3) assignment of possible fillers to class to achieve indexing of frame elements.

It would be nice if a mixture of 2 and 3 were enough, because then we could still index frames interlingually. It is proposed here that such an indexing scheme be tried out for a subdomain, once expert domain knowledge has been tapped to provide at least a tentative list of the relations to be used and a classificatory system thought to be of the right degree of delicacy to achieve the purposes of frame element indexation within the selected subdomain.

#### NOMINALIZATIONS

TRANSMISSION is very clearly a deverbal, and — totally or only partly — inherits its frame from that of the parent verb, TRANSMIT.

The first thing to keep in mind is that terminology is basically geared towards nouns (or, in the majority of cases, nominal groups), rather than verbs. This reflects a more general bias towards nouns and nominalizations in technical language itself. We often find a support verb and a deverbal instead of the verb, as in 22:

22. *transmission of a burst is done at the bit speed used by the TDMA system*

This may lead to richer frames for nominalizations than for the corresponding verb.

A second thing to remember is that nominalizations can exhibit a 0 frame realization, *i.e.* all frame elements are optional. 23 is an example of the 0-frame use of TRANSMISSION:

23. *nevertheless everyone (...) will express his or her opinion on transmission (Telecommunications Networks)*

(Note that in this case the parent verb too can exhibit a 0-frame realization, as witness

24. *a common antenna system serves for transmitting and receiving*

and is often used with arg1 only:

25. *for the earth station when it is transmitting*  
 26. *any user equipment may start transmitting (Digital Communications, p. 113)*

This intrinsically variable number of filled arg slots leads to problems in the case of non-saturated frames. It should first be mentioned that, outside constructed examples, the frame for a verb like TRANSMIT will never be fully saturated, as this would lead to a very poor communication strategy (trying to say it all in a single sentence). However, there are degrees in the unsaturatedness of a frame. We might find that two prepositions are kept apart (convey a different relation) when they are both used as frame element markers in a single sentence, but are interchangeable when used alone in less saturated frames.

Nominalizations create their own indexing problems. As is clear from the frame for TRANSMISSION, a prepended noun or nominal group may express very different relations (arg1, arg2, method, rate, means), and an OF-phrase can express both arg1 and arg2, although not simultaneously, of course. The indexing must therefore call on additional clues, like ordering (in the case of prepended elements the core args — arg1 and arg2 — keep as near the deverbal as possible) and whether a relation is satisfied somewhere else in the clause. We again observe that the more saturated a frame is, the easier to index are its elements:

*frequency modulation*  
*colour television transmission*  
*high speed data transmission*  
*satellite transmission of television*  
*television transmission by satellite*  
 \* *television transmission of satellite*

#### Note

1. This paper owes a lot to brain-storming sessions within the then Eurotra Dictionary Task Force, in particular to discussion with *Henrik Sørensen* from EUROTRA-DK (Denmark). *Eric Lee* (ISTI, Brussels) suggested stylistic improvements, not always minor ones. I wish to thank them both.

#### REFERENCES

- HARRIS, Z. S. (1968): *Mathematical Structures in Language*, New York, Wiley  
 SAGER, N. (1986): "Sublanguage: Linguistic Phenomenon, Computational Tool", in Grishman, R. and Kittredge, R., (Eds), *Analyzing Language in Restricted Domains: Sublanguage Description and Processing*, New Jersey, Lawrence Erlbaum

#### Corpus

- CCIR Report, 1982, Vol IV: *Fixed Service by Satellite*, Geneva  
 Digital Communications — BARTEE, T.C., (Ed), *Digital Communications*, Indianapolis, Howard Sams, 1986  
 Telecommunications Networks — MURPHY, R.J., *Telecommunications Networks: A technical Introduction*, Indianapolis, Howard Sams, Indianapolis, 1987