

On-line syntax: Thoughts on the temporality of spoken language

Peter Auer

Deutsches Seminar 1, Universität Freiburg, 79085 Freiburg, Germany

Accepted 17 October 2007

Abstract

One fundamental difference between spoken and written language has to do with the ‘linearity’ of speaking in time, in that the temporal structure of speaking is inherently the outcome of an interactive process between speaker and listener. But despite the status of “linearity” as one of Saussure’s fundamental principles, in practice little more than lip-service is paid to the temporality of spoken language, which is treated as having few if any consequences for syntactic analysis. It is trivial to point out that a structuralist definition of the sentence is incompatible with an on-line model of syntax processing. A structuralist analysis, even of ostensibly spoken language, is carried out not from a real-time emergence perspective but as if it were – like a written text – a finished product. This article suggests that a significantly untraditional approach to syntax is required when one focuses on its on-line emergence, and outlines such an approach.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Emergent grammar; Language and time; Speech *vs.* writing; Syntax of spoken language

1. Introduction

It is almost commonplace to assume that one of the fundamental differences between spoken and written language resides in the ‘linearity’ of speaking in time.¹ Of course, this is not meant to imply that the act of writing – like the act of speaking – does not require time; rather, it means that the temporal structure of speaking – unlike that of writing – is inherently the outcome of an interactive process, a dialogue between speaker and listener; it is this temporal aspect of speech to which the term *on-line* in the title of this paper refers.² Writing is a private act; only the finished product of this act comes into contact with the addressee. Spoken language must (and can) reckon with interactive openness from the very outset.

Despite “linearity” being one of Saussure’s fundamental principles of language,³ only second to his much more often-cited “arbitrariness”, reference to the temporality of spoken language is often nothing more than

E-mail address: peter.auer@germanistik.uni-freiburg.de

¹ Cf., among many, Chafe (1979) and Pawley and Syder (1983).

² Some background on the temporality of language can be found in Auer et al. (1999) (particularly in the first chapter), in Auer (2005); as well as in Hopper (1989).

³ Cf. Saussure (1974, p. 103) 1st part, 1st chapter, Section 3 = Second principe; caractère linéaire du signifiant: Le signifiant, étant de nature auditive, se déroule dans le temps seul et a les caractères qu’il emprunte au temps: (a) *il représente une étendue*, et (b) *cette étendue est mesurable dans une seule dimension: c’est une ligne*“ (emphasis in the original).

lip-service in linguistics and has few consequences for analysis. The question of whether the approach to syntax would have to differ significantly from the traditional approach when one focuses on its on-line emergence has until now hardly been asked (much less answered). The following article is an attempt in this direction and thereby follows a recommendation of Silverstein (1984, p. 182, *italics P.A.*): “We should identify surface forms of what we might call sentence-scope, *as these unfold in realtime discourse*”.

The first part of this article deals with the meaning of temporality for spoken language research – indeed, the issue is not as clear as one might expect. In the second part, I examine the consequences which an on-line approach to spoken language has or should have; in doing so I will explore some typical features of spoken syntax.

2. Temporality as a feature of spoken language

The temporality of spoken language can be viewed as both interactive and cognitive, and from at least three different perspectives.

2.1. Transitoriness (“rapid fading”)

The first perspective refers to the transitoriness of spoken language (its “rapid fading”, as Hockett (1960) has called this “design feature” of language). The ‘present’ of spoken language is limited to the time span within which the speaker and the hearer can retain it in memory. This is true of all direct (face to face or telephone) communication as long as its transitoriness is not suspended by technological means.

Transitoriness has an interactional side, since it limits the speaker’s responsibility for what has been said and the hearer’s responsibility for what has been heard. Whereas there will rarely be a dispute over what has been written (to someone) as long as the parties have access to the original or a copy of it (a written message can be retained, at least for a relatively long time and can be preserved and produced; it is a record), references to what has been allegedly or supposedly said or not said, and in general retrospective reconstructions of past interactional events (that is, stories about previous interactions) are a permanent problem in interaction. ‘Remembering’ what has been said is an interactive process which in itself is costly and prone to error and disagreement. Not only does it lead to quarrels, it also is part of interpersonal arguments, where the transposition of the actual event to the meta-level of who said, or did not say, or did not understand, or did not intend, is routinely an issue.⁴

That our memory is poorly adapted to the storage of speech, and is subject to quite evident limitations even for the content of what has been heard, but certainly for its form, has implications for the production and processing of language as well. One superficial correlate for language production is the fact that speakers sometimes ‘lose the thread’ of/in complex constructions, and that hearers presented with such constructions may have corresponding problems of processing. Normally both are indicative of an overload of short term storage. Spoken language copes with these conditions of oral communication by having shorter basic units of processing⁵ and by avoiding types of construction that require processing *against* time. Especially pertinent here is the preference for so-called right- as opposed to left-branching (let alone center-embedded) constructions, which holds for constituents that exceed a certain degree of size or complexity (as, for instance, relative clauses). Syntactic expansions that unfold in time (‘right-branching’), however, are possible in spoken language and can be both extensive and elaborate. (The widespread opinion that spoken language is less hypotactic than written language is a myth; cf. Auer (2000b).)

An explanation for the asymmetry between right-hand and left-hand branching can be seen without difficulty to lie in the linearity of on-line syntax: left-branching structures can often only be processed completely when the head of the construction has been reached (i.e., when they have been produced completely). In rightward branching constructions, which proceed ‘in time’, syntactic processing can begin right after the head constituent has been identified, since its dependent constituents will not usually make a revision of the external syntax of the construction necessary. According to Behaghel’s (1932, p. 4) famous First Law of language seri-

⁴ Cf. research on reconstructive interactional genres, e.g. Bergmann and Luckmann (1995).

⁵ Roughly in the size of Chafe’s *idea units* (cf. Chafe, 1979).

alisation, cognitive constraints are also responsible for the difficulty of interpreting discontinuous constituents when they are very distant from one another. They also underlie the preference for ‘natural’ (iconic) linearization processes (such as Cause before Effect, Condition before Consequence, Earlier before Later in time) in spoken language.

2.2. Irreversibility

Transitoriness does not, however, cover all aspects of temporality, i.e. of the on-line character of spoken language. A second important aspect is the irreversibility of speech, which can also be viewed from both interactional and cognitive perspectives. From an interactional point of view, the simple yet consequential fact of spoken language is that which is said is said and cannot be undone. Of course, this does not mean that we cannot retrospectively try to revise the interpretation of what was said through our co-participants in the interaction (for instance, by contending that we did not mean what the co-participant claims to have understood), or to undo the particular speech act, i.e. to interactively erase it by repair or reformulation. Whether or not this is successful depends on the willingness of the co-participant to ‘delete’ this event from the record. Usually, our co-conversationalists are willing to allow repairs for harmless slips of the tongue, referential ambiguity, vague utterances etc., without their and our face being threatened, but will be more unwilling in cases of *faux pas*, ‘accidentally’ spilled secrets, insults, etc. Written language is of course completely irredeemable as soon as it is received; until then much time will have passed. During the process of writing, however, the writer (other than the speaker) is usually protected from such face-threatening accidents due to the privacy (and thus reversibility) of the writing process. He or she will be free to revise what was written, until the written message is dispatched.

The consequences of the irreversibility of speech for language processing are especially apparent when planning problems become visible through so-called editing phenomena such as hesitations, self-repairs and restarts. These editing phenomena either do not exist or are erased in written texts. (Here, too, new technologies have created intermediate forms however: e-mail messages often do not show such surface polishing/correction, although they are written.) The constraints of temporality for syntax also translate directly into well-known features of spoken syntax such as constituents that are produced later than they should be. Such ‘misplaced’ elements (according to written syntax) are especially common where syntax is governed by strict serialisation rules. They occur when some thematic or rhematic information needs to be added to the utterance beyond the slot where it ought to have been produced. Obvious examples come from spoken German with its constraints on argument placement (all non-clausal constituents have to be placed either in the position directly preceding the finite verb, or between the left verbal brace – the finite part of the verb –, and the right brace – its non-finite parts). Although standard German (written) syntax allows only few exceptions to this rule (such as relative clauses, see above), spoken German is full of ‘debraced’ materials that occur in the position *after* the right brace (cf. Auer, 1996). More generally, ‘afterthought’ positions seem to be a common feature of all languages (even strict SOV languages).⁶

2.3. Synchronization

A third angle on the specific temporality of spoken language significantly involves the relationship between speaker and hearer and the immediacy of the face-to-face situation. We have known since Albert Schutz that in direct communication as in no other form of interaction the stream of consciousness between the ego and the other is *synchronized* (cf. Schütz and Luckmann, 1975; also cf. Auer et al., 1999, p. 115ff). My speaking partner recognizes my own action with minimal delay by my behavior, and conversely I recognize his action by his behavior. Time of production and time of reception approach each other asymptotically. By means of this synchronization, emerging syntactic constructions of the speaker are processed with only a short delay by the recipient. By contrast, production and reception in written language are separated by a considerable distance of time. Moreover, writing and reading occur for most people at extremely divergent speeds, whereas speaking and hearing are relatively similar in tempo (Chafe, 1984). (Once more, as a result of modern

⁶ The recent debate about so-called increments has shown this for many languages to be true (cf. Couper-Kuhlen and Ono, in press).

technological media, intermediate forms have come into being that permit the delay in written communication to be minimized and even allow accessibility by the recipient to the on-going text production. Many forms such as e-talk, chat rooms, and so on belong in this category.)

The effects on syntax of this synchronization between speaker and hearer in face-to-face communication have become recognized through the incisive research of Goodwin (1979, 1980, 1995) and other conversation analysts during the last decade. It manifests itself in numerous more or less microscopic forms of the dialogic coordination between the activities of speaker and recipient, forms that in the last resort relativize the role of the speaker (Goodwin, 2007). The most obvious testimony consists of the collaborative construction of sentences by several participants in a conversation (cf. Lerner, 1996, 2004).

Taken together, these three perspectives on the temporality of spoken language – transitoriness, irreversibility, and synchronization – undoubtedly point to radical differences between speech and writing. Research on spoken language has often ignored these aspects and their ramifications; this is still reflected in the terminology we use (and for which we have no replacements) which is often explicitly based on the visual, written word. The most blatant examples are word pairs such as “left/right dislocation”, “left/right adjoined”, “left/right headed”, “left/right brace” etc., in which the two-dimensionality of the linguist’s piece of paper is the name-giving component.

3. Projection, expansion and retraction: the basic operations of on-line syntax

I would now like to introduce three basic syntactic operations. They by no means occur exclusively in oral speech, but do take on a very distinctive, typical characteristic in speaking and listening: projection, expansion and retraction. Through the use of syntactic *projections*, the speaker creates expectations in the listener about the further development of the emerging syntactic pattern; a syntactic gestalt is opened up, which in the end will be closed via the production of a more or less predictable element. The projection is then fulfilled. One reason for which syntactic projections are of great importance for oral language is that they (together with semantic and prosodic projections) enable the prediction of possible turn completion points.

Just like syntax itself, projections are always based on hierarchical structures in language. One of the most central projective procedures of syntax is therefore government. Other syntactic projection procedures operate on the basis of discontinuous constituents, such as the verbal brace in German. Small-scale projections seem to be ideal for spoken language: as a whole, they are easily processed since they do not overburden short-term memory, and are especially recipient-friendly. But there are also relatively far-reaching projection procedures found in spoken syntax. For instance, initial conditional clauses are frequently used in spoken syntax to project the apodosis, and between protasis and apodosis, many ‘parenthetical’ clauses may be inserted, such that the time span from one to the other is considerable (cf. Auer, 2000a; Ford and Thompson, 1986).

If we assume synchronization between the time of speaking and that of processing of the linguistic input by the listener, we imply that recipients do not wait until the end of a syntactic structure to begin processing it; rather, they will start from the first word they can identify to take into consideration an array of projectable continuations from which, over the trajectory of the sentence, more and more will be eliminated. Listeners thus first go through a phase of relative uncertainty in which a number of syntactic alternatives (and, accordingly, several syntactic interpretations of the first word(s)) are still ‘in play’ and simultaneously tracked. During the course of the speaker’s further production, this projective variety grows smaller, until towards the closure of the syntactic gestalt by the speaker, the final components of the syntactic unit will become more and more, and often entirely, predicable. (Early starts of next turn-taking speakers often orient themselves at this “recognition point”; cf. Jefferson, 1973, 1984; Wilson and Wilson, 2005.) Predicted courses of linguistic events can of course turn out to have been mistaken, in which case the recipient will have to revise his or her projections – sometimes even the entire syntactic analysis must be redone. However, as soon as a syntactic pattern is identified with reasonable certainty, the recipient is cognitively relieved; he/she now just has to check whether the projection actually fulfills itself. Similar conditions exist from the speaker’s point of view. In the beginning of his or her turn (or turn component), the speaker must build up new syntactic projections while assembling syntactic patterns in a phase of high cognitive stress. Later, over the course of time, the pattern chosen just needs to be lexicalised more or less automatically. However, upon closer examination there are

differences between the speaker and recipient points of view. In interactional terms, one of these differences is related to turn-taking: in certain contexts, the speaker can by all means have an interest in far-reaching structural projections which allow him/her to develop a statement without legitimate opportunities for intervention on the part of the recipient. In cognitive terms, the speaker is ahead of the recipient since s/he can chose a syntactic pattern and stick to it at a point where for the hearer, more alternative developments are still available and in need of being processed.

Let us look at an example. After a syntactic completion point, the following is uttered:⁷

- (1) (Extract from a job interview; the applicant has been talking about his holiday trip through North America, here about Toronto. The interviewer has the next turn.)

Interviewer: in der (–)
 in the

The interviewer starts out with a preposition followed by a definite article; he then hesitates. A German utterance occurring in the present context (i.e., in the beginning of a syntactic unit and a turn, and without contextual help, i.e. excluding an ‘elliptical’ reading), which consists of a preposition and an article, is not complete. Rather, the utterance projects a continuation, and this continuation is relatively well predictable: the hearer has all reason to expect a noun which is governed by the preposition (i.e., the head of the construction). And as a matter of fact, this noun follows (*Gegend*) after a retractive (see below) self-repair; it is additionally expanded (in a non-projected and therefore syntactically non-predictable way) by a genitival attribute (*von Toronto*) and a final particle which introduces some kind of vagueness (*da*):

Interviewer: in der (–)
 in the
 in der (–) gegend von toRONto da; (.)
 in the region of Toronto PART

Up to this point the hearer has at his disposal a sequence of words that can be parsed in an unambiguous way as a prepositional phrase. The falling intonation contour at the end of the utterance (indicated by the semicolon) and the utterance-final vagueness particle *da* achieve some kind of closure. But again, even though packaged as an intonation phrase of its own, this prepositional phrase is not sufficient to form a self-contained turn in the present context. More is projected. At this point, it is impossible to predict precisely how the emerging syntactic structure will continue; some possibilities can be excluded, but among the remaining, two alternative developments, and therefore two projections, continue to be ‘in play’.

Among the possibilities to be excluded is the projection of a finite verb (the left verbal brace) in the next position. The prepositional phrase would under this interpretation occupy the front field of a sentence, and a finite verb would be expected since the front field can only accommodate one constituent in German. This alternative is excluded both on prosodic grounds (the front field constituent, unless it is a clause, rarely is an intonational phrase of its own), and on syntactic grounds: after the final particle *da*, no such continuation is possible: thus, while

in der (–) gegend von toRONto →
in the region of Toronto
leben wenig Frankophone.
live few French speakers

⁷ A more elaborate example is discussed in Auer (in press).

would be possible, resulting in the sentence ‘in the area of Toronto live few French speakers’ (with the verb *leben* governing a subject, *wenig Frankophone*, and a location, *in der Gegend von Toronto*),

in der (–) gegend von toRONto da; →
 in the region of Toronto
 leben wenig Frankophone.
live few French speakers

is not.

But other ways to continue after the prepositional phrase can be projected. The prepositional phrase may be left-dislocated (proleptic) with an uptake (*da* ‘there’) in the following slot, and the finite verb after that:

in der (–) gegend von toRONto da; (.) →
 in the region of Toronto PART
 da leben wenig Frankophone.
there live few French speakers.

(Note that the uptake-*da*, the correlate of the prepositional phrase, is different in function from the vagueness marker *da* closing off the prepositional phrase.) The prepositional phrase may also turn out to be a ‘hanging topic’, the expected continuation being a full sentence the internal makeup of which is open, e.g.:

in der (–) gegend von toRONto da (.)
 in the region of Toronto PART
 → wenn Sie da aufs Land fahren, können Sie noch viel
 Deutsch hören.

*if you travel in the countryside, you will still hear a
 lot of German*

Arguably, hearers will treat alternative projections such as prolepsis and hanging topic according to their likelihood, i.e. on the basis of their previous experience with turn-beginnings of this type; this will make it possible for them to weight the alternatives according to their frequency of occurrence. In such a frequency-based calculation of likelihood, the second projection is probably less relevant than the first.⁸ And in the present case, such a calculus would also lead to the right result, for the actual continuation of the emerging syntactic unit is as follows:

in der (–) gegend von toRONto da (.) →
 in the region of Toronto PART
 da: (1.0) .h leben ja noch ne ganze REIhe dieser; (–) Amish people
there (1.0) .h still live a whole lot of these; (–) Amish people

Of the two projections, only one (prolepsis) remains in play, as soon as the local adverbial *da*: has been produced. At this point, the speaker hesitates again (one second silence plus a subsequent hearable in-breath). The hearer can expect the left verbal brace element of the German declarative sentence now; and this projection is in fact fulfilled with the production of the finite verb *leben* (1st or 3rd person of the plural). According to the case frame of the verb, the subject of the sentence is now expectable, and delivered in the form of a complex noun phrase (*ne ganze Reihe dieser Amish People*). This noun phrase has an internal hierarchical structure, which makes additional projections possible on smaller scale. In particular, *ne ganze Reihe* ‘a whole a whole array/lot’, although being a full noun phrase, projects a genitive attribute.

So far, the interviewer’s utterance has been analysed in terms of projecting and projection-fulfilling (projected) elements in the emerging syntactic gestalt. It is obvious that this is not sufficient, since there are other elements which cannot be classified as such. One group of syntactic items which do not fall under either category I call *expansions*. We have already seen that the genitive attribute *von Toronto* modifying *in der Gegend* is one of those expansions. Other examples of elements which neither project nor fulfil projections occur in the

⁸ More on oral constructions at the left periphery of German sentences can be found in Selting (1993), Scheutz (1997), and Auer (1997).

middle field of our example, i.e. the modal particle *ja* (suggesting that the information contained in this sentence should be known) and the adverb *noch* ('still'):

da: (1.0) .h leben ja noch ne ganze REIhe dieser; (–) Amish people

We can now have a look at the complete utterance of the interviewer. What he actually says is this

(1) (complete version)

Interviewer: (a) in der (–) gegend von toRONto da (.)
 in the region of Toronto PART
 (b) bissl wEstlich davon,
 a litte to-the-west of-it
 (c) =also in Waterloo
 that-is in Waterloo
 (d) =(w)o (–) .h auch ne (–) beKANNte universITÄT isch (–)
 where also a well-known university is
 (e) da: (1.0) .h leben ja noch ne ganze REIhe dieser; (–)
 Amish people
 there (1.0) .h still live a whole lot of these; (–)
 Amish people

Between the projecting utterance component (a), and the projected continuation in (e), three other components occur within the projection span (a) → (e). The last of these components (d), the relative clause *wo auch ne bekannte Universität isch*, expands the noun phrase *Waterloo* internally without fulfilling a projection (and without creating its own projections). It therefore is another example of an (internal) expansion. But what about the others?

At this point, we need to introduce a third type of basic operation of spoken syntax which I refer to as *retractions*. Like expansions, they are attached to existing structures without being projected by them, but they do not add syntactic structure. Rather, retractions refer back in time to already existing syntactic structures which they reactivate and change. In example (1)', this applies to components (b) and (c): in a context in which the prepositional phrase *in der Gegend von Toronto da* is already available, both *bissl westlich davon* and *in Waterloo, wo auch ne bekannte Universität isch* retract to the slot in which this prepositional phrase has occurred, in order to insert into the same slot a different, but syntactically equivalent constituent. (The retraction is marked by the discourse marker *also* in the case of (c).) What on the semantic level can be described as an elaboration of the description of the locality, is on the syntactic level a retraction to a previous slot in the emerging syntactic gestalt.

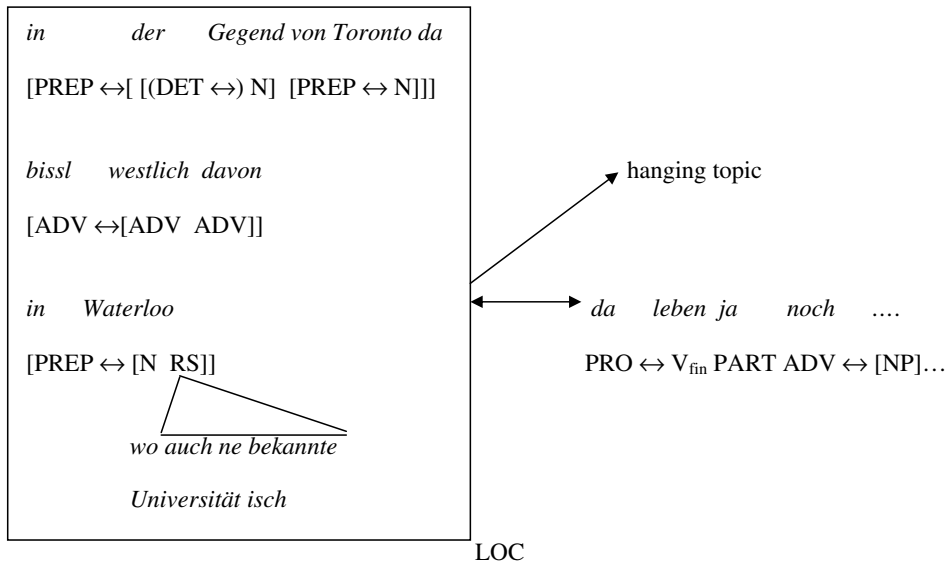
In sum, I claim that any syntactic structure in spoken language at any point during utterance can be categorized according to:

- whether or not it modifies an already-existing construction by retracting to a slot in it; if this is not the case,
- whether the structure is projection-building or projection-fulfilling (projected).⁹ Since one constituent may be both projected and projecting, one must distinguish between maximal projecting structures (i.e. those which build a new syntactic structure from the start), maximal fulfilling structures (i.e. those which are non-projecting) and those structures which fall somewhere in between.

⁹ The formulation leaves it open whether a retraction (repair) can be projected as well. There is some evidence that upcoming repair work can be projected. For instance, the discourse marker *da*, which rounds off line (a) of example (1), introduces a certain amount of vagueness in the referential description 'in the region of Toronto'. It thereby foreshadows additional conversational work which aims at making the description tighter and interactionally more successful. In this sense, it can be said to project the following repair work. However, this kind of projection is on the interactional level only; the exact syntactic format of the repair is not projected, and in fact, it could be done in a non-retractive way.

- Constituents that are neither projection-building nor projection-fulfilling, but not retracting either, are either internally or externally expanding (within syntactic constructions or beyond syntactic completion points, respectively).¹⁰

A visual illustration of the projection and retraction operations in the beginning of the utterance above might now look something like this:



The arrows mark projections (to the right) and projection fulfillments (to the left). Smaller constituents are only partly labelled, and the strength of the projections is not marked. Alternative projections are only indicated as discussed in the text above. As noted, the projection starting with the constituents *in der Gegend von Toronto da, bissl westlich davon* and *also in Waterloo, wo auch ne bekannte Universität ist* is underdetermined in the present context. Only one of the two possibilities finally becomes relevant (double arrow), while the other is not realised (single arrows pointing to the right): As soon as the resumptive locative pronoun *da* is produced, the already-uttered structure becomes that of a left dislocation. Now only one further development is possible – namely the ensuing finite verb, i.e. *da* both is projected and projects. The same applies to the following finite verb *leben*, which likewise cannot close the syntactic structure on its own (the verb requires a subject). The particle *ja* and the free temporal adverb *noch* are neither projecting nor projection-fulfilling (no arrow), but expanding. The terminating/concluding complex noun phrase *ne ganze Reihe dieser Amish People* fulfills a projection without being a projecting itself; it thus completes the syntactic structure. Retractive syntactic operations are indicated in the diagram by listing constituents in their paradigmatic substitution-relationship above and below one another in the box.

4. Some typical phenomena of spoken syntax

Some further examples may illustrate how projection and retraction work together in typical spoken constructions, and how the approach sketched in the last paragraph can explain the occurrence of these constructions.

¹⁰ External expansions are discussed in more detail in Auer (in press).

4.1. Contaminations

Any valid approach to the temporality of spoken syntax must take into account the structurally awkward but nevertheless everyday fact that the syntactic status of a particular constituent can change during its production. An example of one such a change in construction (often called a contamination) is the following:

(2) [group therapy session, M is one of the clients]

M: <<p>kann ich dich total gut nach (-) NACHvollziehen;> (5.0)

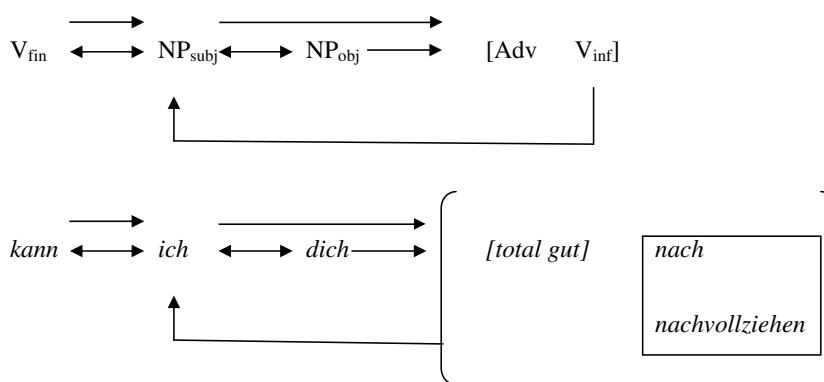
The speaker wants to say that she can perfectly well understand the person she is addressing. There are two verbs in German equivalent to English *to understand* in this sense. One of them, *verstehen*, requires a personal or non-personal object (someone/something who/that is understood), the other one, *nachvollziehen*, requires a non-personal object (referring to the thing or propositional content understood). The speaker starts out with a personal noun phrase in the accusative (*dich*) which, in combination with the formulaic, emphatic particle *total gut* ('totally well') projects the verb *verstehen*, but excludes *nachvollziehen*. However, the speaker ends up with the verb *nachvollziehen* in the right sentence brace (non-finite form, infinitive). The verb and the personal pronoun in the position of the direct object are not compatible. The two constructions that are merged are thus the following:

(da)¹¹ kann ich dich total gut verstehen
(there) can I you_{ACC} totally well understand

and

(das) kann ich total gut nachvollziehen
(this) can I totally well comprehend.

It is easy to explain such a contamination in terms of the projection approach to on-line syntax: it can be understood as a clash between the structure that is projected and the item that fulfills the projection. Using the symbols of the diagram in Section 3, the picture is the following:



¹¹ Note that the utterance starts with the finite verb (*kann*), which in written standard German is impossible in a declarative main clause: the front field usually must be filled by (exactly) one constituent and cannot remain empty. In spoken German, however, this constraint is lifted in certain contexts (cf. Auer, 1993). Since the necessary front field constituent would be a different in the version with *verstehen* than from the version with *nachvollziehen*, the possibility of verb-initial sentences in spoken German is a prerequisite for the kind of contamination we are discussing here.

The finite verb *kann* which begins this syntactic structure allows a number of different projections; in addition to the possibility of a subject pronoun immediately following in the middle field (which is confirmed by the actual development), there are other possible projections according to which an object pronoun (*kann das total gut verstehen*) or a second-person pronoun (*kann dich total gut verstehen*) follow. However, after two elements of the emerging structure are available (*kann ich*), the possibilities are narrowed down; now there is a choice between a non-finite verb (such as *nachvollziehen*) or a personal object (*kann ich dich...*) only.

The actually uttered third element (*dich*) clearly selects this last option so that there is only one way of continuing left, namely the infinitive of a verb which requires a personal subject (*ich*) and direct object (*dich*). The verb *nachvollziehen*, which is now produced in two steps (retraction), does not fulfill this projection, but rather (along with *total gut*, which as a non-projected pivot is compatible with both the projecting and delivered continuation) reactivates a different, previously eliminated structure. For this reason, there is no arrow from *nachvollziehen* pointing back to *total gut*. Instead, the chosen verb joins with the adverb *total gut* in order to attach itself to an already-eliminated structural development. In this way, the constituent *dich* is retroactively cancelled from the emerging syntagma.

4.2. Apo-koinu constructions

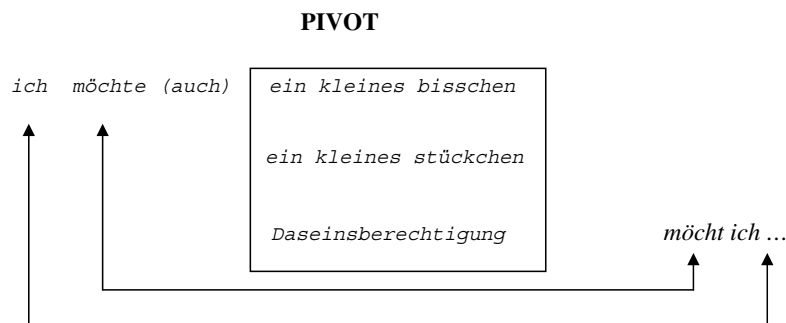
Apo-koinu or pivot constructions are another typical structure of spoken syntax (cf. Scheutz, 2005 for details). They are similar to contaminations in that one element of the emerging syntactic structure changes its status over the course of production and reception time:

(3) [extract from phone-in show/therapeutic context]

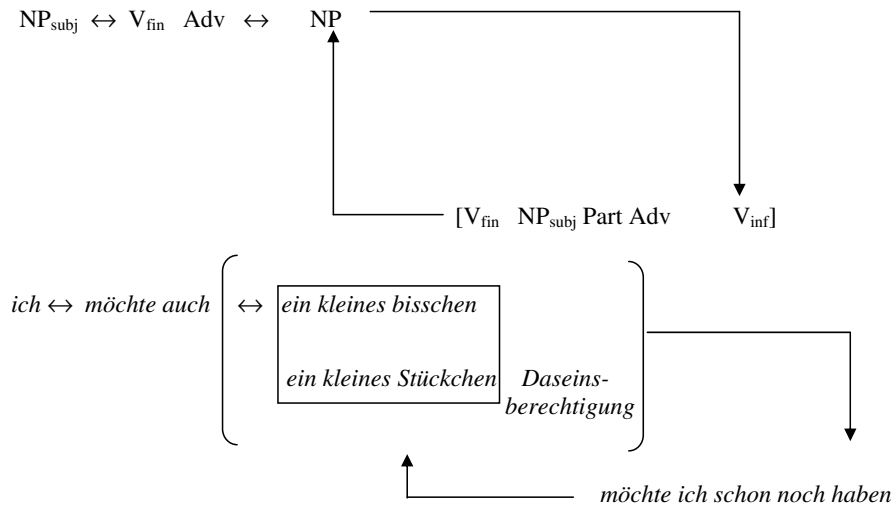
Caller:	und	ich	möchte	auch	ein	kleines	bisschen
	<i>and</i>	<i>I</i>	<i>want</i>	<i>also</i>	<i>a</i>	<i>little</i>	<i>bit</i>
	ein	KLEInes	stückchen	DAseinsberechtigung			
	<i>a</i>	<i>little</i>	<i>bit (of)</i>	<i>justification-to-exist</i>			
	möcht	ich	SCHON	noch	habn;		
	<i>want</i>	<i>I</i>	<i>PART</i>	<i>still</i>	<i>have</i>		

‘and i also want a little bit a little bit of a justification to exist i would also like to have’

What we observe here is a typical ‘mirror construction’: the subject and finite verb preceding the koinon (pivot), i.e. *ein kleines bisschen ein kleines Stückchen Daseinsberechtigung*, are repeated after the koinon in inverted (i.e. mirror-image) order:



The graphical representation in terms of projecting and projection-fulfilling elements – which is somewhat simplified with regard to the alternative projections here – shows this pattern:



Instead of completing the entire construction *ich möchte auch ein kleines Stückchen Daseinsberechtigung* with the only structural part lacking, and highly predictable, i.e. the non-finite (infinitival) verb form *haben*, the actual subsequent structure *möchte ich schon noch haben* matches with one of the possible projection of the NP *ein kleines bisschen Daseinsberechtigung* taken on its own, i.e. detached from the previous utterance part. It retrospectively deletes large parts of this utterance which do not fit into the pattern, since *möchte ich schon noch haben* can only retrospectively accommodate one additional element to form a sentence – a constituent in the front field (the pivot now parsed as the beginning of the sentence before *möchte*).

4.3. Self-repair

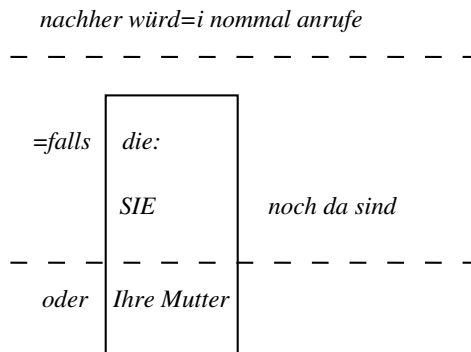
The previous examples have primarily demonstrated ‘fuzzy’ relationships between projections and projection-fulfillments, as is typical for spoken language. Another typical characteristics of spoken syntax is the high amount of retractions which play an especially important role in self-repair. A final example can show this.

(4): [telephone conversation between a telephone technician and the daughter of a client]

Technician: nachher würd=i nommal anrufe=falls die: (.) SIE noch da sind, oder ihre mUtter
 afterwards would I again call if the: (.) you still there are, or your mother

A technician tells the daughter of the phone-owner that he will call again, and he wants to make sure that either the daughter or her mother will be at home. The conditional clause introduced by *falls* ‘if’ leads into trouble in the formulation of the subject noun phrase which follows the subjunction *falls*. This is indicated by the elongation of *die*: (‘the’), a break-off, a retraction to the beginning of the noun phrase, and the replacement of *die*: (‘the’) by *Sie* (‘you’). A very strong structural projection which had been achieved by the determiner – a following noun, resulting in a noun phrase – is thereby eliminated. But in spoken language, that which is said is said: we (just like the daughter) know that the technician was probably about to say *die Mutter* (‘the/your mother’).¹² This in turn needs further repair. After the subordinated clause is brought to completion, the speaker goes back to the trouble spot (in the following diagram, dotted lines indicate syntactic closure, boxes retractions):

¹² In colloquial, particularly southern German, reference to family members is often done with the definite article instead of the (standard German) possessive form *Ihre Mutter*.



Retrospectively, an element of the already-complete construction (*Sie*) is paradigmatically replaced by another (*Ihre Mutter*), long after the appropriate slot for the subject NP has gone. The result is a twofold retraction within and across a projection span: *die: ... → Sie → Ihre Mutter*.

5. Concluding remarks

The observations in the preceding paragraphs are intended as a first attempt at the integration of temporality and the investigation of spoken syntax. Central terms in this area are projection, expansion and retraction. Typical differences between spoken and written language arise from the applications of these operations, especially the various types of non-equivalents between projections and projection-fulfillments, and from the manner and frequency of retractive repair in face-to-face interaction.

An approach to syntax which is based on operations such as projection and retraction will find theoretical support not so much in structural or post-structuralist (generative) syntax research, but rather in the semiotics of Ch.S. Peirce (1960–66): the retractive strategies are an example of iconism, the projective strategies a case of indexicality. It is almost trivial to point out that a structuralist definition of the sentence is not compatible with the on-line model of syntax processing. The structuralist segmentation of a text into sentences is carried out not from a real-time emergence perspective, but rather from the bird's eye view of the text editor, i.e. the finished product.

References

- Auer, P., 1993. Zur Verbspitzenstellung im Gesprochenen Deutsch. *Deutsche Sprache* 3, 193–222.
- Auer, P., 1996. On the prosody and syntax of turn-continuations. In: Couper-Kuhlen, E., Selting, M. (Eds.), *Prosody in Conversation*. Cambridge University Press, Cambridge, pp. 57–100.
- Auer, P., 1997. Formen und Funktionen der Vor – Vorfeldbesetzung im Gesprochenen Deutsch. In: Schlobinski, P. (Ed.), *Syntax des gesprochenen Deutsch*. Westdeutscher Verlag, pp. 55–92.
- Auer, P., 2000a. Pre- and Postpositioning of *wenn*-clauses in spoken and written German. In: Couper-Kuhlen, E., Kortmann, Bernd (Eds.), *Cause – Condition – Concession – Contrast. Cognitive and Discourse Perspectives*. Berlin: de Gruyter, pp. 173–204.
- Auer, P., 2000b. Schreiben in der Hypotaxe – Sprechen in der Parataxe? Kritische Bemerkungen zu einem Gemeinplatz. *Deutsch als Fremdsprache* 39 (3), 131–137.
- Auer, P., 2005. Projection in interaction and projection in grammar. *Text* 25 (1), 7–36.
- Auer, P., Couper-Kuhlen, E., Müller, F., 1999. *Language in Time*. Oxford University Press, New York.
- Auer, P., in press. Why are increments such elusive objects? An afterthought. In: Couper-Kuhlen, E., Ono, Y. (Eds.), *Increments in Conversation* (Special Issue of *Pragmatics*).
- Behaghel, O., 1932. *Deutsche Syntax: Eine geschichtliche Darstellung*. Bd. 4: Wortstellung – Periodenbau. Heidelberg: Winter.
- Bergmann, J., Luckmann, T., 1995. Reconstructive genres of everyday communication. In: Quasthoff, U. (Ed.), *Aspects of Oral communication*. de Gruyter, Berlin, pp. 289–304.
- Chafe, W., 1979. The flow of thought and the flow of language. In: Givón, T. (Ed.), *Syntax and semantics*. Academic Press, New York, pp. 159–181.
- Chafe, Wallace, 1984. Integration and involvement in speaking, writing, and oral literature. In: Tannen, D. (Ed.), *Spoken and Written Language*. Ablex., Norwood NJ.
- Couper-Kuhlen, E., Ono, Yoshida (Eds.), in press. *Increments in conversation* (Special Issue of “Pragmatics”).

- Ford, C., Thompson, S., 1986. Conditionals in discourse: a text-based study from English. In: Traugott, E. (Ed.), *On Conditionals*. UP, Cambridge, pp. 353–372.
- Goodwin, C., 1979. The interactive construction of a sentence in natural conversation. In: G. Psathas (Eds.), *Everyday Language*, New York, pp. 97–121.
- Goodwin, C., 1980. Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Soc. Inquiry*, 272–302.
- Goodwin, C., 1995. Sentence construction within interaction. In: Quasthoff, U. (Ed.), *Aspects of oral communication*. de Gruyter, Berlin, pp. 198–219.
- Goodwin, C., 2007. Interactive Footing. In: Holt, E., Clift, Rebecca (Eds.), *Reporting Talk: Reported Speech in Interaction*. Cambridge University Press, Cambridge, pp. 16–46.
- Hockett, C.F., 1960. The origin of speech. *Scientific American*. 203 (September).
- Hopper, P.J., 1989. Times of the sign lectures on discourse and temporality in recent linguistics. Humanities Institute, Santa Barbara, Lecture Series on Human Temporality, 1989–1990, October 16–21, 1989 (online available from: <http://langs.eserver.org/times-of-the-sign.txt>).
- Jefferson, G., 1973. A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences. *Semiotica* 9, 47–96.
- Jefferson, G., 1984. Notes on some orderliness of overlap onset. In: D’Urso, V., Leonardi, P. (Eds.), *Discourse Analysis and Natural Rhetorics*. CLEUP, Padova, pp. 11–38.
- Lerner, G., 1996. On the ‘semi-permeable’ character of grammatical units in conversation. Conditional entry into the turn space of another speaker. In: Ochs, E., Schegloff, E., Thompson, S. (Eds.), *Interaction and Grammar*. Cambridge University Press, Cambridge, pp. 238–276.
- Lerner, G., 2004. Collaborative turn sequences. In: Lerner, G.H. (Ed.), *Conversation Analysis: Studies from the First Generation*. Benjamins, Amsterdam, pp. 225–256.
- Pawley, A., Syder, F.H., 1983. Natural selection in syntax. Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics* 7, 551–579.
- Peirce, Ch.S., 1960–66. In: Burks, A. (Ed.), *The Collected Papers of Charles Sanders Peirce*, vol. 8. Harvard University Press, Cambridge.
- de Saussure, F., 1974. *Cours de linguistique générale* (critical edition by de Mauro, T). Paris: Payot.
- Scheutz, H., 1997. Satzinitiale Voranstellungen im gesprochenen Deutsch als Mittel der Themensteuerung und Referenzkonstitution. In: Schlobinski, P. (Ed.), *Syntax des Gesprochenen Deutsch*. Westdeutscher Verlag, Opladen, pp. 27–54.
- Scheutz, H., 2005. Pivot constructions in spoken German. In: Hakulinen, A., Selting, M. (Eds.), *Syntax and Lexis in Conversation*. Benjamins, Amsterdam, pp. 103–128.
- Schütz, A., Luckmann, T., 1975. *Strukturen der Lebenswelt*. Luchterhand, Neuwied.
- Selting, M., 1993. Voranstellungen vor den Satz. Zur grammatischen Form und interaktiven Funktion von Linksversetzungen und Freies Thema im Deutschen. *Zeitschrift für Germanistische Linguistik* 93 (21), 291–319.
- Silverstein, M., 1984. On the pragmatic ‘poetry’ of prose: parallelism, repetition, and cohesive structure in the time course of dyadic conversation. In: Schiffrin, D. (Ed.), *Meaning, Form and Use in Context*. Georgetown University Press, Georgetown, pp. 181–199.
- Wilson, M., Wilson, T.P., 2005. An oscillator model of the timing of turn-taking. *Psychonomic Bulletin and Review* 12 (6), 957–968.