Part I

Production
1 Utterance structure

Wolfgang Klein and Clive Perdue

1.1 Introduction

This chapter deals with the way in which learners put their words together – that is, with the ‘syntax’ of learner varieties.¹ The term ‘syntax’ is normally used to refer to particular formal constraints on the structure of utterances, stated in terms such as verb phrase, subject, case agreement, and similar ones. Thus, syntactic descriptions of a particular language consist of statements such as

- the finite verb is clause-final;
- genitive attributes precede their head;
- verbs of type x govern an indirect object;
- an indirect object is morphologically marked as dative;
- the subject immediately precedes the finite verb;
- there is number agreement between finite verb and direct object;
- adjectives agree in person, number and gender with their head nouns,

which have constituted ever since Priscian, a descriptive language which every linguist is held to understand. But they are of limited use in investigating learner varieties, as we may illustrate by a short look at what such a learner variety typically looks like. Imagine you are an Italian learner who has been living and working in England for about six months. Then, you might know from everyday contact

- a number of proper names, such as John, Peter, Mary,

¹This chapter is a synthetic and necessarily simplified summary of joint work by Mary Carroll (the acquisition of English and German), Josée Coënen (Dutch), José Deulofeu (Moroccan learners of French), Thom Haeberer (English), Wolfgang Klein (Dutch and German), Clive Perdue (English and French) andAnne Trévis (Hispanic learners of French) brought together in the Final Report VI to the European Science Foundation (Klein and Perdue 1988). It comprised a pilot study, which has since been published (Klein and Perdue 1990) and a main study, also published (Klein and Perdue 1992), as well as numerous articles by all these authors (see the ESF Bibliography in Volume I).
4 Adult language acquisition: results

- a number of noun-like words, such as beer, bread, work;
- a number of verb-like words, such as work, love, see, give;
- a few adverbial-like words, such as then, Christmas, today;
- a few numerals, one, two, three;
- and a few complex rote constructions, such as how are you?

At this stage, your learner variety most likely has no inflexion, hence no case morphology, no finite morphology and no agreement whatsoever. Further, it will have no, or very few, function words, such as determiners or prepositions (for case marking). In such a variety – and this is a variety through which virtually all untutored learners pass – ‘syntactic’ constraints in the Priscianian sense play no role at all or a relatively minor one. Even if you happen to have concepts such as ‘finite verb’, ‘agreement’ or ‘dative marking’ from your first language, you cannot apply them in your second language production. Many early productions even lack verbs. Therefore, government by the verb cannot be operative, except to the extent to which it is inferable from the context. The same is true for prepositions and NPs governed by them.

Does this mean that learner varieties are just chaotic collections of words, thrown together at random? This is, at first blush, an empirical question. But a closer look soon shows it not to be the case. The utterance structure of learner varieties is governed by other organisational principles, which are also present in fully-fledged languages, but with less weight – for instance principles based on what is maintained from a previous utterance and what is freshly introduced (‘referential shift’), on what is topic information and what is focus information, on the semantic role property (‘thematic role’) of an entity, etc. So utterances rather follow a constraint such as ‘focus last’ or ‘new information first’ or ‘agent precedes patient’ than a constraint such as ‘finite verb is clause-final’. Hence, any serious investigation of the internal structure of learner varieties, and as a consequence, any deeper understanding of the nature of ALA, requires going beyond purely ‘syntactic’ constraints in the narrower sense and including organisational principles of the latter type.

The central idea of the present investigation, then, is this: There is a limited set of organisational principles of different kinds which are present in all learner varieties, including the borderline case of fully-fledged languages. The actual structure of an utterance in a learner variety is determined by an interaction of these principles. The kind of interaction and hence the specific contribution of each principle may vary, depending on various factors, for example the
Utterance structure

speaker’s language of origin. In particular, the interaction changes over time. Picking up some components of verb morphology from the input may cause the learner to modify the weight of purely ‘syntactic’ factors in his utterance organisation. From this perspective, learning a new feature is not adding a new piece to the puzzle which the learner has to put together. Rather, it entails a – sometimes minimal, sometimes major – reorganisation of the variety, where the balance of the various factors change in a way which, eventually, brings it close to the balance characteristic of the target language.

This approach holds that the utterance organisation in learner varieties is characterised by a two-fold systematicity. There is, first, a ‘horizontal’ systematicity – the balance between various principles which obtains at any given point in time; it is this balance which constitutes the ‘syntax’ (in the broader sense of the word) of a given learner variety. There is, second, a ‘vertical’ systematicity which leads from one learner variety to the next. Such a change is induced by the intake of new information from the input – so long as the learner is able to take new input in. For some learners, this process comes to an early halt. They are either unable or unwilling to further modify their system: their language fossilises. Fossilisation in this sense does not necessarily mean that the learning process has come to an absolute halt. The learner may still enrich his vocabulary, for example. But he does not add features which would lead to a potential structural reorganisation. Other learners however, may do so: they modify the balance reached at some point, and set out to construct a different type of interplay between the various organisational principles. Such a venture is a risk, and this might explain why so many learners are reluctant to abandon a variety which, though still far from the target, allows them to express themselves in a way they feel sufficient for their communicative needs.

The approach we take here is quite different from perspectives dominant in ALA research in at least two ways. First, it does not look at the acquisition process from the end – the alleged properties of the target language – but rather at the internal structure of a learner variety at a given point in time. A learner variety is not so very much viewed as a rudimentary, imperfect or faulty simulation of the target but as a system in its own right, ruled by a particular interplay of the same principles as any other language. Second, these principles are not confined to Priscianian syntax (i.e., properties such as the ones mentioned above), but include any type of constraint which might influence the structure of an utterance.
6 Adult language acquisition: results

The reason why we have taken this perspective is not so very much a general discontent with Priscianian syntax but the simple fact that we see no other way to understand what is systematic about early learner varieties, and what is systematic about the way in which later learner varieties emerge from them. In more practical terms, the investigation has taken the form of a number of inductive and hypothesis-guided case studies. Based on a first, explorative study with three learners, we developed a number of hypotheses of what learners with different source languages could do to put their words together. Then, we studied what they did do.

The remainder of this chapter is organised as follows. In section 2, we shall briefly characterise the informants and data used for this particular study. Section 3 presents the hypotheses for organisational principles – those observed in the explorative study. Section 4 contains the general results and a discussion of possible causal factors, for example source language influence.

1.2 Informants and data

For various reasons, the Swedish data could not be included in the study. For all other SL/TL pairs, a minimum of two informants were analysed in detail over a period of three cycles. Occasionally, data from other informants were included in accordance with the ‘procedure of mutual compensation’ explained in Volume I:3.1. The main informants are: Madan, Ravinder, Santo, Lavinia, Tino, Angelina, Çevdet, Ayşhe, Ergün, Mahmut, Fatima, Mohamed, Abdelmalek, Zahrar, Berta, Gloria and Paula, who are described in Appendix B of Volume I.

Since it was simply impossible to exploit the whole range of data collected and available, we decided to concentrate on film retelling, described in detail in Volume I:6.3. An abridged version of Chaplin’s Modern Times was used, whose plot may be found in Appendix C of Volume I, together with text samples. The procedure is particularly interesting for present purposes. Firstly, it is a complex verbal task, in which the learner retells a series of events whose relationship to each other must be specified. Within each event, he has to say who did what to whom, introducing new characters and maintaining characters who are already on stage. The main characters are male (Charlie Chaplin) and female (the young girl), and they act and are acted upon. Therefore, the informant has to deal with referent in-
roduction and maintenance in a wide range of semantic functions. Secondly, we have some control over the learner’s retelling of the story, in that we have the (abridged) film to compare his production with.

Each learner did the retelling three times (once per cycle; see Volume I:5.3), at an interval of about ten months on average. His or her retelling was recorded and transcribed for analysis. Depending on informant and cycle, the length of this transcription varied between thirty and three hundred utterances. The word (token and lemma) count for thirteen of the learners analysed here may be found in Table 8.7 of Volume I.

1.3 The pilot study

In a first, exploratory study, we had a close look at the learner varieties of three informants, who are not part of the main sample: Vito (Italian-German), Rudolfo (Italian-English), and Ramon (Spanish-French), concentrating on the ‘shipyard episode’, in which Charlie causes the premature launching of a ship. The idea of this pilot study was to develop some ‘guiding hypotheses’ on which principles might underly the learners’ utterance organisation. In what follows, we illustrate the procedure in some detail for one informant, Vito. In section 4, we sum up the ‘guiding hypotheses’, resulting from the entire pilot study.

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 did not have very much contact with the German population. He worked in a kitchen of an Italian restaurant, was married to an Italian woman and had no children.

On the other hand, he was talkative, self-confident, lively and very interested in questions of language. As a consequence, his metalinguistic behaviour during the retelling is quite elaborate. 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
8 Adult language acquisition: results

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 were kept apart in analysis:

(a) Purely metalinguistic speech which largely consists of 'rote forms' such as Was ist der Name?. Their inclusion would have led 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.

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.

Vito’s linguistic repertoire
The following description of Vito’s repertoire relates to the entire retelling (he retold the shipyard episode in thirty-one utterances).

Morphology
Vito has no inflexional morphology, hence no case marking, no agreement, no tense.

Lexicon
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 sixty different nouns, the most frequent ones being Mädchen (referring to the girl, twenty-seven occurrences within 1050 words of running text), Polizei (or Policia-Mann) ‘policeman’ (22x) Frau ‘woman’ (17x), Gefängnis ‘prison’ (15x), Auto ‘police car’ (13x),
Utterance structure

Holz ‘wood’ (13x), Brot ‘bread’ (12x), Schiff ‘ship’ (10x). Clearly, the main protagonists and features of the film show up most often. The name ‘Charlie Chaplin’ occurs twenty-four times; Chaplin is by far the most important referent, mostly denoted by a personal pronoun, sie.

Verbs:

He uses about forty different verbs, some of them with a rather overgeneralised meaning (nouns, in contrast, are rarely overgeneralised, 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 sages ‘say’, the most common verbum dicendi in other German learner varieties.
(c) Komme ‘come’ (23x), geht ‘goes’, nehme ‘take’, bringe ‘bring’ (1x). He totally misses the deictic component of these verbs;
(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 (fände); hence, there is no reason to assume that it is used to mark past or perfectionality in opposition to present; it is interesting, though, that a verb with inherent perfective meaning is the first attested perfect form.
(h) He uses a number of compound verbs, such as rausgucke ‘look-outside’, aussteige ‘get-out’,
10 Adult language acquisition: results

_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 above). 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 spatial 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).
Utterance structure 11

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 ten 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 what follows, we will illustrate the way in which he proceeds by a closer look at the first ten utterances of the shipyard episode. These utterances are somewhat ‘edited’, that is, we have omitted obvious false starts and breakdowns, hesitations, interjections, and metalinguistic comments. In addition, random 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.)

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/had 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:

(a) 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