The aim of this chapter is to deal with “hard” syntax, that is, the aspects of the forms of the utterances in a given language that can be described by hard rules. Within the structure of the proposed theory of production and interpretation, hard syntax plays two roles.

Its primary role is to be part of the account of production, that is, of the process speakers go through when they produce an utterance on the basis of some communicative intention. While any account of production that would not take into account the speaker strategy of producing only utterances that have the intended interpretation as the most probable one in the context would be seriously incomplete, utterances must also satisfy certain hard syntactic rules. The focus of the chapter is on the part of the production process where utterances receive their morphology and word order. In this area, which one could, following a similar usage in the field of natural language generation, call syntactic realisation, there are important differences between languages. For a proper model of production, in addition to syntactic realisation one also needs higher-level generation, the area where the inputs for syntactic realisation are constructed and lexicalised. In addition, the further constraints on utterances developed in Chapter 3 Self-Monitoring need to be incorporated.

The second aim of this chapter is to deal with hard syntax as part of interpretation. Bayesian interpretation incorporates the production probability of an utterance given a hypothesis. It is assumed to be estimated by a simulation of the utterance under the hypothesis. It is further assumed that this Bayesian interpretation is incremental, which requires that the partial hypotheses arising from initial fragments of the utterance to be evaluated with respect to their ability to cause the observed initial fragment. Higher-level generation is much less involved in these evaluations and the considerations of Chapter 3 Self-Monitoring not at all: monitoring defines the adaptation to the bias for the most probable interpretation that is characteristic of the hearer strategy. It does not need to be checked again in interpretation.

Syntactic realisation—possibly with a fragment of higher-level generation—is then what is shared by production and interpretation. The main result of this chapter is that production can be done with a quasi-linear
algorithm based on a specification of a plan to achieve a communicative task and that this same algorithm can also check whether the initial segments of the sentence are syntactically correct given a partial hypothesis about what they want to achieve. The first goal, the algorithm, needs to be reached in order to show that production can be effortless and instantaneous, the second goal, checking initial fragments of interpretation, again, is important for using production as an efficient filter in an incremental interpretation algorithm, especially such as the one defined in Chapter 4 Interpretation. If production were not linear, the interpretation algorithm also would not be linear. Human interpretation is incremental (as originally shown by Crain and Steedman (1985)). The main aim of this chapter is to show that there exists a production algorithm that has this combination of properties.

Many—perhaps even all—proposals for syntactic description may turn out to allow for a generation algorithm that has both of these properties. The primary reason for using a restricted version of optimality-theoretic syntax as below is that the generation algorithm is simple and the demonstration of both properties is straightforward, in the second case even trivial. For other frameworks, it would be necessary to develop a concept of production to the extent that these properties could be proven and, to the best of my knowledge, such proofs are not available.

This chapter is organised as follows. Section 2.1 gives a minimal introduction to optimality theory. In subsection 2.2.1, a simplified optimality-theoretic grammar is given for word order in a Dutch clause. This is then generalised in subsection 2.2.2 to a treatment of German and English, demonstrating that the constraint set for Dutch has some degree of typological validity. In section 2.3, it is shown that the particular set of constraints specifies a linear production algorithm. Section 2.4 demonstrates how the grammar can be extended to morphology and higher-level generation, and thereby lead from a task specification to the string. Section 2.5.3 discusses the checking of initial fragments.

The optimality-theoretic grammars are important for Chapter 3 Self-Monitoring because they also form the basis for the requisite proofs which

---

1 Joan Bresnan's optimality theoretic LFG (Bresnan, 2000; 2001) is a successful reformulation of LFG in optimality theory. The account of Dutch, German, and English below is strongly influenced by Gazdar et al. (1985) and, with some charity, could be seen as a translation of Generalised Phrase Structure Grammar into optimality theory. The treatment of Grimshaw (1997) is optimality-theoretic syntax with much influence from Government and Binding. There is no principled reason why such translations into optimality theory could not be achieved also for other formalisms.