

2 Framework

Wolfgang Klein and Clive Perdue

2.1 Introduction

The procedure followed here¹ results from the general perspective on the acquisition process which was sketched at the beginning of the first chapter. The central idea, to repeat it briefly, is the following. The acquisition of a language does not consist in the addition of grammatical or lexical features to the existing knowledge of the learner, for example in the addition of TL specific word order rules ("adjective follows noun"), TL specific case marking ("dative is marked by umlaut") or TL specific category distinctions (say between animate and inanimate nouns). It is not like putting a puzzle together where the final outcome is predetermined, one piece is added after another, and occasionally a wrong or wrongly set piece is removed, or re-set. Instead, we assume that there is a set of principles which apply in all learner varieties, which operate on a given repertoire of linguistic devices, and whose interplay leads to the specific utterance structure. The interplay and the relative weight of these principles may vary, depending on factors like the source language of the learner.² In particular, it varies over time, though in a systematic way. This "vertical systematicity" - in contrast to the "horizontal systematicity" of each learner variety in itself - constitutes the acquisition process, to the extent to which it is scientifically interesting and not just random. Adding a new feature derived from the input, say case-marking, leads to a shift in the balance of the various interacting factors, until, possibly, that interplay of interacting principles is reached which is constitutive of the target language. The questions to be answered here are: (i) What are these principles, how do they interact, how does this interaction develop over time? (ii) Why does a learner variety develop in exactly the way it does? The present study cannot answer either of these questions, but it can contribute to an answer, and hence to a better understanding not only of the acquisition process but of the structure and functioning of linguistic systems in general.

The way in which this is done is an essentially inductive but hypothesis-guided investigation of how learners of different origins put their words together at a given time in a given and at least partially controlled situation. We do not think that this is necessarily the best way to study the development of utterance structure. It is simply a result of the present state of second language acquisition (SLA) research. At present, we do not have testable theories of SLA - that is, theories which are general and consistent, where basic concepts are precisely defined, which make clear predictions about the facts, and for which there are clear criteria of verification or falsification.³ The approach taken here is much less ambitious; but it has the advantage of being controllable and easily applicable.

In the next section, we will explain what we have in mind. Section 2.3 discusses some practical problems of the procedure, and section 2.4 is a very detailed sample analysis. In section 2.5, we present some results of a pilot study; it provides us with some hypotheses which will guide us through the empirical analyses of chapters 3-6.

2.2 The learner's problem of arranging words

At any point in time, an adult learner (as any other speaker) can draw on different kinds of cognitive prerequisites whenever he wants to communicate. To mention the most important ones:

- (1) He already masters a language, that is, he disposes both of the specific expressive devices which constitute that language and the semantic and cognitive categories that underlie them, such as modality, deixis, agency, or whatever.
- (2) He knows about language and communication in general, that is, he knows that he has to control for communicative success, that speaker and hearer have differing background knowledge, that there are social conventions of who is allowed to talk to whom, etc.; he possibly does not know how this is concretely organised in the language he has to learn and in the culture which he joins (or invades), but he knows that there are such rules and constraints.
- (3) He has a lot of non-linguistic information about the world and, thanks to his eyes and ears, about the situation in which he puts forth his utterance.
- (4) He finally knows bits and pieces of the target language.

Obviously, these prerequisites may be very different from learner to learner, and they permanently change for one learner over time. (Actually, this is the point of learning.) Suppose now that you are an adult Italian worker who has been living and working in England for about 6 months; we will take this characterisation as a sufficiently clear specification of (1-3) in the present context. Concerning (4), you know, among other things which we will not specify,

- a number of proper names, such as *John, Peter, Mary*,
- some base forms of verbs, such as *love, come, kill*;
- that the former words denote human entities and the latter words denote actions that these entities can accomplish.

You have no inflexion, hence no case morphology or agreement (that is, your learner variety is quite restricted).

Suppose now that on some occasion and for some reason, you want to express the thought that Peter is in love with Mary. Now, you are able to denote some components of this thought, Mary by *Mary*, Peter by *Peter*, and that relation at least approximately by *love*. But this does not suffice: somehow, the words must be arranged. This follows from the fact that language is linear (there are no hierarchical complications here). But it must also be done in such a way that the listener is able to derive the intended thought from what is uttered.

There are two extreme possibilities with respect to the six theoretically possible arrangements:

Either: There are no restrictions at all in your learner variety.

In this case, there are two apparent consequences:

- a. It is unlikely that you will get your message across since your utterance is in many ways ambiguous.
- b. The analyst has nothing to say with respect to the organisation of your utterances, he must just wait for a more advanced learner variety.

Hence, we can forget this case.

Or: There are restrictions of some sort. The obvious question then is: in what terms can they be stated?

Let us give some examples (we only consider the first position of the utterance):

- (1) a. *The shortest unit* (in terms of phonemes) *is first*. Most people would not seriously consider this possibility, although incidentally, it is close to being true in discourse.⁴ But we take this to be a consequence of other facts.
- b. *The verb comes first*. Note that this makes sense only if you, the learner, indeed distinguish verbs and other word classes in your variety and that the analyst can perceive the distinction. This is by no means trivial; we might ask what a "verb" is, if there is no tense, no inflexion, no agreement, etc. Thus, one might argue that the distinction between *love* (verb) and *love* (noun) totally collapses in your variety. (For a longer argument on this point, see Klein 1986, section 5.2.)
- c. *That NP which is morphologically unmarked ("nominative") comes first*. This is actually impossible in your variety, as it was defined above, since you have no morphological marking at this time. So, if we define "subject" as a syntactic category on morphological grounds, then there is no point in speaking of subject in your variety at all.
- (2) a. *If the thought to be expressed corresponds to an action or activity, then that entity which performs the action ("agent") is named first*. If there is no action or activity, the choice is free.
- b. *Animate entities come first*; if there are more than one of them, the choice is free, unless one of them is human; then this one comes first.⁵
- (3) a. *An entity which was referred to before comes first*; if there are more than one of them, the choice is free.
- b. *That entity which is dearest to the speaker's heart comes first*.
- c. *That entity which the speaker thinks to be best known to the listener comes first*.
- (4) There is a combination of some of these possibilities, such as for example: If the thought expresses an action with an animate and an inanimate participant, then the animate participant comes first unless the inanimate participant was mentioned before; then this one comes first. If there is completion of the action, the verb comes last, unless it was mentioned before; then ... etc.

All of these possible principles are invented, but they are not totally unrealistic. Note that they satisfy more or less the necessity of getting the message across, depending on the degree to which the listener shares the principle in question. It is easy to see that this is probably quite different in all of these examples. We think that most fully developed languages actually represent a quite complex case of (4), that is, the organisation of utterances is not based on some straightforward principles such as for example (1b) or (3a), but on a highly complex "coalition" (Bates and MacWhinney 1987) of different constraints.

Examples (1-3) are just some of the possibilities, but they are not totally arbitrary, of course: they represent what one could call "syntactic" principles, "semantic" principles, and "pragmatic" principles, respectively. One should not attach too much weight to this terminology, but we will use it occasionally as a shorthand. It should be noted, however, that "syntactic" constraints such as (1b) or (1c) in a sense postpone the very problem. One might ask why, for example, certain elements of a total thought to be expressed are encoded as a verb rather than as a noun, why a certain entity gets morphological case marking, or why an entity is expressed by *a* N rather than *the* N or just N. To answer these questions, we must resort to statements of type (2) or (3); for example: Something is expressed by *a* N rather than by *the* N because it was not mentioned before, etc.

2.3 Transfer and universal principles

In (1-3) above we simply listed some possible principles which a learner with the elementary repertoire described above could follow. Where could these principles come from? A number of sources are possible; we will mention and briefly comment upon four of them:

A. *The constraint holds in the learner's first language.*

This is probably the most common explanation, and there is no doubt that transfer (of this sort) in part influences the structure of learner varieties. There are however some problems with this explanation. Compare, for example, the following two cases, where the target language has constraint (3a), that is, the first position is always taken by a NP which was mentioned before, if there is a choice; otherwise, there are no restrictions:

- (1) The source language has (2a), i.e. "agent first". This the learner can easily transfer, and it is accidental whether the resulting utterance corresponds

to the target language rule.

- (2) The source language has (1c), i.e. "nominative first". In this case, transfer is simply impossible unless the learner has *what he thinks to be* the appropriate L2 morphology. In this case, the result with respect to correct L2 structures is accidental again, and it also shows that transfer from L1 structures into learner varieties often requires considerable knowledge of L2 structures (here, knowledge of TL case marking is a prerequisite to transfer of case marking): *The more you know about LS, the more detail you are able to transfer from L1*. It may also be that the learner who still lacks L2 morphology indeed happens to follow (3a) right from the beginning, following the general maxim "from known to unknown" (cf. "Behaghel's law" in Behaghel 1923-32); only after he has the appropriate morphology does he apply his L1 rule. In other words: Transfer resembles many other domains in that the more you know, the more kinds of mistakes you are able to make.

The second possible answer is:

- B. *A constraint is used in a learner variety, because the learner assumes that it holds in the target variety.*

This sounds almost trivial. Why after all should a learner use a rule if he does not think it holds in the language to be learned? But in real time communication, the learner is often forced to apply rules he is totally uncertain about or which he even thinks to be false with respect to the TL. This situation is familiar to any speaker of a second language; but it also has been reported for child language (for a discussion, see Klein 1986, chapter 8).

- C. *The constraint is based on a universal.*

Claims to this effect have often been advanced in the literature (amongst many others: Gass 1984; Rutherford 1984); they are sometimes seen as an alternative to transfer hypotheses (although the choice is not mutually exclusive). There are various types of universals that might be operative here:

- (1) Constraints in the sense of Chomskyan "universal grammar", such as "sub-jacency", "specified subject condition", etc. (See Chomsky 1981; Hornstein and Lightfoot 1981; Felix 1984; Flynn 1987; White 1990). If there are indeed constraints of this sort, then they do not say very much about utterance organisation in learner varieties that are as elementary as those discussed

here. It is hard to see how a constraint such as "subjacency" or even universal phrase structure could restrict the possible word orders in the "Peter is in love with Mary"-example...⁶ For a more systematic discussion of this point, see Klein (1986).

- (2) "Statistical" universals, in the sense of the Keenan-Comrie-hierarchy of noun phrase accessibility (Keenan and Comrie 1975; Comrie 1981; Hawkins 1983; Gass 1984). Universals of this kind may well serve as heuristic guidelines: they give the researcher an idea of where to look for interesting phenomena. But they face some of the problems of UG (e.g. not all learner varieties actually have relative clauses) and all of the problems which we briefly discussed in section 2.2: So long as we don't have criteria for what "subject" is in learner variety utterances (and elsewhere), universals of this kind are of little help (Perdue 1984a).
- (3) "Pragmatic" universals, such as "From known to unknown" (Behaghel 1923-32), "Me first" (Cooper and Ross 1975; see also Silverstein 1976), etc. We think that universals of this sort indeed play an important role in the organisation of learner varieties. They are not particularly clear, however, and as has been said in section 2.2, we think it might be more practicable to describe first what is indeed operative in learner varieties in these terms and then look for possible extensions and generalisations of the results, rather than stating a universal and then applying it to learner varieties. Thus the study of learner varieties may help in providing a better description of these universals.

Note that for both (1) and (2), the same objections hold that have been made in connection with transfer and its possibility. To be operative, both kinds of universals need considerable knowledge of the language to be learned. This is much less the case for (3), as the constraints here are less tied to the specifics of a given language (hence their importance in learner varieties) and as such can be assumed to be broadly shared by competent communicators. These constraints therefore tie in well with a fourth, and rather different answer to the question why a learner might prefer a specific constraint over another:

- D. *The learner may assume that following this constraint makes his utterance better understandable (or understandable at all).*

This possible and plausible reason has a slightly different status from A-C, in that it projects these latter onto the assumed state of knowledge of the listener: it

influences production only indirectly. Thus, D might be operative just by causing the learner to choose B or C - or A, but only if source and target language happen to be identical - or close - in some respect.

We think that in reality, all of these factors influence the way in which the learner organises his utterances, and the way in which they interact with each other can only be determined by comparing learner varieties under varying conditions: with different target languages, different source languages, and at different developmental stages. The present study is a step in this direction.

In the next section, we will present a detailed sample analysis of the procedure. Essentially, it is an utterance-by-utterance analysis of a connected piece of text. This text is one learner's attempt at the Charlie Chaplin retelling described in chapter 1.3 above. This procedure defines a controllable background - that is, we have at least a partial control of the speaker's and the listener's shared knowledge at the outset. We also know what the speaker has to talk about, that is, what he (maximally) *wants* to tell. Obviously, this control is not perfect, but is fairly strong. A full text (rather than individual sentences) was chosen because it allows us to control for introduction, maintenance and shift of referents under varying circumstances. We also have a reasonable control over the temporal and spatial organisation of the whole story.

Thus it appears that this kind of data gives us good, although by no means ideal material for verifying a whole series of possible constraints along the lines indicated in 1-4 of 2.2. There are three methodological problems, however, which deserve mention:

- (1) The data analysed here are limited both in size and type. It may well be that the learner, under different circumstances, organises his utterances in a different (or partly different) way. This can only be controlled by including other types of data.
- (2) It is often very problematic to interpret utterances in learner varieties. There is little reliability in learner grammaticality intuitions, nor (and this is far more problematic) in their reports of what a certain utterance that was recorded some time ago really means. As a consequence, most samples of learner varieties contain certain utterances that are wildly ambiguous or simply uninterpretable.
- (3) Finally, we are faced in a particularly blatant way with all the practical problems of analysing recorded spoken language. Learner variety samples show many false starts, hesitations, self corrections, etc. This is no problem in principle (and indeed may provide us with helpful additional insights),

but in practice, it is a challenge for any reliable analysis. We will be confronted with numerous problems of this sort in the following sample.

So much for the general background of the present study. In the following section, we will make a first step from the fog of these theoretical considerations into the swamp of empirical facts. Those readers who do not wish to get their shoes muddy may skip the first part of this section and jump to the conclusion on page 35sq.

2.4 A sample analysis: Vito

Vito is an Italian learner of German. In what follows we shall analyse the first episode of his retelling. This passage is only about the seventh part of his entire retelling but it suffices to illustrate virtually all typical features of his learner variety. It also exhibits most of the problems an analysis of a learner's utterance is faced with: ambiguous verbs or structures, uninterpretable sequences, inconsistencies, etc.

2.4.1 *The informant*

Vito was born in 1948 near Palermo (Sicily). He went to Germany in 1981. The data on which the present analysis is based were recorded about one year and a half after his arrival. At this time his command of German was still highly limited. This is due to the fact that he does not have very much contact with the German population: he works in a kitchen of an Italian restaurant, is married to an Italian woman and he has no children. On the other hand, he is talkative, self-confident, lively and very interested in questions of language. As a consequence, his metalinguistic behaviour is quite elaborate. While retelling the film, he often interrupts himself and asks for a word or expression, mostly with a formulaic question: *Was ist der Name?* "What is the name?", *Was Name diese?* "What name this?". Occasionally, he checks whether his own expression is correct: *Richtig spreche?* "Correctly speaking?". There are other, less apparent but more interesting traces of his metalinguistic awareness. Very often, his production gives the impression of being carefully planned, with clear prosodic phrasing of each word. It is also interesting to note that quoted speech, another "metalinguistic" device which he often uses (as other informants do) appears to be closer to the German standard than utterances in which he reports

events or provides background information. Thus, three types of utterances should be kept apart in analysis:

- a. Purely metalinguistic speech which largely consists of "rote forms" such as *Was ist der Name?*. Their inclusion would lead to a distorted picture of his own utterance organisation.
- b. Direct narration, including both events and background information.
- c. Quoted speech which is not totally different but a bit more advanced towards the German standard.⁷

The difference is best illustrated by the use of the copula. It never occurs in type b. utterances, that is, in his "genuine" production. It occurs sometimes in type c., and it is rather frequent in type a., due to its presence in rote forms.

2.4.2 *Vito's linguistic repertoire*

Morphology

Vito has no inflectional morphology, hence no case marking, no agreement, no tense.

*Lexicon*⁸

It is obviously impossible to estimate the exact size of Vito's lexicon on the base of the text studied here, since his active use is clearly determined by the nature of the task. Still, it gives us some idea of how his vocabulary is composed. For some of the following forms, there are a number of phonological variants; we give only one. The word class assignment given here is highly problematic and should be treated with caution especially with respect to closed class forms.

Nouns: He uses 60 different nouns, the most frequent ones being *Mädchen* (referring to the girl, 27 occurrences within 1050 words of running text), *Polizei* (or *Policia-Mann*) "policeman" (22x) *Frau* "woman" (17x), *Gefängnis* "prison" (15x), *Auto* "police car" (13x), *Holz* "wood" (13x), *Brot* "bread" (12x), *Schiff* "ship" (10x). So, clearly main protagonists and main features of the film show up most often. Incidentally, the name Charlie Chaplin occurs 24 times; Chaplin is by far the most important referent, he is mostly denoted by a personal pronoun, *sie*.

Verbs: He uses about 40 different verbs, some of them with a rather overgeneralised meaning (nouns, in contrast, are rarely over-

generalised, as far as the data allow any conclusions here). The most interesting forms are:

- a. *gucke* is the most frequent verb and means "to perceive", "to realise", "to look for" or "to look at", even "to imagine".
- b. *spreche* "speak" (18x), *rufe* "call" (2x) basically introduce quoted speech; *spreche* may be used with an addressee (*sie spreche diese* "he said to this one") or without ("he said"); it is interesting to note that he never uses *sage* "say", the most common *verbum dicendi* in other German learner varieties.
- c. *komme* "come" (23x), *geht*⁹, *nehme* "take" (1x), *bringe* "bring" (1x). He totally misses the deictic component of these verbs;¹⁰
- d. *mache* "make" , *höre* "hear", *lasse* "let", *brauche* "need", *rauche* "smoke" and some others roughly correspond to standard usage.
- e. *habe* "have" is used only as a full verb, never as an auxiliary.
- f. there are two modals, *muß* "must" and *wolle* "will".
- g. there is one perfect participle form, *gefunden* "found", but no corresponding present form (*finde*); hence, there is no reason to assume that it is used to mark past or perfectivity in opposition to present; it is interesting, though, that he chooses a verb with inherent perfective meaning on first introducing a perfect form.
- h. He uses a number of compound verbs, such as *rausgucke* "look-outside", *aussteige* "get-out", *wegmache* "get-rid- of"; sometimes, the separable particle is used alone, in varying positions.
- i. There are no auxiliaries (*sein*, *haben*, *werden*), except in rote forms and sometimes in quoted speech (see 2.4.1). Hence, there is no passive, although one occurrence could be interpreted as a past passive.

Negation: *nix* (16x) and *keine* (7x) are used interchangeably as sentence negation; sometimes they are used together.

- Adjectives:** He uses about a dozen adjectives, both in attributive and in predicative function; in the former, they may be before or after the noun.
- Adverbs:** (a) spatial: *weg* "off/away" (20x) and *zurück* "back" are rather frequent; basically, they are a kind of verb remnant (*weggehen*, *wegnehmen*, *zurückgehen*, etc.); there is one occurrence of *hinten* "behind"; but the most striking fact is the lack of deictic local adverbs;
- (b) temporal: there are only four of them, *sofort* "at once" (8x), *dann* "then" (8x), *später* "later" (2x) and *immer* "always" (1x); again, the lack or rare use of deictic forms is rather striking; most narratives in learner varieties (and elsewhere) are structured by "and then" and related connectives, which Vito almost never uses.
- (c) others: among the five or six other adverbs, two are particularly interesting: *vielleicht* (9x) means something like "something like" or "approximately"; *zusammen* (10x) often replaces the personal pronoun "they", e.g. *zusammen spreche* ("they talked to each other").
- Determiners:** He regularly uses three determiners: *diese* (50x), which marks definiteness, *de* (36x), with many phonological variants, which marks definiteness, too - the difference will be discussed below - and *eine* (32x) for indefiniteness.
- Quantifiers:** They are very rare: *viel* "many, much", (5x), *all* "all" (3x), *zwei* "two" (1x).
- Pronouns:** *sie* (28x) means "pronoun third person" - it mostly refers to Charlie Chaplin; *ich* "I" (13x), *du* "you" (3x), *mir* "to-me" (2x) - only with prepositions - and *seine* "his/her" (9x).
- Conjunctions:** *und* "and" (13x), *oder* "or" (3x), *aber* "but" (1x).
- Prepositions:** There is only one frequent preposition, *in* (30x), which, just as the six or seven other ones he occasionally uses is strongly overgeneralised to denote all sorts of spatial relations.

There are a number of other words, which are rare, however, and hard to classify. He very rarely uses Italian words, with one exception: **alora** is used about 10 times to mark a restart.

So much about his words. How does he put them together, given that he has no case, no agreement, no case-indicating prepositions? In the next section, we will illustrate the way in which he proceeds by a closer look at a sequence of about 30 utterances.

2.4.3 *The shipyard episode*

The following sequence of utterances is somewhat "edited", that is, we have omitted

- obvious false starts and breakdowns
- hesitations, interjections, etc.
- metalinguistic comments.

In addition, phonological variants are "standardised" to one (the most frequent) form. Obviously, this "falsifies" the original transcript; but otherwise, a sensible analysis is almost impossible. (All utterances are numbered; + denotes a short pause, xxx a short, acoustically unclear passage, ...M... an omitted metalinguistic passage. For complete transcription conventions, see the Glossary.)

Some observations

The informant had been asked to start with the scene where Chaplin left the prison with a letter of recommendation. Both Chaplin and this letter had been mentioned in the immediately preceding utterance of one of the interviewers.

(1) *sie habe brief + brief für gefängnis*

'she have letter + letter for prison'

The intended meaning is quite clear. Chaplin has/had¹¹ a/the letter - the letter from prison. If we ignore the attributive complement *brief für Gefängnis* for a moment, the utterance structure is NP₁-V-NP₂. We will consider these three components in detail:

1. NP₁ refers to Chaplin. The form *sie* corresponds to a pronoun of the target language; its appropriate form there would be *er*; it is unclear whether *sie* is derived from the corresponding feminine or from the corresponding plural form - both are *sie* in standard German. It sounds peculiar to start a story with a non-deictic personal pronoun; but in this case, its use is simply explained by the fact that Chaplin was mentioned immediately before by the interviewer; hence the informant uses anaphoric elements in first position, at least if the referent was thematic immediately before.¹²
2. V denotes a stative relation; hence, it makes no sense to call NP₁ an "agent".

3. NP₂ refers to an entity which in Standard German could be denoted either by a definite or an indefinite NP (*den Brief- einen Brief*), depending on whether the speaker shapes his expression on what has been said before or not. Since he has done so in case of NP₁, we would have to assume that this language behaviour is inconsistent when assuming that NP₂ is indefinite; hence we should assume that "bare" NPs may - but need not - be definite.

At this point, many would be inclined to say that *sie* is the subject and *brief* the object of this utterance; but the only argument at this point is *that this is so in the corresponding target variety*, hence we would automatically assume that the same regularities that characterise standard German also obtain in this learner variety. But any such assumption is at best a useful heuristic, given that those characteristics that mark "subject" and "object" in German - such as case, agreement, position - either do not apply in this learner variety (case marking) or are dubious (agreement, position).

NP₂ is actually more complicated, since it also contains, or is related to, the complement *brief für gefängnis*. It is hard to see, at this point, whether this construction is simply a kind of "postscript" or rather a disguised relative clause "...letter - the letter from..." as opposed to "...the letter (which was) from ...". In any case, it is interesting to note that *gefängnis* again is definite and has no determiner.¹³

- (2) *komme in eine baustell*
'come in a building site'

The intended meaning is clear again: Chaplin comes to a building site. The structure is V-PP, where PP is directional. We will comment on the "missing" agent, on V and on PP.

1. There is an agent here, since *komme* denotes an action, and it is clearly Charlie who does it. It is not made explicit, however. The agent is the same individual that was referred to in initial position in the immediately preceding utterance. Note that this does not allow us to state that (2) has a "zero anaphor" in *first* position. What we have so far, are two hypothetical conditions for leaving a referent unexpressed:

- a. It was thematic¹⁴ immediately before;
- b. It was in initial position before.

We shall return to these conditions below. It is interesting, though, that an analysis according to which there is a "subject", realised by zero, in first position, comes so naturally. But any such assumption about a learner

variety includes various uncontrolled presuppositions and is possibly more the result of a language-specific interpretative bias than of factual evidence. Italian, the source language in this case, easily allows the "subject" in non-initial position; hence, there is no sufficient justification to assume that the learner, were he to borrow from his first language here, indeed places his empty subject in first position. German, on the other hand, has a certain preference to have the subject first if there is no adverbial that could take that position. But it would be mere speculation to say that the learner already "has" the German strategy at this point.¹⁵

2. The verb clearly denotes an action; *komme* seems to have the standard German meaning here.
3. The PP corresponds to the target language pattern, except that it is not case-marked. Note that the referent of the NP *eine baustell* was not thematic before. It is marked, as should be the case, by an indefinite article.

(3) *baustell* *vielleicht*
 'building site perhaps'

The intended meaning is something like "a sort of building site" (actually, it is a shipyard). Prosody clearly marks this utterance as being separated from the preceding one; it cannot be a syntactically integrated part of a complex NP *in eine baustell + baustell vielleicht*; rather it is a kind of postscript (cf. 1 above). Note that the contextually given part of the construction, *baustell*, comes first here; it is not marked by any article.

(4) *diese mache schiff*
 this make ship

There are two possible readings here, depending on whether *mache* is given a specific or a generic reading: "this one was building a ship", or "this one was one of those that builds ships". Both interpretations are justifiable in this context. Given the whole plot the speaker could mention that "they" are building a particular ship; but he might also elaborate on what he had said before: it was a sort of building site, one of these places where they build ships. There is no way to decide which of these two interpretations is more appropriate.

The structure is clearly NP₁-V-NP₂. We will briefly comment on all three positions.

1. the NP₁ *diese* refers to an entity, the shipyard, that was already thematic in the immediately preceding utterance, but in a different position, and with a different function. Hence, we might say that *diese* goes with "referent

maintenance" but "position shift" and "role shift", leaving aside the question of what "role" precisely means in this learner variety. Note that in this respect, *diese* corresponds to at least one of the uses of the demonstrative pronoun *diese* in standard German.

2. *mache* clearly denotes an action, hence *diese* is an agent, although one would not normally consider a shipyard to be an agent; hence, standard semantic processes that allow to go from *the shipyard* to *the people at the shipyard* also apply to *diese* in this learner variety.
3. *schiff* introduces a new referent; it is unclear whether it is specific or generic, likewise whether it is singular or plural; according to the two interpretations mentioned above, the Standard German counterparts would be "...baute ein Schiff" and "...baute Schiffe", respectively. So, we again must leave open whether there is a special function to the "zero article".

(5) *kleine schiff mache*
'small ship make'

Just as in the preceding utterance there is a specific and a generic interpretation, roughly corresponding to "The shipyard was building just a small ship" and "It was a shipyard for small ships". Clearly, the latter interpretation is much less natural here; one small ship is shown in the film, and there is no reason to state that it is a shipyard just for small ships. Hence, we will assume the first interpretation.

The structure is clear (NP-V) but perplexing. The verb is in final position, the NP refers to an object; the agent is still the shipyard, but it is not referred to explicitly. Note that both conditions from (2) above for having a referent unexpressed are satisfied. Again, there is no clear evidence for putting a "zero subject" in a specific position, for example before *Weine schiff*. Note that any such assumption would have immediate consequences for the positioning of objects and the relation of the learner variety to the target variety. In Standard German, an object may precede the (finite) verb in simple sentences, but the subject must then follow the verb: the verb must not be preceded by two major constituents. But it is not clear whether Vito's learner variety also disallows this, especially as his variety does not have tensed verbs; having a zero anaphor in initial position may well be compatible with the structure of his variety.

It is hard to see why he prefers this order rather than having "mache kleine schiff". It might be random and hence uninteresting (see 2.2). If it is not random, several possibilities come to mind:

1. There might be a structural principle roughly saying that, if there is only one NP, that NP comes before the verb, irrespective of its possible function. This is clearly falsified by (2) and by numerous other examples, as we shall see.
2. It may indicate a special function of the constituents involved. Now, the NP *schiff* in (4) seems to be in a similar sense "direct object" of *mache* as *kleine schiff* in (5). Hence, the function "object-hood", is not associated with the choice of position. A more plausible reason for the different arrangement in (4) and (5) might be a difference in the "given-new-distribution". Now, the verb *mache* is clearly "given" in (5) and "new" in (4), and the only new part in (5) is *kleine*, i.e. a part of the object NP. Hence, the arrangement in (5) is clearly at variance with standard assumptions about "given-new-order": it goes from "new" to "given". Interestingly enough, in Standard German a sequence such as *Sie baute ein Schiff. Ein kleines Schiff baute sie.* sounds at least as natural as *Sie baute ein Schiff. Sie baute ein kleines Schiff.*¹⁶ if the second sentence is meant as a specification, rather than a correction, of the first sentence. Standard assumptions about given-new-order are perhaps too gross.

To sum up: (5) clearly contradicts two straightforward views about utterance structure in learner varieties:

- a. Grammatical functions, such as being a direct object, are clearly related to positions.
- b. Utterances proceed internally from "given" to "new".

We will see below that there is an interpretation of (NP-V), which fits a number of other utterances.

- (6) *chef arbeiter rufe "charlie chaplin"*
 'boss-worker call "Charlie Chaplin"'

The intended meaning is obvious: the foreman called: "Charlie Chaplin". It is clear from the intonation that *chef arbeiter* is one constituent (Vito uses a similar strategy elsewhere: a station-master is *kaiser bahnhof* in his variety), and that the second NP *Charlie Chaplin* is a vocative, rather than an object.

The structure is quite clear. The NP introduces a new protagonist; this is done by a lexical NP without an article. Clearly, this NP is definite, although the referent has not been explicitly introduced; what is meant is "the foreman of the shipyard", the latter not explicitly referred to again. The verb denotes an action, hence the NP denotes an agent.

- (7) *ich brauche eine holz*
 'I need a wood (log)'

The intended meaning is clear; the whole construction, as often in quoted speech, corresponds to the target language pattern. The structure is again NP₁-V-NP₂.

1. NP₁ is realised by a deictic pronoun, denoting the speaker introduced in the preceding utterance.
2. The verb is clearly not agentive, hence *ich* is no agent.
3. The object NP is *indefinite*, its referent being freshly introduced.

- (8) *ich brauche eine ...M... keil*¹⁷
 'I need a wedge'

The preceding pattern is exactly repeated.

- (9) *sie nix verstehn*
 'she no understand'

The meaning of (9) is not fully transparent. It could mean "He was not understood"; with *he* referring to the foreman; this is somewhat unlikely since the informant never uses passive elsewhere. The other and more plausible interpretation is "He did not understand", with *he* referring to Charlie. But this forces us to assume that a personal pronoun, in this variety, may jump over an appropriate referent, here *chefarbeiter* and take up another referent introduced some utterances before.

- (10) *nix komme eine keil +*
 'no come a wedge'

- (11) *eine holz + lang + zu lange*
 'a log + long + too long'

Again, several interpretations are possible; (10-11) could correspond to Standard German *er bekam nicht einen Keil*, (*er bekam*) *ein Holz*... "he did not get a wedge (he got) a log", or *es kam nicht ein Keil* (*es kam*) *ein Holz* "it wasn't a wedge that came (it was) a log". There is no way to decide between these alternatives. Note, however, that in the second case, a "subject" would appear in final position; this is not unlikely with verbs that express something like "appearance on a scene", such as *komme*, as we shall see.

The whole construction is interesting in that it actually consists of two adverbial components, roughly 'come not wedge - come wood', the latter NP being expanded by a post-posed attribute. It is interesting to note that the negation *nix* precedes the whole first clause although it only applies to the NP *eine Keil*. Note that both NPs are marked by an indefinite article although at least the first one was mentioned before; but this would be possible in the target language, too (in this case, both definite and indefinite NP would be appropriate).

- (12) und *sie spreche*
'and she speak'

The context makes clear that *she* refers to the foreman. Hence, if we assume that *sie* in (9) refers to Chaplin, we would again be faced with an anaphoric jump. The verb again introduces quoted speech:

- (13) *diese nix*
'this no (nothing)'

The meaning is clear, the structure too. The initial NP *diese* refers to the log which was thematic before, although in a different function and position. The case resembles (4), although there is a difference in that *diese* belongs to a stretch of quoted speech; thus, we will treat this use as deictic.

Apparently, the whole utterance is highly elliptic; it could mean "This one is not good" or "I don't want this one" (to mention two possibilities), but any such interpretation would be arbitrary. Note that in Standard German, the corresponding elliptic utterance sounds quite natural, although the reversed order (*nicht diese*) is also possible, if not better.

- (14) *ich wolle eine ...M... keil*
'I want a wedge'

The meaning is clear, the structure corresponds to the target language.

- (15) *sie gucke eine keil*
'you/she look a wedge'

The meaning is either "He (=Chaplin) looks for a wedge". Here, there cannot be any doubt that *sie* jumps over an appropriate referent and takes up an item introduced some utterances before. The object is again indefinite, as it should be in the target language. The verb denotes an action, hence the "subject" is an agent. Or, we have quoted speech: "You look for a wedge". However, this interpretation necessitates analysing *Sie* as the polite address form, which Vito never uses elsewhere.

- (16) *hinten + eine große holz*
'behind (adv) a big log'

The meaning is quite clear "Behind/further back, there was a big log". The structure corresponds to the target language, except that there is no copula or copula-like verb: *Hinten war/befand sich/lag ein großes Holz*. But it clearly differs from almost everything we have had so far. The initial position is taken by an adverbial, the only NP - which would be considered as subject in the target language - is in second position. This NP is indefinite, and it introduces a new referent. Hence, we may conclude that at least in "copula"-utterances, NPs in subject-like function need not be bound to the first position, if they are

indefinite.

- (17) *komme diese nix weg*
'come this no off/away'

The meaning is clear, at least to some extent; he does not get this log (to move) away (in the film, Charlie tries to remove a wedge-like log which supports the ship). But it is unclear whether *komme* denotes an action which would require an (implicit) agent, or whether no such agent is implied, the whole utterance meaning "this one did not come away". The latter analysis corresponds to the preceding uses of *diese* referring to an entity which has already been introduced, but in a different function. Here, it functions as a kind of "subject", which is not an agent, of course. Its position corresponds to the "subject" of *komme* in (10) above.

- (18) *seine hand nix habe keine kraft weg diese*
'his hand no have no strength off/away this'

This is by far the most complex utterance in the whole episode. It means "his hand was not strong enough to remove it". It consists of two "clauses": NP neg V neg NP, and Particle NP, where the particle represents a full verb - at least semantically.

The first NP, although clearly not agentive, does correspond to a target language subject and it is definite. It is difficult to say whether the second NP should be considered as definite or as indefinite; in Standard German, *keine* is a combination of *neg* and *eine*, and it would be appropriate to say "hatte keine Kraft, um..."; but Vito uses *keine* often as a regular negation which corresponds to *not*; hence, it is unclear whether he indeed marks *kraft* as indefinite.

In the "complementiser clause", the order of *diese* and *weg* from (17) is reversed. The referent of *diese* was thematic before but, if our analysis above is correct, in a different function; here, it is a kind of object to *weg* which functions as an agentive verb. This could account for the reversed order. But it does not, as we shall see in a moment.

- (19) *sie gucke eine ... M ...*
'she look a' M

Vito does not know the word for "hammer", and having asked for it, he goes on:¹⁸

- (20) *gucke eine hammer*
'look a hammer'

Both meaning and structure are quite clear. Note that *sie* again jumps over several utterances which however, in this case, do not contain appropriate referents that would compete with Chaplin. In (20), no protagonist is mentioned, but it is clear that it is the same agent.

(21) *probiere diese weg*
 'try this off/away'

Again, the agent is maintained and left implicit. The second NP refers back to the big log introduced in (16), rather than to the hammer. This shows that *diese*, too, may jump over appropriate referents, and that it does not necessarily require a change of function: when last mentioned in (18), its function was apparently the same. What differs, however, is its position: it precedes *weg* which apparently represents a verb, again. There is no reason to assume that the grammatical relations between *weg* and *diese* differ in both cases, the varying order is either random or it reflects some other principle. But just as in the case of (4) and (5) above, it is hard to conceive of any clear difference in terms of topic-hood, givenness, focality and similar notions. We have no real explanation at this point.

(22) *sie nā nix gucke*
 'she no(?) no look'

The second word is difficult to understand, and hence, the whole utterance is hard to interpret. But the global structure is clear and does not add anything new to what we have seen before.

(23) *diese holz sicher schiff*
 'this log safe ship'

It is unclear whether *sicher* is a verb or an adjective. There is a German verb *sichern* "to make safe" which could be quite appropriate here; but it sounds very elaborate, and it surely is not a word one would expect in Vito's vocabulary, moreover, its form would then be *sichere*. It seems more plausible to assume that *sicher* corresponds to the common adjective *sicher*, and that a causative component such as *mache* is omitted (or rather not made explicit). In any case, *sicher* in (23) *Junctions* like a transitive verb in that it binds two arguments, one of them corresponding to a "subject" and one to a direct "object". We make this point here to highlight again the fact that analysing a constituent in a learner variety as a verb, and object etc. is very often based on uncertain decision and speculative analogies. The first NP *diese Holz*, although not animate, is a kind of agent here. It refers to an entity that was thematic in (18) and (21), but in a different role. Just as in the case of (21), *diese* crosses a previous utterance; but no ambiguity is possible, first because (22) does not contain an appropriate referent (as far as we can see), and second, because it is completed by a lexical item, *hob*. One wonders why he adds *holz* here, but not in (21) where it would indeed resolve an ambiguity.

The second NP, *schiff* could take up the ship introduced in (4) or (5) and

hence be definite; but as the next utterance will show, this is not so clear.

(24) *diese schiffarbeite neue schiffbau + neue schiff*
'this ship work(er) new ship build + new ship'

Both meaning and structure of (24) are unclear. Obviously, he wants to give some background information about "that" ship, roughly "this ship was being worked on, it was in the process of being built". The last part *neue schiff* is distinctly separated by intonation; it sounds like a repetition: a new ship was being built - well, a new ship. If we ignore this segment, (24) consists of two parallel clauses, each consisting of a NP and a verb (or a constituent which one is inclined to interpret as verb-like). What is peculiar about these two clauses is the relation of NP to V. In both cases, the verb denotes an action, and one would interpret NP as the "object" rather than the "subject" of this action. Hence, the construction is a kind of passive without the least morphological marking. It resembles the "ergative" use (Keyser and Roeper 1984) of actions in constructions such as

Der Laden schloß
'The shop closed'

Das Radio spielte
'The radio was playing'

These relate to normal transitive use, as in "Someone closed the shop", "Someone was playing the radio", on the one hand, and to passive use, as in "The shop was closed" and "The radio was being played", on the other. We would assume that the semantic relation between the verb and the non-agentive NP is the same in all three constructions, but that there are some differences on another level, differences which are hard to grasp.¹⁹

We will not try to pursue this point in detail here. But it seems not implausible to assume that Vito indeed has such an "ergative construction" of action verbs, rather than a passive. This assumption would also easily account for the mysterious word order in (3) above: he uses *mache* in the same way in which he uses *arbeite* or *baue* in (24) or in which "to close" or "to play" are used in target language examples above.

It should be clear that all of this is highly speculative; but it seems to yield a consistent picture.

(25) *zwei holz + diese zwei und all andere*
'two log + this two and all other'

Again, the global meaning is quite clear, although the construction does not look particularly transparent at first glance. What he (Vito) wants to express is something like: "There were two logs; these two were connected to everything

else". (In the film, the half-built ship is held by two logs, one of them the wedge which Charlie tries to remove.)

Apparently (25) consists of two clauses, the first one being just an "existential" NP, which represents something like *Da waren zwei Hölzer* "There were two logs". Normally, the function of such "presentational clauses" is just to introduce a referent, the focal part being obviously the expression for that referent. Vito restricts his utterance to that focal expression, in this case without associating a local adverbial with it, as in (16), as the referent is already spatially located.

The second clause has no overt verb. It consists of two NPs related to each other by "and". The first NP is again constructed with *diese* and just as in previous instances, it takes up a referent that was previously introduced, but now has a different function. One might argue whether *und* has a verb-like status here ("these two - connected to - everything else") or whether it is just a normal phrasal connector, and Vito is simply omitting an appropriate verb ("these two and everything else hang together"); but there is no strong argument for either analysis.

(26) *sie weg eine + diese schiff weg*
'she off/away one + this ship off/away'

There are two "clauses", again, which are prosodically interrelated. What Vito seems to express, is: "When he took away one (of the logs), the ship went away". The relation is not overtly marked, however; hence one might analyse (26) just as two independent clauses.

In both cases, *weg* functions as a kind of verb, but in two different ways. Following the lines suggested in connection with (24), we may argue that in the first clause, it is transitive, and in the second, it is ergative. Alternatively, one might say that there are two verbs *weg* in Vito's variety, roughly corresponding to *wegnehmen* "to take away" and *weggehen* "to go away". The first analysis seem simpler, since we need not assume a lexical ambiguity in Vito's repertoire but only a different use of one and the same unit, and this use is found elsewhere, too.

Note that the second NP in the first clause, *eine*, does not function as an indefinite article but as a numeral ("one of them"); the first NP *sie* again "jumps" some clauses, but takes up that protagonist that was last referred to. The NP in the second clause shows a familiar pattern.

(27) *diese schiffweg + sofort meer ...*
'this ship off/away + straightaway sea' ...

Once more, there are two clause-like constituents, the first one being identical to the last part of (26). (27) roughly means: "When it moved away, it went

straight into the sea". The second clause only consists of an adverbial and a bare noun; this noun represents a directional adverbial. There is no reference to the "actor" - the ship - and no overt verb of movement.

There is an alternative analysis - the second clause could just be a "presentational" with an omitted copula, corresponding to: *da war gleich das Meer* "just there, there was the sea". We will come back to this kind of analysis in a moment.

(28) *diese schiff neu*
'this ship new'

Meaning and structure are clear; the copula is omitted. Note that the referent is re-introduced explicitly.

(29) *nix keine fertig*
'no no ready'

The utterance is parallel, the "subject" of the previous copula-clause remains and hence is omitted. Note the double negation.

(30) *sofort komme wasser*
'immediately come water'

There are two possible analyses of (30). It could correspond to "The ship immediately went into the water". Under this interpretation, *wasser* is a directional adverbial, the ship is "subject" and not made explicit, since it already was "subject" in the preceding clauses, and *komme* is a verb of movement which differs from the Standard German *kommen* just by the lack of the normal deictic orientation.

Alternatively, (30) could be viewed as the counterpart of Standard German *Da kam sofort das Wasser* "there immediately came the water", i.e. "the water was nearby". This analysis is equally consistent with what we have observed so far; in particular, (30) would be the exact counterpart of (27) under the "presentational" analysis of the second clause; it also would correspond to (17) and to one of the possible analyses of (11). Again, there is no straightforward way to decide between these two competing analyses. Note also that the first analysis - with *wasser* as a directional adverb - would be consistent with the analysis of *meer* in (27).

(31) *kaputt*
'kaputt'

The general message of this utterance, which concludes the episode, is clear. What is not clear, however, is the "topic" of *kaputt*. He could mean "The ship was 'kaputt'", but also "Everything was 'kaputt'", i.e. "it was just a disaster". The first interpretation only seems plausible if we analyse (30) in the first way, i.e.

with "the ship" being maintained as the "subject". We would then have a chain from (28-31) where each utterance has the same "subject", this subject being zero except for the first occurrence. But nothing excludes the other analysis.

This is probably the right point to stop these considerations and to sum up the major observations. We have been extremely detailed here in order to highlight the manifold problems which one faces in the analysis of learner varieties, that is, of utterances on a very elementary level of organisation, but also to get an idea of the principles behind Vito's way of putting words together.

Three conclusions

At least three lessons may be drawn from our previous considerations:

1. There are two major methodological problems. First, many utterances allow for several analyses. This concerns, among many others,

- the segmentation of longer stretches of words into smaller units (cf. utterances 1, 2, and 24);
- the lexical meaning of individual words (cf. the meaning of *komme* in 30 or of *weg* in 26);
- the option of positing zero elements at various places, such as zero subjects (cf. 2 and many others).

The second problem is a kind of methodological trap - namely the tendency to analyse the learner variety utterances as minimally deviant from target language utterances. Thus, we would be inclined to analyse (10) as "He did not understand" or (11) as "He did not come with a wedge" because the corresponding target language utterances look quite similar. Such an analysis may be correct, but it may also miss the internal properties of the learner variety itself. Example (16) shows that *komme* may have a special function - a function not totally alien to the target language but less familiar there.

The favourite prey of this trap are analogy inferences of the following sort: There is a learner utterance NP_1-V-NP_2 ; the target language has an apparently related construction $NP_1'-V'-NP_2'$, where NP_1' is considered to be the subject; hence, NP_1 is the subject of the learner utterance. Note that this latter analysis may indeed be appropriate for the learner variety; but it is not sufficiently justified by the alleged analogy to a target variety construction.

2. The way in which the utterance is put together is related to lexical properties of the verb, especially to whether the verb denotes an action, a static relation, etc. So far, we have encountered at least the following cases:

- a. There is a group of verbs which denote static relations, such as *habe* (1), (17), *brauche* (7), *wolle* (13), *sicher* (in 23, if it is indeed a verb). They are preceded and followed by a NP, where the first NP corresponds to a "subject" and the second to an "object". It goes without saying that all reservations concerning "subject" and "object" fully hold in this case: criteria such as "nominative", "agentive" do not apply to "subject", and characterisation by means of position is circular. Note that more indirect features such as "animacy" do not apply, either: the "object" may be animate, too, and the "subject" need not, as (23-25) show.
- b. Some other verbs denote an action, such as *machte* (4), *gucke* (14), (18), *bau* (24), possibly *weg* (16), (17), (26). They have either one or two overt NP actants.²⁰ In the latter case, there is a NP in initial position which denotes the agent, and a NP in final position which refers to the "object". Depending on the nature of the action; this object may be an effected, affected, or other entity. The agent may be missing. There are two structurally different cases. Either the agent is the same as in the immediately preceding utterance; then everything else remains as described above; the most straightforward analysis of this structure is to say that the agent is denoted by zero. Or the verb allows for ergative use; then the "object" - being the only NP - is in first position.
- c. There are also "inherently" intransitive action verbs, such as *komme* in (2) (as opposed to *komme* in 16). Here again, the agent is in first position - where it can be omitted if it is the same as in the preceding utterance; there might be a directional in final position; there is only one case in question among the utterances considered so far, but in fact there are many examples of this sort in Vito's complete story.
- d. *komme* may also denote something like "appearance on the scene", such as in (10) or in (30). It then has one NP actant that denotes what "appears"; this actant is in final position.
- e. Finally, there are a number of "copula" constructions, which are either equational or presentational: in both cases, the copula is mostly omitted. Equational constructions consist of two arguments; the first one denotes what is defined or characterised; it is always a NP; the second is either a property, as in (28), or a "predicative" NP (no example so far, but there are a number in the subsequent text). The first NP may be omitted if the same NP-referent was in first position in the preceding utterance. Presentationals consist of

a NP which denotes what is "presented"; this NP may, but need not be preceded by an adverbial (cf. 15, 25). In a sense, this construction should rather be seen as a variant of *komme* in the sense of d, the difference being that the verb is a copula and hence omitted.

Obviously, this is a very gross characterisation; but it makes clear what the minimal distinctions are.

3. The central constituents are verbs and noun phrases. Verbs systematically occur in one form. NPs, on the other hand, show a considerable variation in form. For present purposes, we may distinguish five types:

- a. lexical NPs which consist of a noun, which may be preceded by *eine* or by a numeral (in the subsequent text, we often find *de* N, where *de* is apparently a definite article);
- b. *diese*; it may be followed by a noun, too, so this construction overlaps with a;
- c. *sie*; it always refers to one of the protagonists and always has "subject"-function;
- d. zero (0); it is used if a referent in first position is omitted; unlike *sie*, however, (cf. 28 and 29), this referent need not be an agent. Note that the assumption of "zero anaphora" is already a (hypothetical) result of the analysis of the learner variety; it will allow us to state some acquisitional principles in a much easier way.
- e. Finally, there are some occurrences of the deictic pronoun *ich* in quoted speech.

Both *diese* and *ate* require that their referent be mentioned before, although not necessarily in the immediately preceding utterance (as in case d). The difference seems to be that *diese* is used if the function of the NP changes - most typically from "object" to "subject", such as in (15) and in (16), as in (22) and (23).

All that has been said here is based on a single episode which consists of 30 utterances. Taking these observations as a starting point we will now state some more general regularities based on an analysis of the whole text.

2.4.4 *Vito's basic patterns*

An inspection of the whole text shows that all of the patterns observed so far regularly occur, though with varying frequency, and that there is no really

different pattern; there are true problematic cases, however. We will first survey the main findings, and then "condense" them to some principles and finally consider some exceptions.

Main patterns

In the interests of clarity we will first ignore particles and adverbials. Utterances then consist of verbs (always base form), occasionally a copula (only in quoted speech) and NPs; NPs may be lexical, pronominal, demonstrative and - under very restricted conditions - zero elements (see the end of section 2.4.3). There are essentially four patterns:

A. $NP_1 - V - NP_2$

As has been noted at the end of 2.4.3, there are two possibilities here: the verb may denote an action or a static relation; there are some intermediate cases. This construction is relatively frequent (about 35% of the approximately 200 utterances of the retelling). Here are some examples:

sie habe brief...
'she has letter'

ich brauche eine holz
'I need a log'

sie gucke eine keil
'she look for a wedge'

\emptyset *nehme seine jacke und fez*
' \emptyset take his jacket and fez'

ein mädchen bringen eine brot
'a girl bring a bread'

die polizei gucke diese
'the police look for this'

diese polizei lasse charlie chaplin
'this police let Charlie Chaplin'

diese holz sicher schiff
'this log safe ship'

charlie chaplin gucke diese mädchen
'charlie chaplin see this girl'

The different kinds of NP are unequally distributed over both positions:

- a. *sie* and, by definition, zero are restricted to NP₁, i.e. there is no pronoun for NP₂.
- b. There is a clear preference to have those NPs in first position which denote a referent that was thematic before; there are indefinite NPs in first position, however, as well as "old" referents in last position.

If there is an agent, it is always in first position. In fact, this solves all arrangement problems for "action utterances". There is no such straightforward solution for "static relations", as expressed by *wolle, habe, brauch*. Almost all instances in the text could be accounted for by stating that the first NP is animate and the second is inanimate; but there is at least one counterexample (*diese holz sicher schiff*), and furthermore it seems just random that there is no animate entity that is needed, wanted or owned. It seems difficult to say what the functional difference between the first and the second NP in "a man needs a woman" is, or in other words: what it means to say that the first one is functionally subject and the second one object. Apparently there is no difference in the degree to which they are "agent" or "patient", neither of them may have been thematic before, and it is hard to see why this utterance should be more *about* "a man" than *about* "a woman". We have no solution to this problem. Let us mention before moving on that this static construction is never enriched by an adverbial; there is one (double) negation which immediately precedes the verb.

B. NP - V

This very frequent pattern (about 45%) comprises at least three different cases:

- a. "inherently" intransitive verbs, i.e. verbs which normally admit only one NP actant such as *spreche* "talk, speak" or, more frequently, verbs of movement; in the latter case, there is often an additional locative or directional adverbial, which, in Vito's variety, may have the form of a bare noun. Some examples:

Ø komme in eine baustell
'Ø come in a building site'

sie spazieren
'she stroll'

charlie chaplin spreche
'Charlie Chaplin speak'

diese frau sitze platz
'this woman sit place'

sie aufstehen
'she get up'

zusammen geht in diese haus
'together go into this house'

If there is a (directional or locative) adverbial, it is in final position.

b. omission of the second NP of action verbs, i.e. only the "agent" is left behind; this happens if there is no specific "object" of the activity in question.²¹ This case is quite familiar; some examples:

ich gucke
'I look'

ich esse
'I eat'

Ø immer rauche
'Ø always smoke' (=Chaplin smoked a lot)

vielleicht mädchen vergeß
'perhaps girl forget'

There are a couple of cases where it is difficult to decide whether (a) or (b) obtains, for example in the case of *spreche*. It is not fully clear whether he treats *spreche* as an intransitive verb (as in Standard German).

c. "Ergative" use of action verbs. In this case, no agent is expressed, nor taken from the context (as in the case of "zero subject"). We discussed this case in connection with (24) in section 2.4.3. Some examples:

kleine schiff mache
'small ship make'

essen keine bezahle
'eat(ing) no pay'

geschenke kinder keine bezahle
'presents children no pay'

Most of these cases are disputable, but still, this analysis sounds more consistent than granting a - totally unmarked - passive to Vito's learner variety.

It goes without saying that the functional role of the only NP may be quite different; in (b) it is an agent; in (a) and (c) it is less clear, as the small collection of examples shows. (a) and (b) allow for all forms of NP, including *she* and zero; (c) only has lexical NPs. In the case of (b) and (c), adverbials are rare; there are several occurrences of *vielleicht* which is always clause initial and makes the content of the clause as "approximate", "uncertain" or "speculative". There are a couple of negations; they immediately precede the verb.

C. V-NP

The structure is limited to "presentationals"; the only possible verb is *komme* (note, however, that the copula may have a similar function, cf. D. below). V may be preceded by an adverbial. Their construction is not very frequent (7 utterances in all), but there are a number of clear cases; we will give a few examples:

nix komme eine keil + eine holz
'no come a wedge + a log'

in diese ecke komme charlie
'at/to this corner come Charlie'

sofort komme chef bäckerei und frau
'immediately come boss bakery and woman'

komme auto polizei
'come car police' (=police-van)

V may be preceded by a particle or by an adverbial. The crucial feature of the construction is that there is just one NP which *follows* the verb.

D. "Copula" constructions

The problem with these constructions is that the copula is mostly absent; but there are some occurrences (in quoted speech), and there *would be* a copula in the corresponding target language utterance. (But see the first conclusion of section 2.4.3.)

There are two types of copula constructions, which we may label predicative and presentational:

a. Predicatives have the form NP-Cop-Pred, where Pred is either an adjective or a NP; sometimes a locative adverbial (apart from that, adverbials - except *vielleicht* - do not occur). Some examples:

diese nix
'this no'

diese schiff neu
'this ship new'

vielleicht sie keine gangster
'perhaps she no gangster'

ich bin mann brot
'I am man bread'

diese ist keine palast
'this is no palace'

There is one vexing case, in which the adjective is reversed:

schön diese haus
'beautiful this house'

It may be that it is a kind of exclamation, i.e. words attributed to the characters "beautiful, this house" (Chaplin and the girl are standing in front of the house, admiring it).

b. Presentationals resemble type C. above (cf. "there came a man" vs. "there was a man") except that there is no V but an - empty - copula. They consist of a NP which may be preceded by an adverbial:

hinten + eine grosse holz
'behind a big log'

eine policia-man
'a police-man'

nahe fluß eine haus
'near river a house'

Copula constructions constitute about 14 utterances in all.

This completes our description of the learner's utterance structures; there are some exceptions to which we will turn in a moment.

We may now sum up the learner's problem in putting his words together in two quite simple principles:

- I. *If there is only one NP, put it in final position in presentationals and in initial position otherwise.*
- II. *If there are two NPs, put that one which denotes the agent into first position, and the other one into final position.*

The two principles solve almost all of Vito's arrangement problems. However, they do not tell him what to do with "static relations" (see A.); as mentioned above we have only ad hoc answers here.

Positional restrictions of different NP types

Up to this point, we have not discriminated between different types of NP, such as indefinite NP, personal pronoun, etc.; for a survey of Vito's forms, see 2.4.2. There are some clear restrictions of occurrence which we may sum up as follows.

- (1) "Ergative" verbs only take lexical NPs without any determiner, presentationals only take lexical NPs (mostly with *eine*) or proper names.
- (2) *sie, ich* only occur in first position.

- (3) *eine* N is in complementary distribution with N, except that both can be used in the NP₁-V-NP₂ pattern with indefinite reference.

All other NPs may occur in all positions.

There are clear differences, too, between the three anaphoric constructions 0, *sie* and *diese* (N):

- (4) 0 is possible only if the same referent was in initial position in the immediately preceding utterance (excluding metalinguistic asides).
 (5) *sie* is possible only if the same referent was in initial position in some preceding clause; it need not precede immediately, but it is difficult to say how many utterances may intervene between *sie* and its antecedent. (*sie* mostly refers to Chaplin.)
 (6) *diese* takes up an entity which was referred to shortly before (mostly in the immediately preceding clause), but not necessarily in the same position. Typically, its usage corresponds to a shift from NP₂ position to NP₁ position.

Some exceptions

In a number of cases, we cannot be sure that our analysis is indeed correct; alternative interpretations are possible; but in these cases, at least one interpretation is consistent with the general picture given above. However, there are a few utterances which simply do not fit the description given above. We will mention two of them here.

- (1) Occasionally, Vito pre-poses an "object" in type A. constructions; such a case is

zusammen unfall mache
 'together(=they) accident make'

He immediately continues with the reversed order (*zusammen mache Unfall*). This may just be an error which he then corrects. The other clear counterexample is even more vexing:

(*chaplin*) *gucke zwei kinder* (..) *ein zwei drei geschenke*
 '(Chaplin) see two children one two three present'

geschenke mache geschenke kinder
 'present make present children'

The penultimate *geschenke* precedes the verb *mache*; but the whole sequence seems somewhat blurred, including the preceding and the following parts. We have no plausible analysis for this sequence: note, however, that it is the only

clear example of an intended utterance with a three-actant verb. We might speculate that at this point, Vito cannot handle more than two nominal arguments per utterance.

(2) In two cases, the initial NP is separated from V by another NP referring to a protagonist; both cases are similar:

polizei mit charlie chaplin geht in gefängnis
 'police with Charlie Chaplin go in prison'

charlie chaplin mit andere mann oder frau geht in gefängnis
 'Charlie Chaplin with other man or woman go in prison'

It may be that *mit* functions as a kind of coordination, thus bringing a coordinate NP "the policeman and Chaplin" and "Chaplin and other people". Constructions of this sort are known from pidgins and Creoles. In Standard German, they are non-standard, but not totally unacceptable (cf. English "John with his brother went to the police"). But at this point, such an analysis is but a speculation.

2.4.5 Summary

The preceding sample analysis was excessive in detail. It was our intention to illustrate the analysis, the speculations, ambiguities, interim results and even errors as explicitly as possible. The analysis presented in chapters 3-6 follows this procedure, but of course, we will not go into that much detail there. However, the reader should keep in mind that the actual analysis follows the same procedure and, obviously, is also faced with the same problems.

There is one important difference, though. The normal analysis here was maximally "open", i.e. we made a minimum of assumptions on what the learner's principles might be. The real analysis in chapters 3-6 was much more hypothesis-guided. Based on a first study which included Vito and two other learners, with language pairs Italian-English and Spanish-French, we worked out some more specific ideas about what the actual constraints might be. The result of this pilot study, and our main candidates for acquisitional principles, are presented in the following section.

2.5 Some guiding hypotheses

The results of an in-depth analysis like the one presented above are essentially three sets of observational facts:

- (a) Observations on the distribution of phrasal patterns. They concern
- the order of major constituents; Vito, for example, has V-NP, NP-V, NP-V-NP, but not NP-NP-V or V-NP-NP;
 - the form of these constituents themselves; Vito has NP realised as DET+N, bare N, Adj+N and others, but never as DET+N+Adj, although his native language allows this order.
- (b) A set of more or less systematic observations on semantic and pragmatic factors; they include, for example, animacy of the referent of NPs, referent introduction and reference maintenance, etc.
- (c) A set of observations on exceptions, complex cases, ambiguities, etc.

Observations of these three kinds are an interim result. The researcher's task is then to "condense" them to a set of general principles²² whose interaction determines the learner's utterance structures.

We shall illustrate this now for Vito and for two more learners, Ramon and Rudolfo. In Vito's case, the three types of observational facts are fairly complete (cf. 2.4); in Rudolfo's and Ramon's case, we shall only give some of them. (For somewhat more detail, see Klein and Perdue 1989.) We start with "phrasal" conditions, i.e. those which can be stated in terms like NP, PP, etc., and then link them to others.

2.5.1 *Phrasal conditions*

Any description of possible phrasal constraints depends on which phrasal categories we assume to exist in the given learner variety. This is no trivial problem, and we will not go beyond most elementary assumptions. As the detailed analysis in section 2.4 has shown, and as is confirmed by an inspection of the complete data sets, there are at least the lexical categories N, V, Art(icle), Adv(erb), Pro(noun), Pre(position) as well as the syntactic categories NP and P(rep+N)P. Other categories are disputable. Among the three learners of the pilot study, Ramon clearly has complex verbs, consisting of Aux+V, and relative clauses; this is less clear for Rudolfo and unlikely for Vito.

The analysis yields the following distributional facts. There are constructions with verb (either with one or with two NP-actants) and copula-like constructions. This gives us the following six basic patterns (subscripts are introduced for ease of reference):

	A1				A2
NP₁	V		V		NP₂
	B1			B2	
NP₁	V	NP₂		NP₁	NP₃ V
	C1			C2	
NP₁ (Cop)	{ PP Adj NP₂ }			{ PP Adv } (Cop) NP₂	

All constructions may be preceded by a conjunction or some other sentence connector; all four V-constructions may be completed by an adverbial, i.e. a locative or temporal, sometimes modal, Adv or PP. This adverbial is normally sentence-final; it may also appear in initial position, however.²³ Among all six patterns, A1, B1 and C1 are frequent; A2, B2 and C2 are rare. Before considering them in more detail, we will first see how NP is expanded. This is obviously different for the three informants. It also depends on where NP appears in the pattern:

	Vito	Rudolfo	Ramon
NP ₁	Ø <i>sie</i> <i>diese</i> (N) <i>de</i> N <i>ein</i> N N name -	Ø <i>he,they</i> - <i>the</i> N <i>one/a</i> N N name -	Ø <i>il,el</i> - <i>le</i> N <i>un</i> N N name NP + <i>il</i>
NP ₂ :	all forms but Ø., <i>sie</i>	all but Ø, <i>he,they</i>	all but Ø, <i>il</i> ; in addition Prep <i>lui</i>
NP ₃ :-	-	-	<i>lé</i>

In other words:

- All informants have three types of lexical NPs: the/a N, N (and equivalents); they occur as NP₁ and as NP₂;
- All informants may have proper names in NP₁ and NP₂;

- All informants have two anaphoric NPs, namely \emptyset and *he* (and equivalents); they occur only as NP=1.
- Vito has in addition *diese* (N), which is anaphorical and may, but need not have a lexical noun; it occurs as NP₁ and NP₂.
- Ramon has anaphoric elements (*lui, lé*) as NP₂ and NP₃, too.

Let us return now to the six patterns above. A closer analysis of all available information shows that the findings can be reduced to two general principles, depending on whether the utterance has a V (which governs one or two NPs) or a Cop:

I. The basic pattern with V is

NP₁-V-(NP₂)

except in presentationals, where the pattern is

V-NP₂

II. The basic pattern with Cop is

NP₁-Cop- $\left\{ \begin{array}{l} \text{PP} \\ \text{Adj} \\ \text{NP}_2 \end{array} \right\}$

except in presentationals where the pattern is

$\left\{ \begin{array}{l} \text{PP} \\ \text{Adv} \end{array} \right\}$ -Cop-NP₂

The exceptions only hold for Vito and Ramon, not for Rudolfo. There are some more exceptions which we will not discuss here (cf. pattern B2 above, which only occurs in Ramon's case).

It is an open question whether the category "presentational" should be listed as an exception from the normal phrasal pattern. In fact, it seems more reasonable simply to list both patterns and to account for their distribution by special pragmatic factors. We will come back to this point. In any way, it does not affect the general idea.

2.5.2 *Semantic constraints*

If the learner wants to express that Charlie has seen the policeman, rule (I.) provides him with some, but not all necessary information of how to put his words together: it tells him to put his word for "see" between his expression for "Charlie" and his expression for "the policeman"; but it does not tell him which NP comes first. There must be additional criteria. In this section, we will consider possible semantic factors.

Semantic factors may have to do either with inherent semantic properties of the referent - e.g. whether the referent is animate, human or whatever - or with properties relating to the verb or the whole activity, such as agentivity; we will call them "role properties". The example above suggests that inherent properties may be of little help: it is hard to image any semantic feature of either Charlie or the policeman which could serve as a base for their different position. In fact, an inspection of all examples shows that inherent semantic features play at most an indirect role: animate (which, in these particular texts, coincides with "human") NPs have a distinct preference for the first position. But this is simply due to the fact that referents which function as an agent tend to be animate (cf. Silverstein 1976). An agentive verb like *to make* may have non-animate agents, however, as is illustrated by the following example from Vito's text:

- (4) *diese mache schiff*
'this (=shipyard) make ship'

The crucial semantic factors (if any) are role properties rather than inherent properties. If there is a clear asymmetry between the two NPs, for example the one being an agent, the other not, then the former comes first. (We agree that this is not much of a surprise.) The problem is that not all sentences express actions with a clear agent. A principle such as "agent first" does not work for examples like the following ones (all taken from Vito):

- (7) *ich brauche eine holz*
'I need a wood'

- (14) *ich wolle eine keil*
'I want a wedge'

- (1) *sie habe brief*
'she (=Charlie) have letter'

- (23) *diese holz sicher schiff*
'this log safe (holds) ship'

One might certainly argue that the relation between the two NPs in (23),

though static, is more action-like than for example the one in (14), and the log, although no agent, is at least more agentive in (23), when compared to the other NP-referent "ship". So, one might replace categorical distinction such as "agent" and other "case roles" by a scale which also extends over non-agentive relations.

Looking through our examples, it looks as if the "degree of control" might provide us with such a scale: *It reflects the degree to which one referent is in control of, or intends to be in control of, the other referents.* The degree of control varies with the (non-negated) relation: thus, *to make* provides us with a stronger control asymmetry than, for example, ownership, as in (14) or intended ownership, as in (7). But in all of these cases, it gives us a semantic role asymmetry²⁴ - a NP-referent with "higher" and another with "lower" control. This allows us to state the following rule:

S. The NP-referent with highest control comes first.

We shall refer to this semantic (S.) condition as the "controller-principle".

Admittedly, the relational property "being in control of" needs a more precise characterisation, for example in terms of verb classes.²⁵ But this being granted, rule S. solves a great deal of arrangement problems in sentences with two NPs.

It does not work, however, for copula constructions with two NPs. The relation "x is y" seems to provide no semantic asymmetry. Hence, the difference between *The girl is the thief* and *The thief is the girl* - must be due to other factors, to which we will now turn.²⁶

2.5.3 Pragmatic factors

Remaining problems and observations

There are two arrangement problems left, which cannot be accounted for by L, IL, or S. These are the symmetric (copula) constructions and the specific occurrence restrictions of NP-types: anaphoric NPs, including \emptyset , are subject to specific positional constraints. We will start with the latter problem.

Let us first briefly recapitulate the facts. All informants have three types of NP:

- names (actually only *Charlie*); they occur in any position;
- NPs with a lexical noun; there are three types: (a) bare N, (b) *the* N, (c) *a* N (and equivalents); they may all have an additional modifier, e.g. an adjective; they occur in any position;

- purely anaphoric NPs; there are two types: (a) \emptyset , (b) *he* (and equivalents); they occur only in initial position;

There are several other constructions, not shared by all informants. The important ones are:

- Vito has *diese* (N) "this (N)"; it occurs in any position;
- Ramon has anaphoric or cataphoric NPs in second NP-position: *lé* in NP-*lé*-V, and *lui* in NP-V-Prep-*lui*. In addition, full NP₁ may be "copied" by a pronoun: NP+*il*.

As we have seen in the analysis of Vito's text, and as one would expect to find, the occurrence restrictions of these various NP-types are closely related to whether a referent is first introduced, re-introduced or maintained from some preceding utterance. An inspection of all examples leads to the following - quite straightforward - conclusions:

- a. *the* N (and equivalents) and names are used when the referent can be assumed to be known to the listener, either because it was referred to before, or because it is associatively linked to some entity referred to before, or because it can be assumed to be part of the listener's general knowledge; the first case is clear enough; the second concerns examples such as Rudolfo's

the police take the children

where *the father* was talked about before; so, *the children* means: "the children of that father". The third case is exemplified again by Rudolfo:

in the morning go to the work

- b. *a* N (and equivalents) is used for first introduction of a referent.
- c. The use of bare N is not totally clear. It is often used to introduce or to maintain non-countables. But it also occurs with countables, and then, it *tends* to be used for first introduction; but there are many clear counterexamples where it goes on a par with *the* N. It may well be that the use of bare N reflects still previous learning stages (for a more detailed investigation of this problem cf. Carroll and Dietrich 1985).
- d. For completeness' sake, let us state that \emptyset and *he* are used to maintain a previously introduced referent; so do *diese* (N) in Vito's case, and *lé/lui* in Ramon's.

This leaves us essentially with one important question: What causes the different types of maintenance, namely name, *the* N, \emptyset , *he* (plus *diese* (N) for Vito and *lé/lui* for Ramon)? In Ramon's case, there is also the question under which condition he uses NP+*il*.

In all cases, the referent has already been introduced. The difference between them apparently has to do with the nature of the referent (in these texts, human or not), and the position where it was referred to before: in which utterance, and where in this utterance. We can illustrate this with the first utterances of Vito's text.

- (1) (Charlie is introduced in the previous discourse, which is not part of Vito's narration.)
sie habe brief...
- (2) \emptyset *komme in eine baustell ...*
- (4) *diese mache schiff*

For \emptyset to apply, the referent must be maintained:

- (a) from the immediately preceding utterance of the narrative text
- (b) where it is in initial position.

If (a) does not apply, *sie* must be used; if (b) does not apply, Vito uses *diese* (we will come to the other informants in a moment). There are two problems here. First, condition (a) is often violated in that \emptyset may jump over intervening utterances; but these are cases of utterances which do not belong to the "plot line" of the story; they may give background information or be metalinguistic comments. In order to make (a) more precise, we have to account for this difference in pragmatic function within the text. Second, condition (b) seems to reflect some functional difference: intuitively speaking, \emptyset and (perhaps) *he* seem to require "topic maintenance", whereas the switch from *in eine baustell* to *diese* seems to reflect a transition from "focus" to "topic". Now all of these terms are highly disputable. In what follows, we will work out a simple proposal to account for these problems in the present context.

Topic, focus and quaestio

Very often, a statement is used to answer a specific question, this question raising an alternative, and the answer specifying one of the "candidates" of that alternative. For example, the question "Who won?" raises an alternative of "candidate" persons - those who may have won on that occasion, and the answer specifies one of them. A question may raise all sorts of alternatives, e.g. actions ("What did Charlie do?"), contents of prepositions ("Was Charlie *before* or *behind* the ship?"), etc. Let us call "focus" that part of a statement which specifies the

appropriate candidate of an alternative raised by the question, and "topic" the remainder of the answer.²⁷

Now, not all texts are question-answer-sequences. But we may assume that any statement is an answer to an (implicit or explicit) *quaestio*. Thus, Vito's *Ø komme in eine baustell* is an answer to a (implicit) *quaestio* "What did Charlie do at that time?", whereas *diese mache schiff* answers an (implicit) *quaestio* "What did this baustell do?". Note that the two statements serve quite different functions within the whole narrative: the first indeed belongs to the "plot line" - the foreground of the story - whereas the second gives (relevant) background information: it does not answer the "key question" of the whole text, which is: "What happens with Charlie (and possible other protagonists)?" Thus, all utterances which answer an implicit *quaestio* "What happens with Charlie at time t_i ?" belong to the foreground - they push forward the plot line - and all other utterances, no matter which (possibly very important) *quaestio* they answer belong to the background.

More generally, the producer of a given text-type can be seen as answering an (implicit or explicit) question, the *quaestio* of the text. Each utterance which supplies information directly relevant to the *quaestio* belongs to the foreground of the text-type. Foreground utterances follow on from each other in a non random way insofar as the information they introduce and develop is constrained by the *quaestio*. (Background utterances can be thought of as "answering" various subsidiary questions.) The different sources of information introduced and developed in the text are derived from the operations a speaker effects when producing an utterance, namely:

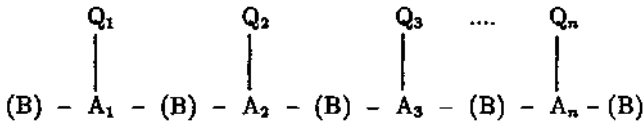
- expressing a propositional content comprising a verb and its actants;
- instantiating this content, in particular by spatio-temporal contextualisation, and determination of the entities involved;
- assigning a modality.

The way in which information from these "referential domains" of events (actions, processes, states, or 'events' for short), entities, space, time and modality is introduced, maintained or developed is dependent on the *quaestio* of the particular text-type. For example, objects are introduced and spatially contextualised in the foreground of living room descriptions (Ehrich 1979).

In this context, a narrative - or for our immediate concerns, a film-retelling - is seen as a complex overall event whose foreground comprises singular events (and their protagonists). It is an answer to a question-function Q_i where i ranges over time intervals:

- Q₁: What happens with p at t₁?
- Q₂: What happens with p at t₂?
- Q_n: What happens with p at t_n?

This question-function²⁸ defines the foreground of the narrative: all utterances which are answers to one of these questions are "foreground-utterances". It also defines topic and focus within each of these utterances: The topic of a foreground clause includes a time span t_i (which is mostly not explicitly specified but given by the sequential order, except for t₁) and the protagonist, or some more or less presupposed protagonist (see the "Focus principle" below). The focus specifies the action or event at that time span²⁹. As background clauses are normally linked to a foreground-clause; their internal focus-structure is quite different, depending on what information they specify. The overall structure of a narrative is then (A = foreground, B = background).



This structure plays an important role in the narratives of all informants. There are also some immediate implications for word order to which we will turn in a moment.

We may now restate our observations in connection with (Vito:l-4): 0 maintains a referent from T in A_i to T in A_{i+1}, *sie* a maintenance from somewhere to T in A_i, *diese* (N) from F in A_i to T in A_{i+1}. In other words: what matters for the use of the various types of NP is not only whether something is maintained or introduced, but also whether it goes from T to F, from F to T etc., and which clauses intervene.

An analysis of the whole texts shows the following regularities of referential movement (M) for entities:

- M. 1 Transition from "nothing" to T (=introduction in topic): lexical NP** (except *diese* (N) in Vito's case). Note, however, that *the* is quite rare.
- 2 Transition from T to T ("topic maintenance"): Ø, he, the N; Ø** only applies when the two clauses are adjacent, and where B-clauses do not normally interrupt adjacency of A-clauses. The difference between *he* and *the* N is not totally clear; it seems, however, that *the* N is used when there might be an ambiguity, e.g. when there is an intervening NP which could be misinterpreted as coreferential to *he*.

3 Transition from P to T: *he* and, in Vito's case, *diese* (N).

4 Transition from "nothing" to F ("Focus introduction"): lexical NP (or name), except *diese* (N); again, *the* N is rare (it is only used when the referent can be assumed to be known, although it was not mentioned before).

5 Transition from T to F: there are few clear examples, probably lexical NP (or name) for Vito and Rudolfo as the following utterance shows:

one man ... go to the work
his wife kiss the man.

6 Transition from F to F: few clear examples, probably lexical NP (or name).

This leaves us essentially with Ramon's two peculiarities: The difference between *lé/lui*, and the use of constructions such as *Charlie il*. We cannot solve the first problem here: it is the problem of what a "clitic" is in Ramon's variety. One might argue that *lé* is part of the (larger) topic-component (cf. footnote 27), whereas *lui* belongs to the focus, or alternatively, that the focus-component has a different internal structure in both cases; in any event, this would need a more detailed treatment of focus-structure, which goes beyond the scope of this section. The *Charlie+il* pattern consistently serves to reintroduce a topic: it "highlights" that there is a new topic with respect to the preceding clause; but that the referent of this topic has already been introduced into the discourse.

So far, we have used the T-F-distinction only to describe the regularities in the use of various types of NP. But apparently, it also has a direct bearing on the word arrangement problems. The learners establish a formal reflexion of the T-F distinction, grouping the F-information towards the end of their utterances. The crucial rule is quite simple:

P. F comes last

We shall refer to this pragmatic condition (P.) as the "Focus principle". In its present form, it is gross; in particular, it does not take into account that topic-component and focus-component may have, and indeed mostly have, an internal organisation. Still, it helps to answer some open questions.

One of them is the "copula-problem" mentioned above: If a copula has two NP-arguments, then there is no semantic criterion to decide about their order, as in the case of NP-V-NP. Here, rule P. puts that NP which is (or belongs

to) F at the end. Rule P. also gives a possible explanation of the word order in Ramon's and Vito's stative (Cop-NP) and dynamic (V-NP) presentationals: It seems at least plausible to assume that in presentationals, the NP-referent is not presupposed, so the NP is the answer to the (usually implicit) *quaestio* where the protagonist is not specified, see 2.5.4 below, and hence, it comes last. Rudolfo, on the other hand, mainly sticks to a strictly phrasal principle, the principle: "NP before V", defined in purely categorial terms, wins over the pragmatic rule P. And finally, rule P. provides us with an explanation of a problem which was mentioned only in passing: the position of (temporal and locative) adverbials, which - as was said in 2.5.1 - may appear in initial or in final position; as an inspection of the examples shows, this difference in position is clearly related to their function as part of the topic-component ("orientation") or as part of the focus-component.

Summing up, we have found that three "pragmatic" factors play a role for the utterance organisation in learner varieties:

- Familiarity: Can a referent be assumed to be known to the listener, either by world knowledge, or by contextual information of various sorts? This factor "familiarity" is least important: it basically decides about the use of *a* N vs. *the* N.
- Maintenance (vs. introduction): Is a referent first mentioned, or was it already referred to in a previous utterance?
- Topic-focus-structure: Does a constituent specify a candidate of the alternative raised by the *quaestio* of the utterance? This *quaestio* may be asked explicitly; but it may also be given implicitly, for example by the "key question" of the whole text which at the same time defines which utterances of the text are foreground and background.

2.5.4 Summary

In the pilot study, we have found five organisational principles. Two of them can be stated as purely phrasal patterns, i.e. in terms like NP, V, Cop, etc.; one uses semantic notions ("Controller first"); and two operate with categories like topic, focus, introduction and maintenance of reference; among these two "pragmatic principles", the first seems to be rather a cluster of different factors, whereas the other is simple and looks rather general ("Focus last"). These principles will guide us through the empirical analysis of the subsequent chapters. We assume that they must be modified, refined and completed in various aspects.

There may be, first, additional phrasal patterns; for example, the verb may not govern one or two, but three nominal actants.³⁰ It may be, second, that the "controller principle" is too crude, especially if there are more than two entities referred to. An appropriate refinement will be suggested in chapter 5.2.3.

And finally, there may be complex hierarchical organisations of topic-focus-structure. This becomes immediately clear as soon as we look at the word order possibilities of a fully developed language like German, where an utterance with the same lexical elements, the same verb-argument structure and the same case roles may have variants like these

Ich habe drei Bücher gekauft.

Drei Bücher habe ich gekauft.

Bücher habe ich drei gekauft.

Gekauft habe ich drei Bücher.

Bücher gekauft habe ich drei.

All of these reflect subtle, but distinct differences in the topic-focus organisation, which obviously cannot be grouped by a rule as crude as P. It may well be, however, that at least some salient patterns of learner varieties can be uncovered by supplying the appropriate (implicit) question which the utterance is meant to answer. It may be, for example, that "presentationals" with the special pattern have this specific order because they answer a different question from normal foreground utterances in narratives (like: "What was next on stage?" compared to "What happened with p next?"). All of this will be considered in the empirical analysis of chapters 3-6. There, we shall also consider cases where the various principles are in "competition" (Bates and MacWhinney 1987). What happens, for example, when a clear controller is in focus? Then, the "controller principle" would require the NP which refers to that referent in initial position, whereas the "Focus principle" requires it in final position.

Apparently, informants are quite skilful in avoiding such situations. But if they occur, the relative weight of the various principles may be quite different, depending, for example, on the source language. It also seems that even at early stages the informants are quite sensitive to the specific balance which the target language has in these conflicting situations. Thus, Rudolfo and Vito have the same language of origin. But Vito is much more willing to give up rigid phrasal patterns in favour of semantic or pragmatic factors than Rudolfo. This seems to reflect an early "tuning" to the respective target languages, German and English, in this particular case. Hence, we will pay special attention to these exceptional but particularly interesting conflicting situations.

Notes

1. This chapter is largely based on and at least partly identical with Klein and Perdue (1989). We are grateful to our colleagues in the project and to the participants of several conferences and workshops where we presented our ideas for their help and advice. We would especially like to thank Maya Hickmann, Eric Kellerman, Brian MacWhinney and Christiane von Stutterheim for their detailed comments.
2. Obviously, there are other factors of variation, for example the type of task which the speaker tries to solve. We will come back to at least some of them.
3. Chomsky's "Universal Grammar" is a much-cited candidate in SLA publications. However it is not apparent to us that it makes clear predictions that are *relevant*, as we will explain.
4. See Behaghel's "Gesets der wachsenden Glieder" (Behaghel 1923-32) which predicts that longer constituents *ceteris paribus* come later.
5. Both for agenthood and animacy, one could refine this principle by talking about "degrees" of agentivity and animacy. For simplicity's sake, we will ignore this possibility here, but will return to it in section 2.5.2 below.
6. Note that this is no argument against the existence of such constraints or even the constraining force of universal grammar in language acquisition, both first and second. It only defers its influence to later stages. Universal grammar in this sense resembles a husband who comes to the kitchen and offers his helping hand when the washing up is done except for three spoons and a saucepan.
7. Some instances of quoted speech are apparently induced by inter-titles in the film. But this only accounts for some cases, and it raises a number of problems, as well. Why does he reproduce these forms - which he obviously understands - without taking them over into his own production? For example, at one point, he quotes Charlie Chaplin: "*Ich bin Mann Brot*", which is apparently shaped after the inter-title "*Ich bin der Mann mit dem Brot*"; but a few utterances later, he repeats this episode and then, he says: "*Ich Mann Brot*", which clearly corresponds to his own production rules. - Incidentally, all inter-titles were presented for a rather short time only and simultaneously in three languages (German, Turkish, and Italian); one cannot exclude that Vito's attention was focussed on the Italian version; but he clearly registered and reproduced the German version.
8. For a detailed analysis of Vito's lexicon, see Dietrich 1989a.
9. The normal base forms end with -e; this corresponds to the infinitive, first person singular and the first and third person plural in the local (Heidelberg) dialect. But for some mysterious reasons, he uses *geht* "go" (11x) more often than *gehe*.
10. *komme* has some peculiar uses; it often marks the achievement of a state of affairs (roughly like *to become*); sometimes, it goes with untypical word order; we will come back to this point.
11. Obviously, it cannot be said whether the whole narrative is in the present or in the past, since there are no such distinctions here. For simplicity's sake, we will always use present forms in the glosses.

12. Since this may sound trivial, it should be mentioned that it is far from being common in elementary learner varieties; see, for example, Klein and Rieck (1982).
13. Standard German requires a definite article here; this article must be in its elided form: ...*Brief vom Gefängnis*; forms such as ...*von einem Gefängnis*, *von dem Gefängnis* are inappropriate, *von Gefängnis* is impossible.
14. This term is used in an everyday sense, that is, a referent is thematic, if it has been explicitly referred to, if it has been pointed to or if every normal speaker would infer it rather than any other possible referent, in the given situation.
15. It is rare, however, to have the verb first in declarative clauses. Hence one is inclined to assume that there "must" be something else in first position which is just elliptical. Thus, when asked to give the "complete" version of stereotype elliptical statements such as "*Kann schon sein*", people would always say "*Das kann schon sein*" rather than "*Kann das schon sein*". - For a careful discussion of "subject in German", see Reiss (1982). She concludes that the different features associated with subjecthood (Keenan 1976) do not give a consistent picture in German and that this concept could be easily eliminated from descriptive terminology, as far as German is concerned.
16. Note, incidentally, that the specification of an entity already referred to demands here an expression containing the indefinite, rather than the definite article.
17. *Keil* is inserted after a metalinguistic passage (...M...), in which he asks for that word.
18. As has been shown in Dietrich (1989b), Vito is relatively weak in his noun repertoire (compared to his verb repertoire). He compensates for this by frequent metalinguistic questions (*was ist name diese?*) and then inserts the appropriate item. It is interesting to note, that the information given by the interlocutor only has "local" effects, that is: Vito uses the word and then forgets it again, as is clearly the case with *keil*, in examples 8 and 14.
19. For an extensive discussion of some of these parallelisms, see Keenan (1984).
20. The term "actant" is borrowed from Tesnière (1959). It includes NPs, but also PPs governed by V, in contrast to "free" PPs ("circonstants", to use Tesnière's term) which may be used, for example, to situate an event in space and time. In what follows, we shall often ignore the difference between NPs and PPs in the former sense (i.e. as "actants") and simply talk of "NP".
21. Note that this case is quite different from "aero subjects"; in the latter case, there is always a specific referent maintained (and hence not explicitly referred to again) in initial position. Here, we have an "absolute" use of a verb, without a specific "object"; cf. "I have already eaten" vs. "I have eaten a chicken",
22. There is no special theoretical claim connected to the term "principle". We interchangeably speak of general "principles" which the learner follows, of general "conditions" determining his learner variety, or of the general "constraints" on his utterance structure.
23. It is easy to see that all of these patterns, including possible complements could be put in the form of one or two phrase structure rules, of the type S --> ...; we will not do that here, since it gives no additional information.
24. One might wonder, of course, who is in control of whom in *Peter loves Mary*. But note that we are talking about real and intended control.

25. In the linguistic literature, there are a number of semantic/pragmatic features which relate to word-order, for example Silverstein's (1976) "person/animacy/agentivity"-hierarchy, or Hopper and Thompson's (1980) transitivity hierarchy. The problem with these hierarchies is, however, that they are based on a whole set of quite heterogeneous properties, and it is not at all clear why these properties converge - if at all, or whether there is some principle behind their clustering. This, incidentally, is not to deny the genuine insight which these hierarchies may give us.
26. This would apply to the relation *to have*, too, in cases such as *This cup has one handle*, i.e. in pure property assignments. (Note that a sentence such as *One handle has this cup* is perfectly natural in German.)
27. It goes without saying that this treatment is very crude at this point; it would be more appropriate to talk about "focus component" and "topic component", which in turn have a hierarchical organisation.
28. Narratives may differ to some extent with respect to the "key question". We have chosen the relatively neutral formulation "happens with p", although it gives the protagonist (or protagonists) a somewhat passive role. Let us add that it might be more appropriate to characterise a narrative by two "key questions" [Q_i, O_i], where Q_i refers to the "rooting" of the whole event in time and space (Labov's orientation): "When and where did a happen", when a is the total event (one may indeed imagine a third "key question": "so what?!"). - This whole approach, which also applies to other types of texts, is worked out in more detail in Klein and von Stutterheim (1987) von Stutterheim and Klein (1989); here we give only the rough idea.
29. This means, incidentally, that foreground clauses normally cannot be imperfective or stative: one could characterise exceptions to this "norm" to a certain extent: - when an event involving a central protagonist itself serves as a temporal frame for another event involving him/her or where an utterance acts as a "conclusion", giving the result of events involving him/her, as in Rudolfo's "conclusion" of the bread-stealing episode: *Chaplin is free + the girl is/go to the prison*.
30. We have seen one example of Vito's in 2.4.4, which posed problems of analysis. As will be seen in the following chapters, arrangement problems with three-actant verbs constitute a serious obstacle for some learners.