17 Generative capacity and grammar formalisms

In several of the preceding chapters, the complexity hierarchy for formal languages was mentioned. The simplest languages are so-called regular languages (Type 3), they are followed by those described as context-free grammars (Type 2), then those grammars which are context-sensitive (Type 1) and finally we have unrestricted grammars (Type 0) that create recursively enumerable languages, which are the most complicated class. In creating theories, a conscious effort was made to use formal means that correspond to what one can actually observe in natural language. This led to the abandonment of unrestricted Transformational Grammar since this has generative power of Type 0 (see page 88). GPSG was deliberately designed in such a way as to be able to analyze just the context-free languages and not more. In the mid-80s, it was shown that natural languages have a higher complexity than context-free languages (Shieber 1985; Culy 1985). It is now assumed that so-called *mildly context sensitive* grammars are sufficient for analyzing natural languages. Researchers working on TAG are working on developing variants of TAG that fall into exactly this category. Similarly, it was shown for different variants of Stabler's *Minimalist Grammars* (see Section 4.6.4 and Stabler 2001, 2011) that they have a mildly context-sensitive capacity (Michaelis 2001). Peter Hellwig's Dependency Unification Grammar is also mildly context-sensitive (Hellwig 2003: 595). LFG and HPSG, as well as Chomsky's theory in *Aspects*, fall into the class of Type 0 languages (Berwick 1982; Johnson 1988). The question at this point is whether it is an ideal goal to find a descriptive language that has exactly the same power as the object it describes. Carl Pollard (1996a) once said that it would be odd to claim that certain theories in physics were not adequate simply because they make use of tools from mathematics that are too powerful.[1] It is not the descriptive language that should constrain the theory but rather the theory contains the restrictions that must hold for the objects in question. This is the view that Chomsky (1981b: 277, 280) takes. Also, see Berwick (1982: Section 4), Kaplan & Bresnan (1982: Section 8) on LFG and Johnson (1988: Section 3.5) on the *Off-Line Parsability Constraint* in LFG and attribute-value grammars in general.

There is of course a technical reason to look for a grammar with the lowest level of complexity possible: we know that it is easier for computers to processes grammars with

---

[1] If physicists required the formalism to constrain the theory

Editor: Professor Einstein, I'm afraid we can't accept this manuscript of yours on general relativity.

Einstein: Why? Are the equations wrong?

Editor: No, but we noticed that your differential equations are expressed in the first-order language of set theory. This is a totally unconstrained formalism! Why, you could have written down ANY set of differential equations! (Pollard 1996a)

lower complexity than more complex grammars. To get an idea about the complexity of a task the so-called 'worst case' for the relevant computations is determined, that is, it is determined how long a program needs for an input of a certain length in the least favorable case to get a result for a grammar from a certain class. This begs the question if the worst case is actually relevant. For example, some grammars that allow discontinuous constituents perform less favorably in the worst case than normal phrase structure grammars that only allow for combinations of continuous strings (Reape 1991: Section 8). As I have shown in Müller (2004c), a parser that builds up larger units starting from words (a bottom-up parser) is far less efficient when processing a grammar assuming a verb movement analysis than is the case for a bottom-up parser that allows for discontinuous constituents. This has to do with the fact that verb traces do not contribute any phonological material and a parser cannot locate them without further machinery. It is therefore assumed that a verb trace exists in every position in the string and in most cases these traces do not contribute to an analysis of the complete input. Since the verb trace is not specified with regard to its valence information, it can be combined with any material in the sentence, which results in an enormous computational load. On the other hand, if one allows discontinuous constituents, then one can do without verb traces and the computational load is thereby reduced. In the end, the analysis using discontinuous constituents was eventually discarded for linguistic reasons (Müller 2005b,c, 2007b, 2015b), however, the investigation of the parsing behavior of both grammars is still interesting as it shows that worst case properties are not always informative.

I will discuss another example of the fact that language-specific restrictions can restrict the complexity of a grammar: Gärtner & Michaelis (2007: Section 3.2) assume that Stabler's Minimalist Grammars (see Section 4.6.4) with extensions for late adjunction and extraposition are actually more powerful than mildly context-sensitive. If one bans extraction from adjuncts (Frey & Gärtner 2002: 46) and also assumes the Shortest Move Constraint (see footnote 30 on page 167), then one arrives at a grammar that is mildly-context sensitive (Gärtner & Michaelis 2007: 178). The same is true of grammars with the Shortest Move constraint and a constraint for extraction from specifiers.

Whether extraction takes places from a specifier or not depends on the organization of the particular grammar in question. In some grammars, all arguments are specifiers (Kratzer 1996: 120–123, also see Figure 18.4 on page 549). A ban on extraction from specifiers would imply that extraction out of arguments would be impossible. This is, of course, not true in general. Normally, subjects are treated as specifiers (also by Frey & Gärtner (2002: 44)). It is often claimed that subjects are islands for extraction (see Grewendorf 1989: 35, 41; G. Müller: 1996b: 220; 1998: 32, 163; Sabel 1999: 98; Fanselow 2001: 422). Several authors have noted, however, that extraction from subjects is possible in German (see Dürscheid 1989: 25; Haider 1993: 173; Pafel 1993; Fortmann 1996: 27; Suchsland 1997: 320; Vogel & Steinbach 1998: 87; Ballweg 1997: 2066; Müller 1999a: 100–101; De Kuthy 2002: 7). The following data are attested examples:

(1) a. [Von den übrigbleibenden Elementen]$_i$ scheinen [die Determinantien $\_i$]
      of   the left.over   elements   seem   the determinants

die wenigsten Klassifizierungsprobleme aufzuwerfen.[2]
the fewest   classification.problems   to.throw.up

'Of the remaining elements, the determinants seem to pose the fewest problems for classification.'

b. [Von den Gefangenen]$_i$ hatte eigentlich [keine $\_i$] die Nacht der   Bomben
   of the prisoners   had actually   none   the night of.the bombs
   überleben sollen.[3]
   survive   should

'None of the prisoners should have actually survived the night of the bombings.'

c. [Von der HVA]$_i$ hielten sich [etwa   120 Leute $\_i$] dort in ihren Gebäuden
   of the HVA held   REFL around 120 people   there in their buildings
   auf.[4]
   PRT

'Around 120 people from the HVA stayed there inside their buildings.'

d. [Aus dem „Englischen Theater"]$_i$ stehen [zwei Modelle $\_i$] in den Vitrinen.[5]
   from the   English   theater   stand   two models   in the cabinets

'Two models from the 'English Theater' are in the cabinets.'

e. [Aus   der Fraktion]$_i$ stimmten ihm [viele $\_i$] zu   darin,   dass die
   from the faction   agreed   him many   PRT there.in that the
   Kaufkraft   der Bürger gepäppelt werden müsse, nicht die gute Laune
   buying.power of.the citizens boosted   become must   not   the good mood
   der   Wirtschaft.[6]
   of.the economy

'Many of the fraction agreed with him that it is the buying power of citizens that needed to be increased, not the good spirits of the economy.'

f. [Vom Erzbischof Carl Theodor Freiherr von   Dalberg]$_i$ gibt es
   from archbishop Carl Theodor Freiherr from Dalberg   gives it
   beispielsweise [ein Bild $\_i$] im   Stadtarchiv.[7]
   for.example   a   picture   in.the city.archives

'For example, there is a picture of archbishop Carl Theodor Freiherr of Dalberg in the city archives.'

g. [Gegen die wegen   Ehebruchs zum Tod   durch Steinigen verurteilte
   against the because.of adultery   to.the death by   stoning   sentenced

---

[2] In the main text of Engel (1970: 102).

[3] Bernhard Schlink, *Der Vorleser*, Diogenes Taschenbuch 22953, Zürich: Diogenes Verlag, 1997, p. 102.

[4] Spiegel, 3/1999, p. 42.

[5] Frankfurter Rundschau, quoted from De Kuthy (2001: 52).

[6] taz, 16.10.2003, p. 5.

[7] Frankfurter Rundschau, quoted from De Kuthy (2002: 7).

Amina Lawal]$_i$ hat gestern   in Nigeria [der zweite Berufungsprozess _$_i$]
Amina Lawal   has yesterday in Nigeria the second appeal.process
begonnen.[8]
begun

'The second appeal process began yesterday against Amina Lawal, who was sentenced to death by stoning for adultery.'

h. [Gegen   diese Kahlschlagspolitik]$_i$ finden   derzeit   bundesweit
    against this   clear.cutting.politics happen at.the.moment statewide
[Proteste und Streiks _$_i$] statt.[9]
 protests and strikes   PRT

'At the moment, there are state-wide protests and strikes against this destructive politics.'

i. [Von den beiden, die hinzugestoßen sind], hat [einer _$_i$] eine Hacke, der
    of   the both   that joined   are has one   a   pickaxe the
andere einen Handkarren.[10]
other   a   handcart

'Of the two that joined, one had a pickaxe and the other a handcart.'

j. ein Plan, [gegen   den]$_i$ sich nun [ein Proteststurm   _$_i$] erhebt[11]
   a   plan against which REFL now a   storm.of.protests   rises

'a plan against which a storm of protests has now risen'

k. Dagegen$_i$ jedoch   regt sich jetzt [Widerstand _$_i$]: [...][12]
   against   however rises REFL now   resistance

'Resistance to this is now rising, however:'

l. [Aus der Radprofiszene]$_i$ kennt ihn [keiner   _$_i$] mehr[13]
    from the cycling.professional.scene knows him nobody   anymore

'Nobody from the professional cycling scene acts like they know him anymore.'

m. [Über das chinesische Programm der   Deutschen Welle] tobt dieser Tage
     about the Chinese   program of.the Deutsche Welle rages these days
[ein heftiger Streit   _$_i$].[14]
 a   hefty   controversy

'Recently, there has been considerable controversy about the Chinese program by the Deutsche Welle.'

---

[8] taz, 28.08.2003, p. 2.

[9] Streikaufruf, Universität Bremen, 03.12.2003, p. 1.

[10] Murakami Haruki, *Hard-boiled Wonderland und das Ende der Welt*, suhrkamp taschenbuch, 3197, 2000, Translation by Annelie Ortmanns and Jürgen Stalph, p. 414.

[11] taz, 30.12.2004, p. 6.

[12] taz, 02.09.2005, p. 18.

[13] taz, 04.07.2005, p. 5.

[14] taz, 21.10.2008, p. 12.

This means that a ban on extraction from specifiers cannot hold for German. As such, it cannot be true for all languages.

We have a situation that is similar to the one with discontinuous constituents: since it is not possible to integrate the ban on extraction discussed here into the grammar formalism, it is more powerful than what is required for describing natural language. However, the restrictions in actual grammars – in this case, the restrictions on extraction from specifiers in the relevant languages – ensure that the respective language-specific grammars have a mildly context-sensitive capacity.