The dynamics of incrementation in utterance-building

Processes and resources

Per Linell
Göteborg University and Linköping University

This article discusses processes and practices in utterance-building in an interactive perspective. It argues for a model of on-line incrementation in which projections, and sometimes their suspensions, play a crucial role. The status of units of language within such a theory of languaging is discussed. Data are drawn from authentic conversational interactions in Swedish, with a special focus on evidence from pivot utterances, deviations from normative grammar and slips of the tongue.

Keywords: Utterance production; dialogue; incrementation; interaction; projection; units of conversational language; Swedish

1. Units in a dialogical and interactional grammar

In this contribution I shall deal with how speakers build utterances in real conversational interaction. It will be argued that utterances are successively built up bit-by-bit in a process in which the speaker incrementally adds to what (s)he, or other participants, have already produced (Hopper 2011, 23). Usually this is done in such a way that the speaker fulfils syntactic and semantic-pragmatic expectations that have been built up, that is, projected, in the prior talk. At the same time, new projections are

1. This paper is based on interactional-linguistic research performed on Swedish data (cf. Linell 2011). The research was facilitated by a grant from the Swedish Research Council (421-2004-1087). This particular text is a considerably revised version of a paper read in the panel on “Units in interaction” organised by Beatrice Szczepak Reed at ICCA in Mannheim in July 2010. Participants there, and at a seminar in Turku in October 2010, made important comments. I am grateful to FRIAS, Freiburg, for hosting me while I was revising this paper in May 2012. I gratefully acknowledge extensive comments by anonymous reviewers, as well as by Elizabeth Couper-Kuhlen, Christine Mertzluft, Kerstin Norén and Sandra Thompson.
of course generated in the unfolding syntactic project. This takes place in a dynamic interplay both with contributions and responses by other participants and with the speaker’s own thoughts and impulses popping up in the course of production. Such a process has come to be called ‘incrementation’ (Schegloff 1996). ‘Increment’ means etymologically that a separate piece is added to what we already have. However, the terms increment and incrementation have been used both in a more general sense, i.e. how utterances are successively and dynamically produced, and in a more specific sense (Schegloff, op.cit.), in which ‘increment’ refers to a ‘turn increment’, when a speaker adds a segment to an already potentially complete turn or turn constructional unit.\(^2\) I shall use the term in its more comprehensive sense.

If utterances are built incrementally, piece-by-piece, these pieces may be seen as ‘units’, more specifically units with syntactic interrelations to other units, which leads to the question of what types of such units we might want to recognise. This must be seen in relation to an interactional approach; a speaker’s utterances are built in interaction with other participants. These utterances usually somehow fulfil action projections from prior utterances (or local situations), and they project further upcoming actions and utterances by others. However, utterances and turns also have a dynamic nature internally, that of turn-internal incrementation and projection. It is this interactional grammar within speakers’ utterances and turns, rather than the responsive and projective relations between participants’ contributions to the interpersonal interaction, that I will focus on in this paper. Therefore, while some sections (in particular Section 5) will deal with interaction across speaker-turns, my point is that a speaker’s turn or contribution to sociodialogue is itself internally dialogical, that is, it comprises responsive and projective relations between parts of turns or utterances (Linell 2009). In other words, I start out from actions, but deal mostly with units and processes of talk.

2. On-line syntax

A grammar for interaction has to be different from a grammar of language structure that ignores the on-line processes of utterance-building. Schegloff, Ochs and Thompson (1996, 24) argue that “the interactional matrix of grammar requires a different understanding of what should enter into a linguistic description and/or a different model of linguistic structure.” Similarly, Peter Auer points out than an “on-line syntax” for

\(^2\) There is an increasing literature on the incremental production, including Ford and Thompson (1996), Auer (2007a, b, 2009a, b), Couper-Kuhlen and Ono (2007), Walker (2004), Imo (2011), and others. Here I will try to contribute a few more variations on the theme.
spoken language must be “a significantly untraditional approach” (2009a: 1). In particular, it must bring the temporal dimension into the grammar of talk-in-interaction.

Our “significantly untraditional approach” must be different from formalist theoretical grammar, including many generativist variants. The latter are characterised by, among other things, the following (Linell 2009):

- Formalist grammar treats language as a set of formal objects, rather than as actions and processes embedded in temporally unfolding communicative and cognitive projects. It belongs to a “written-language-biased” tradition of linguistics (Linell 2005) that represents language in a “deactivated form” (Hopper 2011, 42), and misses out on the temporality of languaging (Thibault 2011);
- It is overly abstract and overtechnicalised, while at the same time refraining from exploring real spoken-interactional language systematically and empirically. It develops a kind of “theory” that does not contribute to the explanation of languaging (“language use”), and may not correspond to anything in the functioning of the brain, let alone the mind (that is, the sense-making processes extending far beyond the boundaries of the brain; Cowley 2011a);
- It is self-contained, with no natural links to context or to alternative or complementary semiotic resources (Enfield 2009; Goodwin 2000).

Dynamics and dialogue are more basic than language and structure. In this paper, however, the focus will be on language, as conceived of within a theory of languaging (cf. n. 5). The syntax of conversational language (and languaging) calls for a processual approach. I will try out a thumbnail sketch of processes and resources in interaction, for example, construction methods, decision points and continuation types within a theory of spoken-interactional languaging. Utterance-building is dynamic and processual, but also planned. However, planning is overwhelmingly partial and local, as we will see below, and processes in impromptu talk often override structural exigences of normative language.

3. Units and elements

Sentence grammar, which was, as we have noted, based on standard(ising) written language, has a theoretical vocabulary of grammatical units, such as lexical categories (noun (N), verb (V), adverb (Adv), pronoun (Pro), preposition (Prep), etc.), phrasal

---

3. I will prefer the term ‘languaging’ (e.g. Thibault 2011) over ‘language use’, since the latter suggests that language structure has priority over processes in “using” language and other semiotic resources (Linell 2009).
categories (noun phrase (NP), prepositional phrase (PP), sentence or clause (S), verb+argument frame), and syntactic functions like grammatical subject and object, etc. The foundational paper of Conversation Analysis by Sacks, Schegloff and Jefferson (1974) similarly employed turn-constructional units of sentential, clausal, phrasal, and lexical types (720).

Interational grammar proposes a set of other kinds of units that can fill different turn-constructional fields: pre-front (‘initial periphery’), front, core, end and post-end (‘final periphery’) fields of turns and TCU's (Schegloff 1996; Auer 1996; Steensig 2001; Lindström 2008). Schegloff, in his paper which explicitly deals with the relation between interaction and grammar (1996), mentions also “elements” like, for example, parentheticals, try-markings, prosodies, gestures, breaths, laugh tokens, “recognizable contexted-silences”, cut-offs, sound stretches, and many more (103–4). At another level, we must talk about “contributions” to discourse of various sizes (Houtkoop & Mazeland 1985). These are relational units with responsive and initiatory properties (Linell 2009). Construction Grammar (Croft & Cruse 2004) argues that we need to work with many more construction types than merely assertives, interrogatives, exclamatives, etc., and other “speech-act-based” types (Searle 1969, 1975). Examples of such other construction types are responsive constructions, such as clefts, reactivives, elliptics, “dense constructions” (Günthner 2006a), and many others.

However, an entifying approach built exclusively on “units” or “elements” tends to portray utterances as being “built” with the use of building-blocks, and structures as “consisting of” constituents. This appears to rest on a house-building metaphor; speakers are assumed to build up utterances from the bottom up, but perhaps with an overall plan of the whole architecture to be implemented. This runs the risk of losing the perspective of agents participating in progressive (inter)action and activities. It diminishes, or even ignores, the role of actions and interactions; units and products overshadow processes. Even among conversation analysts, who share a conviction that the nature of spoken discourse is fundamentally interactional, that is, action-based, there is much talk about turn-constructional units and elements (Schegloff 1996). Is this radical enough? Doesn’t it still smell of a product- rather than process-oriented approach?

Generative linguistics added the notion of ‘rule’ to the basic theoretical vocabulary, sometimes calling these rules ‘processes’. This descriptive innovation was foreshadowed in American structuralism; although the terms were first applied to morphology rather than syntax, ‘item-and-process’ was launched as opposed to the

---

4. It should be admitted that these terms seem to work best for continental Germanic languages (Scandinavian, German, Dutch) and less well for many other languages, including Japanese, Chinese, Finnish, English and Romance language.
more static-sounding ‘item-and-arrangement’ (Hockett 1954). It must be pointed out that this terminology is misleading, in so far as the term ‘processes’ is used about ‘rules’ like left/right dislocation, fronting, subject raising, NP deletion, pronoun dropping, ellipsis, etc. These are simply ways of formulating static relations within the language system (Auer 2009a; 4); they do not involve any real ‘movements’ (cf. dislocation, fronting, lifting), or ‘deletions’ (implying that units have first been present at the location involved, and then removed). By contrast, when we talk about processes and resources within on-line syntax, we refer to constraints on real utterances, or better: (language-specific) methods of utterance-building. Such methods, in particular, grammatical construction methods, should reflect the temporal nature of real utterance-building processes; they are ways of designing utterances on-line. We are concerned with temporality, rather than spatial arrangements. Thus, we should prefer to talk about initial and final, rather than ‘left and right’, peripheries.

4. Interdependence of structures and processes

In the late 19th century, there was a scholarly debate on the “psychology of the sentence” (Blumenthal 1970). Wilhelm Wundt (1900) advocated the theory that language use (in thinking and, by implication, speech) is based on whole sentences, which are hierarchically structured wholes (Wundt’s term was Gesamtvorstellung, ‘comprehensive representation (of content)’). This theory has been legion in logic (philosophy of language) and theoretical linguistics up to this day (e.g. Levelt 1989). Yet, we know, from extensive empirical studies of many spoken languages, that sentences are hardly the only significant units of syntax in interaction; indeed, we cannot build a theory...
of conversational syntax on the ‘sentence’ as the foundational theoretical notion (e.g. Selting 2001; Laury 2009).7

Wundt’s ideas had been challenged already by Hermann Paul (1886), who argued for (what is nowadays often called) an incremental theory of utterances; words are added to words, and this piece-by-piece process finally ends up in the form of sentences. Of course, it is clear that in this process the basic constituent acts involve uttering units, adding new units as the speaker continues from point to point in interaction in-real-time (IRT). However, the units, bits and pieces that are uttered or added are not autonomous atoms but are subject to syntactic interdependencies. This gives rise to projections (Auer 2005, 2009a, b). Put differently, a realistic theory of incrementation cannot mean that utterance production amounts to a concatenation of mutually autonomous words. In some ways this seems to result in a compromise between Wundt and Paul (at least as they were portrayed in Blumenthal’s undoubtedly simplified account).

So, units are relational and structural entities; they generate and fulfil projections that are due to syntactic dependencies. Units occurring at one location in the emergent utterance anticipate other later upcoming units and structures. They foreshadow “more to come” (Auer 2005). In this way, they are a support for the speaker in keeping a consistent line in utterance-building. They are also a resource for listeners allowing them to anticipate the upcoming parts and perhaps even to take over the completion of the utterance-in-progress.

But, as Auer (2009a) argues, the dynamics of projection – the creation and fulfilling of projections – is not omnipotent; sometimes, projection constraints are suspended in ‘retractions and expansions’. In retractions, projections are suspended (or ‘deferred’; Günthner 2011, 173) by the utterance process being backtracked to a prior point, and a new unit of (paradigmatically) the same kind is inserted; in other words, the same structure is ‘recycled with variation’ (Anward 2005). Other retractions (cf. Birkner et al. 2012) involve replacements (Couper-Kuhlen & Ono 2007) as in restarts, anacolutha, and other self-repairs. In inserted expansions, by contrast, projections are momentarily suspended without retraction, and non-projected extra material is inserted. In addition, I would propose the category of redirections, as in

7. The ‘sentences only’ theory was bolstered by the idea that (proper) language structure reflects, or should reflect, (proper) thinking. Hughlings Jackson, a famous aphasiologist, claimed that “to think is to propositionalize” (1878 [1958]). However, that can not hold true for thinking as expressed in speaking or, say, electronic languaging (SMS, tweeting, etc.) Indeed, a growing body of evidence supports the view that “propositionalisation”, i.e. expressing oneself in full sentences (and combinations of sentences and “sentence derivatives” like abstract verb-based NPs), is true only of standard written language in certain prose genres. See also below on pseudo-clefts (Section 6).
pivot constructions (see Section 8), in which a syntactic structure is abandoned and/or continued into another one, without nullifying the content of the first part.

If projections, retractions, expansions, replacements and redirections are basic process types, they are interdependent with unit types. They emerge, by definition, from predictions of upcoming structures (anticipations, prolepsis, cataphora), defined in terms of dependencies primarily within NPs, PPs, verb+argument frames, and clauses. Retractions, expansions and redirections lift, also by definition, projection constraints (see Auer 2009a). Replacements simply render projection fulfillments void. In this sense, projection is basic, as it underlies the definitions of the other process types, which are suspensions or replacements of it.

A theory that seems compatible with projection theory but looks more structurally based, is that of frames, slots and fillers (Anward 2005). By starting with a frame-building part, a slot is created in which a certain kind of filler must be fitted (and can be anticipated, or projected, before being uttered). For example, in German, a preposition like *auf* projects a noun phrase in the dative or accusative, and after *auf dem* (‘on the’), there will be a projection of an (optional) adjective phrase followed by a masculine or neutral noun in the singular, say, *(roten)* Dach(e) (‘red roof’). In Section 9 I will deal with some kinds of “failures” to fill in such frames as projected.

5. Units, decision points, continuation types

If we build up utterances incrementally, by adding more units onto what has already been produced, this process involves passing decision points (transition points, points for optional redirection) of varying importance. At some points, speakers have more options than at other points. In general, major boundaries and new beginnings involve a wider range of options (possible redirections). Before an utterance beginning, there is often a wide range of options. Yet, the speaker may not be free to use any kind of construction, since there are often turn- or utterance-external projections; after some types of prior contributions, there are clearly projections for the next contribution (the new utterance, cf. ‘conditional relevance’ in CA). For instance, after a specific type of interrogative (e.g. a Wh-question), there are clearly fewer preferred ways of responding than after an open declarative question.

8. Sequentiality (in real time) does not mean that all components are linearly ordered in relation to all other components; there is also simultaneity. This applies to prosody, relative to regular syntax, and of course to accompanying gestures. Therefore, co-ordinated timing is in some ways a better term than sequentiality.

9. For an account of how speakers of English deal with responding to Wh-questions, see Fox and Thompson (2010).
After a first constituent of a turn/contribution, which could in itself be (retro) constructed either as a pre-front field filler (“pre-beginning”) or a front field filler (“beginning”), there is still a range of options. In addition, there are several kinds of possible “pre-beginnings” and “beginnings.” Consider an example from a Swedish conversation, with a police officer (P) interviewing a middle-aged woman accused of shop-lifting (W). Here, in the beginning of the interview, the policeman is trying to find out whether the woman wishes to have a defense lawyer present during the interview:

(1a)13
1. P: […] om du tycker att saker å ting e14 if you think that things are
2. besvärliga då ska du veta att (0.5) troublesome then you should know that you
3. [du har möjligheter att (0.5) tala me en advokat=] have possibilities to talk to a lawyer
4. W: [ja [0nä0 yeah no
5. P: om de,15 tycker du att de här e- de e about it, if you think that this is it is

10. Recall Note 5 in this context.
11. The terms ‘beginning’ and ‘pre-beginning’ are due to Scheglof (1996).
12. The Swedish data cited in (1a) and later (except (1b)) are attested spoken examples drawn from private conversations, institutional talk exchanges and media talk (TV, radio) (see Linell 2011 for more details).
13. All authentic examples used in this paper are given in Courier New. I use a somewhat simplified set of transcription conventions current in Conversation Analysis (e.g. Ochs, Schegloff & Thompson 1996). Note the following details:

- **underlining** (e.g. veta) marks a focal stress on a syllabic nucleus of a word.
- * * (asterisks) around a piece of talk indicates laughter in the voice (e.g. Excerpt 11).

The sign => means that the speaker’s turn is continued on a line further down in the transcript. **Boldface** and **grey-shading** are used in some excerpts to draw attention to specific aspects of utterances. Note that neither of these last-mentioned conventions refer to properties in the data themselves; rather, they are the analyst’s devices to increase readability.

14. å is an alternative conversational variant of och (‘and’) e of är (‘is’).
15. de is the spoken counterpart of standard written det ‘it’ in Swedish.
In line 6, the police officer introduces the word *bagatellartat* `trivial’, in sketching one option available (doing without a lawyer) in the current superordinate communicative project. After all, the legal case is one of ‘petty theft’, and it is therefore (legally) evaluated as precisely “petty” or “trivial” (*bagatellartat*). But let us suppose that the suspect reacts to *bagatellartat* as a potentially loaded word that calls for an expanded response from her. This might have caused her to repeat this word (such a repeated word will be referred to as X below). Let us further assume that she picks up on the everyday moral sense of shop-lifting, rather than its legal weight. What are her grammatical ‘options for turn continuation’ (Couper-Kuhlen & Ono 2007) in Swedish? Here are some alternatives (made-up), presupposing that she disagrees with the characterisation:

(1b) (possible continuations of (1a))  

i. *nå, de e inte bagatellartat för…*  
    PRT, 3SGPRO COP.PRES NEG trivial `cause…  
    (`no, it’s not trivial `cause’  
    (a simple negated predicative construction with the unmarked word order of Pronoun + Copula verb (+ Negation) + Predicate (containing the word in focus, = X = *bagatellartat*))  

ii. *nå, de e de inte för…*  
    PRT, 3SGPRO COP.PRES 3SGPRO NEG `cause  
    (`no, it isn’t `cause’  
    (a version of (i) with a pronominal predicate (instead of X = *bagatellartat*) in the front-field position and a contrastively stressed copula verb (*e*) (cf. Stivers 2005))  

iii. *nå, bagatellartat e de inte för…*  
    PRT, trivial COP.PRES 3SGPRO NEG `cause …  
    (`no, trivial it isn’t `cause’  
    (initial repeat of X in a predicative construction, cf. (i) with X in the front-field position)  

---

16. I have italicized these examples to mark that they are made-up (but “natural”) alternatives. Along with idiomatic translations, I have provided glossings here and, when necessary, in some cases below. For explanations of glossings, see Appendix 1.
iv. nå, bagatellartat tycker ja inte att de
   PRT, trivial think.pres 1sgpro neg subj 3sgpro
   € för...
   cop.pres 'cause

   'no, trivial I don’t think that it is but…’
   (initial repeat of X in a biclausal construction with ja tycker “I think”;
   in structural terms: Wh-extraction with an unbounded dependency’
   (in Swedish grammar called satsfläta, ‘interlaced clauses’; embedded
   clause with stressed copula verb (cf. (ii))))

v. nå, de tycker ja inte att de € för...
   PRT, 3sgpro think.pres 1sgpro neg subj 3sgpro cop.pres 'cause
   'no, I don’t think that it is 'cause'
   (pronominal version of (iv))

vi. nå, bagatellartat, de € de inte för...
   PRT, trivial, 3sgpro cop.pres 3sgpro neg 'cause
   ‘no, trivial, it isn’t that ‘cause’
   (X in the pre-front position (“extraposed”), i.e. not syntactically
   integrated in the following clause, but with a pronominal anaphor in
   this clause, cf. (ii)).

As can be seen in (1b), one important decision point in a possible response is before
the first substantial component, whether to put X in the initial position or not, or
whether to use a pronoun instead. After the choice of X, there is another decision
point; whether to construct X as a pre-front constituent (e.g. (vi)), as a fully integrated
front constituent (iv) or as a partially integrated front constituent ((vii) in (1c)).

So there are several alternatives, with the focused word (or a pronominal substi-
tute) as a predicate, in the pre-front-field, the front-field or the middle-field. But these
options do not form an exhaustive list. In fact, the woman in our recorded interview
did not choose any one of them. Here is the alternative that was actually chosen:

(1c) continuation of (1a)

vii.
9. W: [ja bagatellartat vet ja inte för att
   PRT trivial know.pres 1sgpro neg 'cause subj
   well trivial I don’t know ‘cause

10. de e ju (.) egentli fruktansvärt när man
    it is actually dreadful if one

11. tänker på de ef:teråt.
    thinks about it afterwards

Here bagatellartat does not appear in the pre-front position, as in (vi), but in the front
position of the finite clause (it is directly followed by the V2-positioned finite verb).
But note that bagatellartat is arguably only partially syntactically integrated, since it
lacks the rest of a split (Wh-extracted) clause (as in (iv)), that is, something like *om de(t) är/e, om ja tycker de(t) är/e* (‘if it is/if I think it is’) does not follow.

The range of possible options at different decision points is in general inversely proportional to the degree of constrained projection. Typically, projection becomes more constrained as we progress through the utterance and especially through the verb phrase. After the end-field there is a TRP, which makes a turn transition relevant but not necessary. So, after the end-field there is again a strategic choice: to continue with a post-end field filler (a turn increment, ‘post-completion’), or with a new TCU (perhaps a whole clause), or to relinquish the turn (at least temporarily), that is realise the TRP as a turn transition. If the alternative of the increment is chosen, there are still the options either to integrate (subordinate) it with/to the previous clause, or to expose it (set it off as a clear add-on) (see Couper-Kuhlen & Ono 2007 for more details). Another possibility is to retroconstruct the end field as the front field of a new syntactic unit, with the help of what Couper-Kuhlen and Ono (op.cit.) call a ‘glue-on’. This last-mentioned possibility comes close to a pivot construction, which will be the topic of Section 8 below.

In terms of decision points, one might talk about a major distinction between, on the one hand, boundaries and boundary zones, which include TCU boundaries and the zone involving (possible) pre-beginnings and beginnings, and, on the other hand, TCU-internal decision points, which mainly concern the construction of phrases and lexical choices. These different decision points not only concern current speakers, but also current listeners, particularly addressees. TCU-transitional points (boundaries) also constitute response points for listeners, who have to choose between no feedback, limited feedback (mere ‘continuers’), expanded feedback (more elaborated but still limited responses) and complete turn take-over. TCU-internal decision points usually only allow for limited feedback (when the speaker elicits it), unless the speaker calls for a fill-in or when listeners use competitive fill-ins or take-overs.

6. Early identifiability: External responsivity and internal projectivity

With regard to the above-mentioned zone involving pre-beginnings and beginnings, it is interesting that many languages already here allow for the identification or prediction of main ‘sentence types’, which can be regarded as ‘higher-order grammatical constructions’ and often considered to correspond to major speech act types. Here are examples:

(2): Major sentence types in Swedish:
Declaratives, initiated by Subject *V_{lit}; han* {*kan komma, kommer*}
3SGPRO.NOM AUX come.INF, come.PRES
(‘he can come, is coming’)
Y/n interrogatives initiated by $V_{\text{fin}}$ Subj:\textsuperscript{17} $kan$ $du$ $komma?$, 
$aux$ $2\text{sgpro}$.nom $come$.inf, 
$kommer$ $du$? 
$come$.pres $2\text{sgpro}$.nom

du?

Impersonal imperatives initiated by $V_{\text{imp}}$: 
$kom$
$come$.imp

Wh-questions initiated by Wh $V_{\text{fin}}$: 
$var$ $år$ $du?$, $vem$
$Wh$.adv $be$.pres $3\text{sgpro}$, $wh$.pro
$år$ $han$?
$be$.pres $3\text{sgpro}$

WhP questions initiated by WhP $V_{\text{fin}}$: 
$vilken$ $mössa$ $vill$
$which$ $cap$ $aux$
du $ha$?
$2\text{sgpro}$.nom $have$.inf

Exclamatives initiated by WhP Subject: 
$va$ $(bra)$ $du$ $kan$,
$Wh$.adv $(adj)$ $2\text{sgpro}$.nom $can$.pres

$\{va$, $vilken\}$ $fin$ $mössa$

$\{Wh$.adv $which\}$ $adj$ $cap$
du $har$
$2\text{sgpro}$.nom $have$.pres

As we can see, initial sequences (here abbreviated after “initiated by” under each item) 
are analysable as indicators of major types on the basis of the order of (in each case) 
only two constituents (disregarding pre-positioned particles, and the like): the finite 
verb form, the subject NP and/or a Wh-constituent.\textsuperscript{18}

\textsuperscript{17} This syntactic order occurs in some other constructions as well, notably V1 declaratives 
and V1-conditionals (conditional constructions with V1-clauses as protases). I abstain from 
going into more details. However, we must not forget that constructions are also linked to 
particular prosodies and interdependencies with prior utterances, which can help participants 
to keep these constructions apart.

\textsuperscript{18} These facts have long since been observed (Ahrenberg 1978; Anward 1980). However, 
parts of these interrelations have been respecified within interactional grammar, which is a 
kind of \textit{syntaxe de la parole}. In the case of Swedish, they have been explored by Jan Lindström 
(2008).
Major sentence types, taken as constructions (types rather than situated tokens), can occur as first-positioned contributions in topical episodes (or adjacency pairs). Thus, they are not inherently responsive to prior contributions. But the latter is still the case with many other grammatical constructions. Günthner (2006b, 2011) and Auer (2009b) have discussed several such constructions in German. One example is the pseudo-cleft construction (Wh NP V Cop {NP, Subj (that) S}; Wh AdjP Cop {NP, Subj S}), as in (3):

(3) (Auer 2009b: 190)

Wh AdjP Cop Cop NP: was halt toll ist, is die Ostküste
What PRT is fantastic is the eastern coast

Here, the initial constituents (Wh AdjP Cop) forming a dependent clause and placed in the front field of the whole construction, are enough to identify the upcoming utterance type. These constituents therefore project the rest of the construction. Günthner (op.cit.) and Auer (op.cit.) devote extensive discussions to this kind of internal projectivity (Wh AdjP Cop is, in Günther’s terminology, an internal ‘projector phrase’). However, the pseudo-cleft construction also typically occurs when aspects of the contents of the Wh-clause have already been introduced in the preceding discourse. To illustrate this, I borrow a longer example from Auer (2009b):

(4) (Auer 2009b: 182–183) A former owner of a small bakery (M) explains to a female interviewer (I) how prices in bakery shops have gone up recently, for example, on pretzels. The two argue about the reasons for this. M has claimed that it is a lot of manual work to make pretzels. “Pfister” mentioned in line 16 is a well-known bakery chain in Munich:

1. I: aber des is ja net der preisaufschlag but this doesn’t make the higher price ((of the bakery pretzels compared to those in the supermarket)

2. der preisaufschlag is dann wahrscheinlich ladenmiete the higher price is probably ((due to) the shop rent

3. und gehäl[ter etcetera, des kommt ois dazu no ja and the salaries and so on, all this comes on top PRT

((8 lines omitted))

13. M: a:ber (0.5) des is: heute -- but it is today --

19. I have simplified and partly adapted the transcriptions in (3–4) to the norms used in this paper.
20. The original does not include focal stresses for the first four lines.

All rights reserved
14. wenn man schaugt, die verlanga ja
   if you look, they charge

15. (0.4) wo ich also den preisaufschlag überhaupt nicht
    where I don't think the higher price is justified
    (Auer 2009b, 184). It thus ties back to prior talk on higher prices (e.g. lines 1–2: preisaufschlag) and their justification; note the use of the words preisaufschlag (‘higher price’) and gerechtfertig (‘justified’) in the Wh-clause. This external tying back arguably contributes to (pragmatically) projecting what is to come, that is, further arguments concerning the (non-)justification of higher prices (lines 16ff.).

   One feature of the whole pseudo-cleft construction in (3) is that what follows is not a copula plus a dass- (‘that’) -clause, but rather the start of a piece of discourse with several separate main clauses that are syntactically (but not pragmatically) independent of the prior pseudo-cleft segment. Günthner (op.cit.) and Auer (op.cit.) show that in conversational German, this is a common feature of several other grammatical constructions, including die Sache (Pointe, Punkt, Problem) Cop... – construction and conditional constructions initiated by a wenn- (‘if’) -clause (cf. lines 16–17 in (4)). Thus, conversational German tends towards asyndesis and parataxis, rather than the bipartite constructions with dependent clauses that are described in normative grammar books. On-line syntax is often what Auer calls ‘minimalistic’ (NB: not in the generativist sense). Indeed, Auer (op.cit.) demonstrates with ample examples how the incremental on-line syntax in impromptu conversation involves a multitude of syntactic projections, interruptions, abandoned projections etc., pointing to the fact that utterance-planning is partial and local (see Section 11 below). Hopper (2004) makes the same point for pseudo-cleft constructions in English. He notes that normative grammarians would probably regard the fragmented forms current in spoken interaction as “performance errors” and as reduced and “degenerate” forms of the full bipartite sentences in writing and more planned discourses.

   The constructions mentioned in the previous paragraph display a combination of external and internal projectivity (after a few initial constituents). These features are even more pronounced in certain quite specialised constructions, such as those which
have been called reactive constructions (Lindström & Linell 2007; Linell & Norén 2009; Linell 2011). These involve, by definition, a repetition of a particular lexical item \(x\) from a prior contribution (thus, the external responsivity is quite explicit), and are typically used when the speaker feels that (s)he cannot reply by a plain yes or no to a prior question (i.e. cannot give a type-conforming response (Raymond 2003)), or cannot offer a clear affiliation or disaffiliation after a claim. The speaker can then problematise the use of \(x\) by placing it in the initial part of a special construction; consider (5) (which is an authentic example):

(5) Looking at a rough coast line in the countryside, one person A asks another one B:

1. A: kan man bada här?
   ‘can you swim here?’.
2. B: kan kan man men man vill inte—
   aux aux 3SGPRO (‘one’) Conj 3SGPRO Aux neg
   one can but one doesn’t want to

In her reply B (line 2) deploys a peculiar construction that we have so far attested only in Swedish, that of double initial finite auxiliaries \(x_{aux} x_{aux}\). This reactive construction begins with two finite auxiliaries (in this example: kan ‘can’). The first item may be interpreted as a literal quote from the preceding question, while the second one is an identical form serving as the finite V2-positioned verb integrating the expression into a finite clause. Usually, this clause is followed by another clause starting with men ‘but’. This second clause can be seen as an open slot which is also part of the construction. In using the \(x_{aux} x_{aux}\) construction as a whole the speaker concedes that the condition referred to by the quoted auxiliary is true of the situation in question, but it is to be downplayed as more or less irrelevant in the light of some other, more relevant circumstance referred to in the subsequent but-clause.

An alternative for B in a situation like (5) would have been the \(x\)-och-\(x\) construction, in this case kan och/å\(^{22}\) kan, det beror på... (literally: ‘can and can, it depends on...’). Unlike \(x_{aux} x_{aux}\), the \(x\)-och-\(x\) construction negotiates the situated applicability of an expression \(x\) (here: kan), conceding that one could say \(x\) about the present situation, but there are also arguments against it. Regarding these two reactive constructions, there are many more details to account for (see Linell 2011), but the point to be made in this context is that these constructions are, like the general ones given above in (2),

\(^{21}\) German has a close counterpart, using the infinitive in the first position: können kann man aber... For more details, including cross-linguistic comparisons, see Mertzlufft and Linell (fc.).

\(^{22}\) Here, och is the written form, and å the colloquial spoken form.
identifiable on the basis of beginnings; no other utterances could begin with two juxtaposed finite auxiliaries or with a pair of ōch/å-conjoined identical words (x-och-x). Furthermore, both constructions are, as pointed out above, strongly responsive, and their first core segments are also internally projective of particular types of continuations.

Grammatical constructions arise from the experience of recurrent patterns in discourse and interaction. That implies that they are often designed to respond to prior contributions and project possible next contributions. Some constructions are saliently responsive, e.g. the reactive constructions mentioned above. Others are responsive to prior discourse, e.g. pseudo-clefts, but what they more specifically refer to varies. Still other, more general constructions (the major sentence types above) are not responsive on the type level (although situated exemplars may of course be responsive to prior actions). Yet, we have seen, at least for Swedish, that there are constructions in all three categories that allow for early identifiability, through external responsivity (only some cases) and internal projectivity due to specific beginnings.


We are now beginning to see what a “significantly untraditional approach” (Section 2 above) might involve. We are looking for a partly new meta-language and conceptual framework encompassing not only unit types, but also decision points and continuation types (continuations that either fulfil or suspend projections). Decision points (points of option) could perhaps be derived from a syntagmatic-paradigmatic model of unit types (and utterance fields?), but they are nevertheless important in an action-based theory. Many grammatical constructions display early identifiability, which seems to be a major advantage provided by syntactic constraints (Auer 2009b). More specifically, we need a ‘processes-and-resources’ theory of languaging, instead of a ‘units-and-rules’ theory of language. Further, processes and resources are closely related:

a. structures of language have emerged from the primordial processes of languaging (rather than the other way around),

b. when habits or patterns of usage develop into conventions, they come to function as constraints on utterances (Rączaszek-Leonardi 2009). In Selting’s (2005, 21) (and my own; Linell 2009) terms, syntax and prosody are methods for the construction and identification of units.

c. participants in cognitive and communicative practices are concerned with content and expression, rather than building grammatical products (such as sentences), and
d. – partly as a consequence of (c) – utterances may involve non-linguistic components (such as manual pointings, gazings (Ford et al. this volume), gestures or other bodily demonstrations (Keevallik 2013); these form ‘composite utterances’ (Enfield 2009) and sometimes yield unintended (‘ungrammatical’) outputs (more about this below).

In one sense, ‘processes’ must be assumed to take precedence over ‘resources’, unlike in structural theories of language, where units are basic to combinatorial rules. ‘Processes’ are dynamic, temporally distributed ways of dealing with events and projects as they emerge over time (‘on-line’). ‘Resources’ can only be defined in relation to actions or practices, and thus to processes. Their nature is that of potentialities, rather than of actualities (until they become actualised and thereby made relevant in interaction). Linguistic resources always work in interplay with contextual resources in situated languaging; contextual resources include external objects and artefacts, various kinds of background knowledge, knowledge of genres and activity types, etc.

Ford, Fox and Thompson (this volume) have argued along partly similar lines. However, they appear to take a more radical position in avoiding “linguistic units” such as noun phrases and pronouns altogether in their action-based characterisations of linguistic forms. Yet, the possible differences between our approaches may be gradual rather than categorical.

In their paper, Ford, Fox and Thompson analyse a situation with three women, Maureen, Terry and Abbie, who are sitting around a table in the home of Terry (and a fourth woman who happens to be absent at the moment). The authors home in on two particular utterances, the cafe de yin ya:ng? when he was tw- ten? and it’s coo:l, in their linguistic and interactional environments. The first (complex) utterance is said by Maureen, and the second one is (first) uttered by Terry, but in a sequence initiated by Maureen. Both utterances point to pictures on the wall, one with a drawing, allegedly made by a ten-year-old boy, and the other a collage of photos of sheep. The examples clearly show that the linguistic resources are only some of the ingredients of the embodied actions and interactions. In particular Maureen’s gaze direction is used as a pointing resource in the course of establishing the two referents as objects of shared attention. In my terms (Linell 2009), Maureen is acting as the main participant in two temporally distributed, embodied local communicative projects, in which she shows her own focus of attention, seeks and manages to receive the others’ attention to the same objects, and then exploits this shared attention to make positive assessments, which are followed up by second assessments by one or both of the others. Both these local projects are comprised of yet smaller projects, especially those focused on getting attention to the particular referents. In these ‘referring projects’ Maureen uses both gaze direction and language to draw attention to the two pictures.
The theoretical “experiment”, as the authors of Ford et al. (op.cit.) put it, is to avoid linguistic concepts, such as noun phrase, in the analysis. The argument is that noun phrases are not part of participants’ projects. However, I would argue that participants’ actions have two sides, a pragmatic side involving reference and assessments (along with many other aspects) and an expression side, with the use of several semiotic resources (Goodwin 2000), including “language”. Linguistic resources, that is, the “language” you need in order to indulge in languaging, can be conceptualised as vocal (phonological, prosodic), gestural, bodily, grammatical and lexical items and methods. Accordingly, for the purposes of describing the language aspects of talk, linguists can hardly do without some traditional linguistic notions, as Ford et al. observe (such as those mentioned above in Section 3 and later to be used in several sections of this paper). I would also argue that participants’ verbal behaviours are structured not only at a level of phonetic gestures, but also in terms of lexically and syntactically defined chunks. These come in packages of verbal gestures, some types of which may be seen as “phrases”, such as noun phrases. Such packages have several characteristics that show that they are relevant to participants, that is, to speakers and listeners; for example, they are held together by prosodic contours, usually include certain constituents (in the case of noun phrases often articles like the), are used as chunks that can be positioned in various positions (e.g. as “dislocations” in the pre-front field), and constitute domains also for involuntary (but grammatically “permeated”) processes like slips (see Sections 9 and 10 below). At the linguistic-structural level we could still talk about noun phrases and clauses.23

In conclusion, linguistic actions have a material aspect too, as methods to build recognisable utterances. An account of this may make use of some “units” (i.e. units with(in) structure, including responsive and projective relations). To relate to the title of this volume, I see myself, in this paper, as focusing on (some of) these units of talk, yet assuming that they are framed by units of action.

8. Pivot utterances

The logic of the following sections runs as follows. If we want to know something about the “reality” of operative processes in talk, deviations from normative grammar can be quite telling. By contrast, as long as speakers follow standard norms perfectly, we cannot always tell which processes, apart from those which are simply norm-conforming practices, are actually at work. What will be argued in the following Sections (8–10) is

---

23. This is not to say that we do not need to think carefully about terminological issues. For example, it is far from clear that we would want to regard the it of typical assessments like It’s cool as a noun phrase (there is no noun and no phrase).
that processes in the on-line syntax of interactional languaging mostly work with local dependencies within relatively small ‘units’. That is, utterances are built incrementally.

Syntactic structures in particular languages are largely conventional. Utterances are planned in terms of orientation to routinised units and their sequencing, and this also applies to the coordination of different dimensions, such as grammar, prosody and gesturing. But at the same time, utterances are largely automatised in the speaking practices of language-competent individuals. One kind of evidence for this is the existence of unintentional deviations from the norm. Two types of these include non-fulfillment of agreement constraints (Section 9), and involuntary permutations (and other movements) of units (Section 10). But first (Section 8) we shall take a look at a phenomenon, which, I argue, is part of the grammar, yet is often misrepresented by normative grammarians.

A particularly instructive phenomenon as regards incrementation in spontaneous conversation is the frequent occurrence of so-called pivot utterances (or apo-koinou) (Norén 2007). In such an utterance type, the speaker starts with one syntactic construction, but at some point (usually after a possible completion point) switches to another unit that does not fit the beginning syntactically and/or prosodically, and takes as its starting point only the last part of the prior utterance segment, i.e. the so-called ‘pivot’. Thus, retrospectively we are faced with an utterance with mutually incongruent beginning and end and with one segment, the pivot itself (grey-shaded below), which “belong to”, or “is shared” by, both the beginning and the end. The segments preceding and following the pivot are called ‘pre-pivot’ and ‘post-pivot’, respectively. In the examples below, the whole pivot constructions are given in bold.

Pivot utterances do not instantiate one single clear-cut grammatical construction. They are possibly a family of related constructions, among which some are more sedimented as construction types than others. One subfamily consists of (what is often called) mirror-image constructions.

(6) From a telephone conversation: A is asking B if he will be available on the phone any night later in the week, perhaps the upcoming Friday. B answers:

1. B: ja e hemma: ≥ (.) åtminstone vid halv-åtta-tiden e ja hemma.
   *I am home (.) at least at half-past-seven-ish am I home.*

(7) From a TV program about film director Ingmar Bergman’s house-keeper; here she tells about how and when dinner should be served in the late afternoon; Bergman evidently spent the time before dinner upstairs:

1. A: då kom han ju ner då klockan sex (. ) prick kom han ner.
   *Then he came PRT down then at six (. ) sharp he came down.*

In mirror-image variants like (6)–(7), the pre-pivot and the post-pivot are identical, except for the regular change of word order and the unstressed då’s (‘then’) in the
pre-pivot of (7). But post-pivots seem to have a confirming function, and are therefore hardly superfluous. Usually, the pivots contain several content words, and are evidently not planned in detail from the beginning of the utterance. This will arguably make a repetition relevant. In (6), speaker B has been asked for an exact time, but cannot provide it immediately; he projects a continuation by lengthening the final vowel and using a “more-to-come” prosodic contour on hemma, which is followed by a micro-pause before the time adverbial is delivered. In (7) the add-on of prick ‘exactly’ after a micro-pause within the pivot, arguably contributes to making the post-pivot more relevant.

In other cases the post-pivot adds something more, and thus contributes to a perspective shift, sometimes subtle, sometimes more salient (see Norén 2007, 2013, for more about perspective shifts). Here the post-pivot communicates something more and different than the content of the pre-pivot. Such cases have been called ‘Janus-faced’ pivot constructions (they have significantly different pre- and post-pivots). Some examples are (8–10):

(8) During a flight the co-pilot informs the passengers about the weather at the destination:

1. P: vädret i Linköping dom rapporterar svag vind, lätt
2. regn (.) och sju grader har dom just nu.
the weather in Linköping they report light wind, light rain (.) and seven degrees they have just now.

In (8), the pivot consists of a list, with a micro-pause before the last item, which makes it rather complex. This may call for a partial repetition of the content of the pre-pivot, but in this case the restatement comes to include a specifying aspect: the report just communicated is about the present weather conditions, implying that things may later have changed by the time of arrival.

Pivot constructions can be used for changing a conversational turn from one type of communicative action into another one:

(9) A, who has acted as chair in a seminar, asks a colleague if there are any plans for a lunch between sessions:

1. A: e de nära arrangemang för lunchen har vi väl inte va?
are there any arrangements for lunch we don’t have, have we?

In (9), speaker A changes his open interrogative yes/no construction into a declarative question. This is done by transforming the pivot segment (nära arrangemang för lunchen ’any arrangements for lunch’) from a predicate of the pre-pivot into a first-positioned object of the post-pivot, which is tilted from openness towards a negative answer.
Pivot utterances make use of grammatical resources that can occur in other contexts too. Many post-pivots are common-type increments that occur as (prior turn/TCU) expansions in general. Some are devices that are used in places other than final TCUs. Take for example reporting phrases expressing “thoughts”:

(10) The discourse is about a famous person who came to Sweden as an immigrant:

1. A: **han kom till Sverige nittonhundratolv tror ja de va.** he came to Sweden in 1912 I think it was

The epistemic stance specification *tror ja (de va)* ‘I think (it was)’ is linked only to the pivot, which concerns the year of the event told about (that the person had moved to Sweden was already known to participants). The time specification seems to be the important new information (and the only focally stressed part) in the first clause but the specification is epistemically transformed from a factual statement to an item of belief. A change of epistemic stance is accomplished by the *tror ja* add-on.

The post-pivot in (10) is something commonly found in many increments that do not give rise to pivot utterances. Such add-ons also typically appear in parenthetical insertions (Norén 2012). For example, the inserted segment **hörll ja på säga** ‘was I about to say’ in (11) has its more or less exact counterparts in many post-pivots:

(11) (TSYN-L(22); Talsyntax): A conversation with four physicians discussing problems connected to euthanasia, recorded in 1968, here quoted from Norén (2012); D is one of the discussants:

1. D: **[men man] kommer väl in på en mycke *äckliare*** but you are entering a much more disgusting w(a)s (I) (a)bout

2. **h(ö)ll j(a) p(å)säga >för vår egen del<** besvärligare, (.) w(a)s I (a)bout to say >for ourselves< more difficult (.)

3. **.hh::ställningstagande de e ju så mycke som de gäller .hh:: stance there are so many things where it’s important**

4. att ha ryggen fri? (0.4) så att säga va¿ (.) för så många. to watch your back(0.4) so to speak right¿ (.) for so many.

---

24. For a similar argument concerning ‘sluicing’ in English, see Hopper (2011, 39). Sluicing utterances can be regarded as built from two formulaic parts, first I don’t know (or some equivalent), then a “Wh-piece” like why, about what, where etc. as in We knew we were losing oil, we didn’t know where (op.cit.: 35).
5. (.) hh *utav oss*.
( .) hh *of us*.

The speaker produces his talk in exposed increments. This applies not only to his inserted segment in bold (lines 1–3), but also for the last three or four installments preceded by short pauses in lines 3–5. (I will return to this example below.)

Another common type of pivot-like utterances uses tag-like final increments, such as gör dom/jag/han ‘they/l/he do(es)’:

(12) Somebody has dug up ditches around the base of a residence house, putting coarse-grained gravel there instead of flower beds. B lives in the house, A is a visitor:

1. A: ska dom reparera huset också?
   are they going to refurbish the house too?
2. B: dom sätter nya ventiler gör dom.
   they are putting up new ventilators they do

B here confirms that the refurbishment does include some repair work (cf. A’s question), but only a limited amount (only ventilators are being fixed). Note, however, that in this case the increment does not quite work as a typical post-pivot. Rather the increment, with the Pro-verb gör ‘do’, confirms the whole preceding segment. What becomes the pivot is not only the NP but the whole VP in which the NP is the direct object.

In pivot utterances, the post-pivot is most probably almost never planned from the beginning. Instead, the speaker makes a switch in mid-utterance, both grammatically and semantic-pragmatically (the latter constitutes the perspective shift, according to Norén 2007, 2013). The post-pivot is responsive to what the speaker has said just before. It is a case of how actions are responsive to prior actions and projective of future possible actions, but here the dialogical interplay occurs within the speaker’s turn, rather than across turns by different speakers; the speaker is “in dialogue with” her/himself. At the same time, of course, the whole turn is responsive to (another’s) prior action, and may project next actions by others. These aspects, however, have not been analysed in relation to the examples of pivot utterances.

To sum up, pivot utterances are frequent phenomena in conversational languaging. They are clearly built incrementally, and cannot be analyzed as single, complex ‘sentences’. They provide important evidence for an incremental theory, rather than a hypothesis that speakers start out from full sentence plans (Wundt’s Gesamtvorstellung, Section 4).25

---

25. This incremental theory of pivot utterances is spelled out more exhaustively in Norén and Linell (2013) and Pekarek Doehler and Horlacher (2013). The most comprehensive work on Swedish pivot constructions is Norén (2007).
9. Non-fulfillment of agreement constraints (projections)

Pivot constructions are not noticed by speakers and listeners as remarkable in natural languaging, and in our data they are never repaired. There is every reason to believe that they are integral parts of the normal grammar of conversational language. Yet, they have most often been excluded from official grammar books, as in the Swedish SAG (1999).

Let us now move to a phenomenon in conversation that, in stark contrast to pivot utterances, is reasonably seen as grammatically deviant, namely, faulty agreement within noun phrases (NPs) (in which modifiers often precede the nominal heads in Swedish or German), or across adjacent NP boundaries, as in subject + predicate constructions (in Swedish).²⁶,²⁷

For a first illustration, we can return to example (11). We noted that the speaker inserted a parenthetical expression (similar to many post-pivots) (grey-shaded in (11)). This insertion interrupts an ongoing production of a complex noun phrase; the interruption occurs after a segment consisting of an indefinite article (en ‘a’), an intensifier (mycke ‘much’) and an adjective (äckliare ‘more disgusting’), which together project a nominal head that has not yet been produced. When a noun is finally produced, after the insertion, a new adjective in the comparative form (besvärligare ‘more difficult’), a micro-pause and an inbreath (line 2), it does not fit the beginning. Ställningstagande ‘stance-taking’ is a neuter noun, which should have required a neuter article (ett instead of en). Most probably, the speaker had aimed at a non-neuter noun from the beginning, such as problematik ‘problematic’ or inställning ‘attitude’.

The non-agreement in (11) is not repaired. Such non-repairs are quite typical in cases of non-agreements when other linguistic material occurs between the determiner and the head noun. In (11), there is a rather lengthy insertion, but in other cases the interstitial segment may be shorter. We shall look at a few examples here. Let us first take a few more NP-internal non-agreements.²⁸

---

²⁶. Note that in this context ‘agreement’ (a.k.a. ‘concord’) refers to agreement between syntactic constituents, and not to agreement between participants in interaction.

²⁷. German is similar to Swedish, with a differentiation of nouns according to grammatical gender and number. However, in German there is no adjectival predicate agreement with subjects in copula clauses; adjectives are uninflected in this position, unlike in (standard) Swedish (cf. examples (16–18) below). Otherwise, the same kinds of uncorrected non-agreements seem to occur in German (Pfeiffer forthc.).

²⁸. The examples of Swedish non-agreements and slips (Section 9) are drawn from my corpus of actually attested grammatical “errors” assembled mostly around 1980. Such errors occur rather infrequently, and need usually to be noted on the fly. Even today they are very difficult to collect using available methods of searching through spoken language corpuses. It is possible that a few of my examples contained micro-pauses that I was not able to record.
(13) From a TV program about things worth seeing in southern Sweden (2012):

A: idag e Karlsborgs fästning en av norra Europas största befästningsverk

today Karlsborg’s fortress is one of northern Europe’s largest fortifications

A complex predicate NP (in bold) in (13) is initiated by a numeral in non-neuter form, "en ‘one’, which is then followed by an embedded prepositional phrase containing a genitive phrase ("norra Europas ‘northern Europe’s") followed by an adjective ("största ‘largest’"), and only then comes the head noun befästningsverk ‘fortification(s)’, which, however, is neuter and therefore would have required the neuter "ett ‘one’ instead of "en. Arguably, the speaker is influenced by the adjacent, preceding non-neuter noun fästning (which could have been used as the final head noun also in the boldface phrase).

(14) From a political TV debate in 1981:

A: […] en uppläggning som vi varit överens om under den senaste året

a layout that we had agreed upon during the (non-neuter) last year (neuter)

The definite article "den does not fit the neuter "året ‘year’. The speaker might have entertained "den senaste tiden ‘time’ as an alternative, an expression he in fact used in the next utterance to come (not shown here).

Non-agreement can also affect definiteness or number:

(15) From a corpus of spoken Swedish (Talsyntax) recorded and transcribed in the 1970’s; this example deals with national politics:

A: men i en koalitionsregeringen så finns de väl […]

but in a (indef.) coalition government (def.) there is perhaps…

Here, the speaker seems to change from the indefinite form (with the preposed article "en) to definite (the definite ending -en). In normative grammar, these cannot be combined within a consistent construction. Yet the speaker does not go back to correct the NP as a whole.

Other examples of non-agreement in number or gender concern relations between a (subject or object) NP and a predicate:

(16) From a political discussion in TV in 1977:

A: […] den enighet i värderingarna som e nödvändiga.

the (def.sg.) unanimity in the values (def.pl.) that are necessary (def.pl.)

Here, the head noun of the composite NP is in the singular ("enighet"); but the predicate in the relative clause is plural, obviously affected by the more adjacent plural noun värderingarna.
From a political commentary in TV in 1978:
A: de finns nånting i den debatten som ja tycker e oerhört vikti.

there is something (neuter) in that debate (non-neuter) that I think is enormously important (non-neuter)

A neuter head noun (nånting) would have required a neuter adjectival predicate (viktit) in the relative clause, but the interjacent non-neuter noun (debatten) seems to have taken over.

In (16–17), the non-agreeing adjective could arguably have been predicated about the interjacent noun as well. But this is not necessarily the case:

From a political interview in TV in 1980:
A: [...] otåliheten bland dom offentlit anställda e mycke stora.

the impatience (sg.) among the publicly employed (pl.) is very great (pl.)

In (18) stor ‘big, great’ is of course not predicated about anställda ‘employees’, but about the subject otåli(g)heten ‘the impatience’. However, what is common to (16–18) is that the nearest noun takes precedence by a principle of proximity, and the predicate agrees grammatically with that item, and not with the head of the complex NP or the subject.

We noted above that non-agreements are arguably ungrammatical, in contrast to pivot constructions. Nonetheless, in none of the above cases (11, 13–18) are the faulty non-agreements repaired (although such repairs do of course occur in exceptional cases). When non-repaired non-agreements occur, it seems that the speakers’ ‘self-monitoring’ (Levelt 1989) either fails or is ignored. In dialogical terms, the speaker prefers to continue his/her utterance, rather than anticipating or adopting the perspective of a critical listener who might react to the inconsistencies. But actual listeners too seem to remain inattentive most often.

An issue which is so far relatively unexplored in research is whether there are any links between faulty agreement and interactional aspects of turn construction. Yet, we can say that these phenomena can be explained in a processual account, which assumes that utterances are built incrementally using fragments of projected expressions. By contrast, a ‘units-and-rules’ theory, especially one which rests on the notion of grammatical sentences, would provide a less satisfying account, since the resulting unit types do not exist in normative language, something that native speakers know.

10. Planning as local and partial

The phenomena that I have here called ‘non-agreement’ deviate from normal usage: any speaker of Swedish (or German) would recognise that if they had attended to them

All rights reserved
(as is easier to do in writing). Yet, these unintended inconsistencies do not seem to be noticed by the speaker (at least not in the majority of cases, as exemplified above), and they are not (self-)repaired. Nor are they repaired by addressees; other-repair is non-preferred anyway (Schegloff et al. 1977). Things thus seem to proceed communicatively smoothly. The interesting point (for us here) with the ‘non-agreement’ phenomena is that they do show grammatical agreement, but not the normatively prescribed one. There seems to be a grammatical process operating that results in agreement with an adjacent or hierarchically dominant (often grammatical subject) controlling NP.

Unintended agreements are evidence of real processes occurring in the utterance building. (By contrast, when speakers produce grammatically ‘correct’ utterances, it is hard to know what the operating processes are.). In the ‘non-agreement’ cases of Section 9, there is evidence of planning ahead. But this planning is local and partial. Similar evidence for local planning can be drawn from data on so-called slips of the tongue. In the corpus of about 1400 such slips in Swedish analysed in Linell (1982), there are many permutations, anticipations etc., like (19–20):

(19) lyngdtyftning (for intended: tyngdlyftning ‘weight-lifting’)
(20) stord å bolar (for intended: bord å stolar ‘tables and chairs’)

If the corresponding permutations had happened to the English translations weightlifting and tables and chairs, they would come out as leightwifting and chables and tairs, respectively. Such slips are quite different from the ‘non-agreements’ (Section 9) in that speakers almost invariably interrupt themselves, retract and pronounce the phrases correctly (evidently as intended). (Yet, such a repair would point to an orientation to a linguistic norm of immaculate pronunciation; this is not repair done for better comprehension, since it is obvious what the speaker wants to say, in terms of choice of words). But the errors occur predominantly within close-knit phrases (NP, PP, VP) and rather simple clauses (Subject (Copula) Predicate, Subject Verb Object), just like the ‘non-agreements’. That these inconsistencies occur almost only within such structural units indicates that linguistically specified planning tends to be relatively short-sighted. Beyond this local planning, there is only a more diffuse planning concerned with participants’ larger communicative projects.

11. The status of grammatical constructions

What is the role of grammar and language within a theory of languaging, such as on-line syntax? What, specifically, is the status of grammatical constructions? Here

For many more cases, see Linell (1982). The classical work on English speech errors of these types is Fromkin (1973).
(and earlier: e.g. Linell 2005, 2009), I have referred to grammatical constructions in at least three ways, (1) as abstractions from utterances and utterance types, (2) as conventionalised patterns in utterances, recognised by language users and oriented to in situated languaging, and (3) as (participants’) methods of utterance-building. Imo (2011, 152) sees this as inconsistent, and as showing an “indecision about the actual status of grammar”.

Imo may have a point here. But although his verdict was based on an earlier formulation of mine (Linell 2005, 219), I am still inclined to think that the above-mentioned three-fold account is more or less consistent. Grammatical constructions can be seen in different perspectives, depending on one’s point of departure (e.g. specific situations vs. practices/traditions). Yet I would put most of the emphasis on the view on constructions as methods (which I gather is also Imo’s preference). But these methods must be based on participants’ experiences of how people talk and interact in the language in question. Participants do – consciously or, more typically, unconsciously – attend to similarities and patterns across situations (Bybee 2010). Whether these are generalised patterns (formats), or based on analogies with particular concrete utterances (exemplars) is an empirical issue currently the subject of research in language and cognition. As Bybee (2010) suggests, it is reasonable to think in terms of a continuum from general patterns to specific cases.

If we accept the argument about methods and patterns, what about constructions as abstractions? Well, first of all, we should be careful and remind ourselves that we are dealing with abstractions from real utterances and utterance types. We should avoid reifying these abstractions; that is, we should not create or presume another abstract (mental, cognitive) language beyond, or in addition to, external public language. In addition, we could ask whose abstractions we are dealing with. It seems reasonable to think of the language users as making implicit abstractions, when they create habits and patterns in and for their languaging. But linguists have to make the abstractions explicit, when they create their meta-language. This of course involves a recontextualisation, from practices to theorisation. We start out from utterances and interactions in real data, but our theorisations are framed by our – the researchers’ – interests and activities. This is true of CA practitioners as well, when they as analysts speak about taking ‘the members’ perspective’.

12. Some concluding points

1. In this chapter I have shown that utterance building is an incremental process, in which utterances are produced piece-by-piece on-line. But these pieces are not autonomous words or phrases, but ‘units’ with structural relations to surrounding pieces. In this way projections and fulfillments of projections are created, while incrementation is arguably a universal phenomenon, it is clearly subject
to language-specific constraints. This chapter has primarily used data from continental Germanic languages, in particular Swedish and German.

2. An adequate theory of on-line syntax must be able to handle the living processes in language and languages. It will shatter several fundamental wisdoms of formal linguistics. One would be the belief in the strict compositionality principle, that is, that all complex expressions and their meanings can be derived from constituent expressions. On-line syntax (Auer 2009a, b), construction grammar (Croft & Cruse 2004) and emergent grammar (Hopper 2011) will most probably assume some kind of (looser) compositionality, for example, accomplishing incrementation by the use of 'formulaic language' (or prefabs) (Wray 2002), fragments, and constructions with some gestalt-properties. Yet, they all reject strict compositionality. Furthermore, we have reasons to assume that a grammar of real languaging exhibit flatter structures (fewer hierarchies), more fragmentation, and richer content (Bybee 2010) in lexical units and grammatical constructions than in most formal theories.

3. Processes of utterance-building by single speakers are in themselves interactional in nature. In the course of a more or less continuous incrementation, the speaker responds to what others have said and what (s)he has said before, and thus projects what may be said afterwards, by self and other. But equally important, the speaker's own utterance-building and meaning-making reflect an internal dialogue. The silent voice of self-monitoring oftentimes makes the speaker change or supplement his/her utterances in the course of development. One could imagine this as a tension between the speaker's wish for efficiency and expediency in expressing his/her message, and his/her assumptions or anticipations of the recipients' demands for clarity and acceptability of the utterance output.³⁰

4. In building utterances, speakers (and listeners) orient to projected structures of language. Yet, such processes can also produce unintended outputs, which can be regarded as ungrammatical; yet sensitive to grammatical dependencies. Some of these utterances pass unnoticed in interaction (Section 9). As Imo (2011, 150) argues, interactants are not so “picky” about language; they tolerate problems with “small inconsistencies” and “ambiguity” in languaging. In addition, as Breyer, Ehmer and Pfänder (2011) have pointed out, language users are not only seeking and following patterns; they also improvise, and stretch their resources to cover communicative needs in new situations.

5. So, languaging consists of dynamic actions and processes, physical movements, accomplished on-line, dynamically responding to events in the world. That utterance planning is relatively local and partial is obvious from many kinds of

³⁰ Pfeiffer (2012) similarly talks about “competing motivations” (after Du Bois 1985), but chooses to present the issue in slightly different terms.
conversational data; for example, we demonstrated it in pivot utterances and non-agreements within phrases and clauses. These phenomena also speak for an incremental production of utterances.

6. Despite what was stated earlier about ‘units’ and ‘unit types’, languaging builds on real-time activities and processes and resources used in them, rather than on a system of units and rules. We do not build utterances by manipulating object-like building-blocks in analogy with playing with Lego toys. Yet, in the on-line processes participants orient to utterance patterns, “units of talk”, with internal dependencies that give rise to projections. These patterns have been extracted from rich experiences of utterances as perceived products. When such habits or patterns of usage have emerged and developed into conventions, they have come to function as ‘second-order’ (Cowley 2011b) constraints on utterances (Račzaszek-Leonardi 2009; Thibault 2011).

7. Although this paper has focused on situated utterance production, it is important to recognise that in these processes participants orient to patterns that are the result of sociohistorical developments. The patterns have themselves emerged over longer time stretches, in the sociohistorical development over many generations of speakers, or (in some cases) in the individuals’ communicative biographies. Such patterns are also subject to dynamic renegotiations over time. That they change in sociocultural history is due to the fact that active language users can change their own practices over time consciously or, more typically, unconsciously, often of course across generations of speakers.

Using Hopper’s (2011, 27ff) terminology, utterance types are both emerging (in sociohistorical genesis) and emergent (in situated micro-interaction). While Hopper acknowledges both processes, he justifiably puts more emphasis on the latter (nothing can emerge except in and through (repeated) reenactments). But the theory of emergence must be ‘doubly dialogical’; it concerns both situations and traditions (or practices) (Linell 2009).

8. We noted that the units of language are to be seen as ‘second-order’ constraints (see point (6) above). When we have developed this second-order ‘language stance’, our perception of utterance patterns has continuously been influenced by our literate culture, which has also given rise to a ‘written-language bias’ in linguistics (Linell 2005, 2012). Are the ‘units’ and ‘categories’ that are legion in this “stance” merely artefacts based on a misleading way of thinking about language? Perhaps not entirely.

31. Cf. the word-processor-based practices of cutting, moving and pasting words. The analogy with Lego toys has been used by Craib (1992) (quoted by Hopper 2011, 26), and others.

32. See Bybee (2010) on ‘rich memory’.

© 2013, John Benjamins Publishing Company
All rights reserved
References


All rights reserved


All rights reserved


### Appendix 1

**Abbreviations in glossings and formulas (in alphabetical order)**

<table>
<thead>
<tr>
<th>Adj</th>
<th>= adjective</th>
<th>Pro</th>
<th>= pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdjP</td>
<td>= adjectival phrase</td>
<td>PRT</td>
<td>= particle</td>
</tr>
<tr>
<td>Adv</td>
<td>= adverb(ial)</td>
<td>S</td>
<td>= sentence or clause</td>
</tr>
<tr>
<td>Aux</td>
<td>= modal auxiliary</td>
<td>Subj</td>
<td>= subjunction</td>
</tr>
<tr>
<td>Cop</td>
<td>= copula verb</td>
<td>V</td>
<td>= (main) verb</td>
</tr>
<tr>
<td>Inf</td>
<td>= infinitive</td>
<td>VP</td>
<td>= verb phrase</td>
</tr>
<tr>
<td>Imp</td>
<td>= imperative</td>
<td>V_{fin}</td>
<td>= finite verb</td>
</tr>
<tr>
<td>N</td>
<td>= noun</td>
<td>V1</td>
<td>= (finite) verb in clause-first position</td>
</tr>
<tr>
<td>NEG</td>
<td>= negation</td>
<td>V2</td>
<td>= (finite) verb in clause-second position</td>
</tr>
<tr>
<td>Nom</td>
<td>= nominative case</td>
<td>WhP</td>
<td>= phrase determined by Wh-word</td>
</tr>
<tr>
<td>NP</td>
<td>= noun phrase</td>
<td>1/2/3sg</td>
<td>= 1./2./3. Person singular.</td>
</tr>
<tr>
<td>PP</td>
<td>= prepositional phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prep</td>
<td>= preposition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>