1 Introduction

Background and Optimality Theory

In this first chapter, the basic notions and theoretical issues that will be referred to throughout the book are introduced: first, a definitional description of Standard German, the language investigated in the book; second, a review of the constituents of the prosodic hierarchy that constitute the backbone of the domains of phonological alternations and processes, as well as their recursive structure; third, Optimality Theory, the theoretical framework used in the book, is described minimally, for readers who do not know this approach well. The chapter also contains a short description of the structure of the remainder of the book, as well as an overview of the conventions used. It ends with some remarks about what is not part of the book.

1.1 Standard German

Nowadays, at least in Germany, Standard German is used in all administrative and public aspects of life, and it is taught in school. Standard German (or *Hochdeutsch* 'High German') finds its origin in Martin Luther's translation of the Bible that was first published in the first part of the sixteenth century. Luther's written German was mainly based on the Higher Saxon dialect of the time, a Central German dialect that was thought to be understandable by speakers of both the North and the South of Germany. Until the beginning of the nineteenth century, Standard German was almost exclusively a written language. The pronunciation of this dialect slowly became standardized and adapted in different varieties. Nowadays the long-standing standard pronunciation called *Bühnensprache* 'stage language' has established normative pronunciation rules. Standard German is the official dialect of German, even though it is de facto a collection of highly standardized varieties of the language: It is present not only in Germany but also in Austria, Switzerland, Luxembourg and Lichtenstein as well as in German-speaking regions of Belgium, Italy and France. Even if highly standardized, Standard German still consists of three main varieties, German, Austrian and Swiss, that differ in orthography, vocabulary and, most importantly for this book, pronunciation.

For many speakers of a regional variety of German, learning Standard German still amounts to learning a foreign language, especially in Switzerland.

In this book, the version of Duden, volume 6, *Das Aussprachewörterbuch* 'The Pronunciation Dictionary' that was published in 2005 is considered authoritative. Two more pronunciation dictionaries are also important sources of transcription: the *Deutsches Aussprachewörterbuch* (DAWB, Krech et al. 2009), and the *Großes Wörterbuch der deutschen Aussprache* (GWDA, Krech et al. 1982), both first edited in the eastern part of Germany. In some cases, Wiktionary was also consulted (https://en.wiktionary.org/wiki).

1.2 Prosodic Hierarchy

The following chapters cover different aspects of phonology, from the distinctive features of individual sounds to the larger intonational structure of sentences. German is an ideal language for the study of phonology due to its exemplary phonological patterns in different areas. It has well-known segmental alternations such as Final Devoicing and umlauting, the process of fronting a back vowel. It has allophonic relations between sounds, such as the one involving palatal and velar dorsal fricatives (i.e., the so-called *ich*-Laut and *ach*-Laut). It has a crisply clear syllabification across foot boundaries and larger prosodic domains, partly based on the morphemic structure of words and partly based on their segmental composition and signaled by different segmental alternations and glottal stop insertion. It has tense and lax vowels that are largely in complementary distribution in syllables, but not quite. It has a default vowel – schwa – with an interesting distribution and that is epenthetic in only very few contexts. It has an unexpected defective distribution of [s] in the native vocabulary. And it has lexical stress. Most of the phonological facts discussed in this book are well known and have been discussed numerous times in the relevant literature. The singularity of the present work lies in its global approach, establishing a connection across all aspects of phonology, from the small to the large ones. To achieve this aim, the prosodic hierarchy, see (1), is extensively used, as first proposed by Nespor and Vogel (1986) and Selkirk (1984) and also used by Elfner (2012, 2015), Elordieta (2015), Ishihara (2022), Ito and Mester (2013), McCarthy and Prince (1986, 1993a,b), Myrberg (2013), Truckenbrodt (2006) and many others.

(1) Prosodic hierarchy

ι-phrase	intonation phrase	(roughly corresponds to a clause)
φ-phrase	prosodic phrase	(roughly corresponds to a syntactic phrase)
ω-word	prosodic word	(roughly corresponds to a grammatical word)
F	foot	(metrical unit: trochee, iamb …)
σ	syllable	(strings of segments: CV, CVC, …)
μ	mora	(unit of syllable weight)

Most phonological processes, allophonies and alternations, as well as syllabification, stress and tone assignment, take place in the prosodic hierarchy's constituents, in agreement with the "Indirect Reference Hypothesis" that embodies the intuition that phonology is a separate module of grammar, and that phonology applies exclusively in phonological domains. The specific assumption underlying this book is that the spoken and the conceptual parts of language are related yet separate domains of investigation. The main role of phonology is to study the body's oral outputs in language, that is, their manifestations in spoken speech.[1] The constituents of the prosodic hierarchy – mora, syllable, foot, prosodic word, prosodic phrase and intonation phrase – are articulatory gestures, segments, syllables, stress patterns, constituents' weight units, tones, and the like, all elements that have no conceptual content by themselves and that need to be mapped to the meaningful part of language (i.e., morphology, syntax and semantics). As an example, distinctive features define segments as articulatory, perceptual or acoustic objects or events, but they have no relation to the conceptual part of language: [m] or [o] or [ʔ] have no *intrinsic* meaning, although they can be meaningful in some languages, and the same holds for the abstract constituents of the prosodic hierarchy, such as syllables, feet. This separation is evident and uncontroversial in features and segments, but it is also present in all phonological aspects of language. In the same way as distinctive features realize segments (or segments are defined by distinctive features), sequences of segments make up syllables, according to phonotactic principles, but not according to syntactic or semantic ones. The phonological *form* of larger constituents is in principle also devoid of any meaning. Only so-called interface-mapping principles or constraints between prosodic constituents and morphosyntactic words and phrases establish a connection between phonological objects and the conceptual part of language.

The prosodic constituents are divided into two types. First, the lower-level constituents – moras, syllables and feet – have no relation whatsoever to morphosyntactic constituents. They are sometimes called "rhythmic" domains. Inside syllables, some segments are moraic and some others are non-moraic. Syllables are made up of sequences of vowels and consonants that are linearly organized in specific orders according to their sonority value. The moraic status of syllables comes from the segments they consist of and that define their weight. Feet consist of syllables or moras that have more or less weight and strength, partly relative to neighboring syllables. The second type of prosodic constituents are the higher-level constituents, the "interface" constituents that are mapped to morphosyntactic constituents. Prosodic words, abbreviated as

[1] I am well aware that spoken speech is only one of several of the body's possible outputs related to language – sign language and facial or corporal gestures being others. This book concentrates exclusively on spoken language.

ω-words, are primarily defined by their phonological properties: They have a single primary stress and a clearly defined syllabification pattern. ω-words have crisp syllable boundaries at their edges and may have blurred ones medially. The highest levels of the prosodic hierarchy, prosodic phrase and intonation phrase, abbreviated as *φ-phrase* and *ι-phrase* respectively, are also defined by their phonological correlates, such as primary accent and tonal structure. These higher prosodic levels have further properties, such as a rhythmic organization, avoidance of stress clashes, a special melodic pattern, but again, they have no intrinsic meaning as long as they are not associated with a syntactic and semantic content. This is not to say that the interface between prosodic constituents and the morphosyntactic part of grammar is unimportant. On the contrary, to make sense at all, and to be a part of language, the relationship between the phonological component and the conceptual components must be made explicit. This relationship, called *la double articulation du langage* by Martinet (1961) and the "dual pattern of language" by Hockett (1960), is fundamental to all aspects of phonology.

Throughout the book, it will be taken for granted that, at least from the ω-word on, the higher prosodic domains in (1) are recursive, as a consequence of the recursive morphosyntactic structure of German and of well-formedness conditions on prosodic structure. Recursivity characterizes the fact that a category of level *n* can dominate a constituent of the same category *n*. It is left open whether the lower constituents, especially the syllable and the foot, can be recursive as well. The recursivity of prosodic constituents has been denied in some influential prosodic models (Nespor & Vogel 1986, Selkirk 1984, 1995) but has been largely recognized in others (Ladd 1986, Hayes 1990). The min-max model adapted from Ito and Mester (2013) and reproduced in Figure 1.1 is assumed here for the higher prosodic domains, but not for the syllable and the foot, at least not in German. It is illustrated with ω-words and φ-phrases. Here a ω-word dominates a ω-word, and a φ-phrase dominates a φ-phrase. At each level of the tree in Figure 1.1, adjunction of further prosodic material is possible, and it is the adjunction operation that motivates recursivity. The optional adjoined material is noted as X in the trees.[2]

An important property of the min-max model is that it distinguishes between different levels of the same category: A ω^{min} can have different properties from a ω^{max}. The domination relationship between these constituents illustrated in Figure 1.1 is rendered explicit in (2) for the prosodic word and the prosodic phrase.

[2] Ito & Mester (2021) make a further distinction between "coordination" and "adjunction." Coordination is a balanced structure: two elements of the same kind come together to form a larger constituent. In German, however, in such a prosodic coordination, one element is always the head of the larger constituent, the constituent to which a subordinated one is adjoined. This implies that adjunction is a cover term for both recursive patterns.

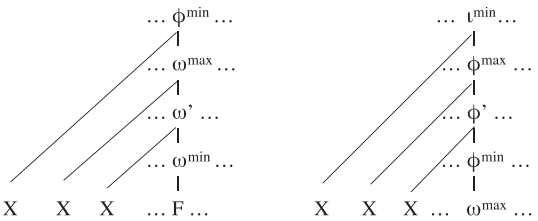

Figure 1.1 Recursivity of the higher prosodic domains, adapted from Ito and Mester (2013: 22). The two writings of phi are equivalent.

(2) Maximal ω (ω^{max}): ω not dominated by ω.
Minimal ω (ω^{min}): ω not dominating ω.
ω': ω dominated by ω and dominating ω.

A maximal ω-word (level *n*) is dominated by a minimal φ-phrase (level *n+1*). According to this representation, only one ω-word can be minimal and only one can be maximal. The third variant, intermediate ω', can appear an indefinite number of times.

The Indirect Reference Hypothesis described in this section is not uncontroversial. Some researchers defend the opposite view, called "Direct Reference Hypothesis," that assumes that domain-sensitive phonological processes can be defined solely with reference to morphosyntactic structure; see Adger (2007), Kahnemuyipour (2009), Dobashi (2010), Pak (2008) and Seidl (2001) among other authors. Such relationships are then defined in terms of syntactic structural relationships such as c-command (Kaisse 1985) or through the presence of intervening nodes between words, which may trigger cyclic or phrasal spell-out (Newell 2008, Newell & Piggott 2014). Some authors deny the existence of higher-level prosodic constituents altogether and consider prosodic structure as inherent syntactic structure. The argument usually invoked for the denial of prosodic structure is the independent and better-motivated syntactic structure that should be enough to explain the prosodic structure as well. However, prosodic and syntactic structure are not identical, as will be demonstrated extensively in this book.

1.3 Optimality Theory

Optimality Theory (OT) is used as the theoretical framework of the entire book. The main differences between OT and earlier grammatical models are the explicit conflict resolution component of OT (Prince & Smolensky 1993/2004) and the obligatory ordered rules component of derivational approaches. Optimality Theory makes the fundamental claim that no linguistic entity, no syllable, no word, no sentence, manages to satisfy

all the requirements imposed by the principles or rules of grammar. It views Universal Grammar as consisting of a set of principles, called constraints, which express universal linguistic tendencies present in all languages. These principles are simple and can be formulated as general statements. They may conflict with each other when imposing incompatible demands on specific linguistic entities. Individual grammars must resolve these conflicts, and they do so by ranking the constraints. A further basic insight of OT is that even if grammars are driven by the same principles, these principles are ranked in different ways in different languages. As a result, the same conflict may be resolved in different ways in different languages. A constraint A may be ranked very high in some language L1, so that grammatical outputs fulfill A, and ranked lower in another language L2, where constraint A is crucially dominated by another constraint B, conflicting with A, to the effect that linguistic outputs fulfill B and violate A. In such a case, A and B conflict with each other and the conflict is resolved differently in L1 and L2. In the OT literature, this kind of conflict is usually visualized by means of so-called tableaux.

Tableau 1.1 shows the ranking of A and B in L1, where A is ranked higher than B. The constraints needed for the evaluation appear at the top of the columns, here only A and B. Their order from left to right corresponds to their weighting in the evaluation process. The ranking is expressed by the ordering of the two constraints: The leftmost one is higher ranking than the following one. In OT one speaks of "dominance": constraint A dominates constraint B (A ≫ B). Suppose now that several candidates compete for the best output, two of which are shown in the tableau, Candidate a and Candidate b. In the OT tableaux of this book, only the most promising candidates are compared to each other. In a complete grammar, many more constraints will also play a role, but this is ignored in Tableaux 1.1 and 1.2. There will also be other candidates participating in the competition that violate or fulfill both constraints, or which violate the constraints more than Candidates a and b. In Tableau 1.1, Candidate a fulfills A but violates B and Candidate b violates A and fulfills B. Violation of constraints by candidates is shown by an asterisk in the corresponding cell. The finger points to the optimal output, the winner of the competition among the candidates.

Tableau 1.1 *Ranking of constraints A and B in L1: A ≫ B*

Input	A	B
☞ a. $Cand_1$		*
b. $Cand_2$	*	

Tableau 1.2 *Ranking of constraints A and B in L2: B ≫ A*

Input	A	B
a. $Cand_1$	*	
☞ b. $Cand_2$		*

L2 is illustrated in Tableau 1.2. The ordering of the two constraints is reversed: B dominates A (B ≫ A) and, as a result, Candidate b is the winner, i.e., the candidate fulfilling the highest constraint. Thus, even if both linguistic principles expressed by constraints A and B are present in both languages, OT predicts that their ranking has an influence on the choice of the best candidate in each language.

The input (or underlying candidate) sits in the upper left-hand cell. It may be conceived as a kind of underlying representation. However, input and underlying representation differ from each other in a crucial way. The principle *Richness of the Base* (Prince & Smolensky 1993/2004) guarantees that the input can take any form it wants, as long as it is a linguistic entity: It can have little structure, possibly underspecified, but it can also be completely specified. Because of *Richness of the Base*, no constraint can limit the form of possible inputs.

OT operates in two steps. First, a set of candidates for the input is generated by a function *Gen* (for *Generation*), the generative part of the grammar, as in (3). The function *Gen* delivers for each input a certain number, possibly an infinite number, of candidates.

In a second step, the candidates are evaluated and compared among each other with the help of a function *Eval* (for *Evaluation*), as in (4). The evaluation determines which candidate is chosen as optimal (out_{real}) and is thus the grammatical and realized output.

(3) *Gen* (in_1) → [Cand a, Cand b, …]

(4) *Eval* ([Cand a, Cand b, …]) = out_{real}

Gen and *Eval* apply in a sequence. As for *Gen*, according to a principle called *Freedom of Analysis*, every thinkable structure is a possible candidate for a specific input. In other words, *Richness of the Base* ensures that an input can have all sorts of forms, and *Freedom of Analysis* ensures that candidates can have all sorts of forms. In all following tableaux, we will only see a small number of relevant candidates for each input, even though an infinite number of candidates may be generated. Most of the candidates are eliminated by high-ranking constraints imposing some level of correspondence (or faithfulness) between the input and the output.

Whether a candidate emerges as the optimal output of a competition depends on the second step of OT. *Eval* consists of a set of constraints by which candidates are evaluated. The function *Eval* in (4) assesses each element of the set of candidates with the help of ranked constraints, as illustrated in Tableaux 1.1 and 1.2.

The optimal candidate is the result of the evaluation. In any evaluation, the winning candidate C_{winner} is the real output (out_{real}), that is, the candidate with the best violation profile such that it violates the constraint hierarchy less than all other competitors in the evaluation set; see (5) for a formal definition. "Less" means that there is no candidate C_{loser} which does better than C_{winner} with respect to the highest constraint on which they differ.

(5) *Best fulfillment* (or *minimal violation*)
A candidate C generated by *Gen* from input in_i is grammatical (out_{real}) iff all candidates C', C' ∈ *Gen*(i), it competes with are such that C violates the highest constraint c from *Eval* on which C and C' differ less often than C' does.

The main principles of *Eval* in OT are universality, violability, ranking and parallelism.

Universality: Universal Grammar consists of a universal set of constraints. All constraints are present in the grammar of each language. A group of constraints care for faithfulness of the outputs to the input. Another group of constraints care for the best phonological form of the output. These latter "markedness" constraints are empirically motivated by independent principles, such as ease of articulation, perceptual contrastivity, typological generalizations, economy or analogy. Depending on the ranking of all of them, the constraints decide which candidate wins in each competition.

Violability: All constraints are violable, but violations are minimal. Violability is a prerequisite for universality, since not all constraints are fulfilled in all languages. Ideally, the best candidate fulfills all constraints although such a candidate does not exist. All outputs violate some constraints, such as those militating against any form of structure and phonetic content. Control over the violations is reached through the notion of *best fulfillment* or *minimal violation*, formulated in (5). The candidate that best fulfills the constraint system or that violates it minimally is called *optimal*, and it is by definition the output chosen by the grammar.

Ranking: Constraints are hierarchically ordered, and minimal violation is determined based on this ranking. Low-ranking constraints can be violated to achieve fulfillment of higher-ranked constraints.

The grammar of each individual language defines a unique ranking of the universal constraints.

Parallelism: The best fulfillment of the constraint hierarchy is determined for each input over the entire constraint hierarchy in a parallel evaluation of all candidates. The most important consequence of parallelism is that all influences exerted on a structure are evaluated at once, for instance strictly phonological influences, such as assimilation, morphological influences, such as morpheme boundaries, prosodic influences, such as the place of a segment in the syllable or in the metrical foot.

Prince and Smolensky also discussed (and rejected) Serial Harmony, a derivational evaluation by which, instead of a single pass through *Gen* and *Eval*, changes are made one at a time. *Harmonic Serialism* has been reintroduced by Kiparsky (2015), Bermúdez-Otero (2018) and McCarthy (2000, 2010), among others, as a variant OT that combines both optimization and derivation. A first candidate is identified by an OT evaluation, and then a second one undergoes a new pass through *Gen* and *Eval* on the basis of the first optimal candidate and the same constraint ranking, and so on, until a definitive optimal surface form emerges. Each evaluation is restricted to one phonological operation. Harmonic Serialism captures phenomena that have proven difficult for classic Optimality Theory grammars, such as feature spreading and opaque interactions. In this book, parallelism is used wherever possible.

Tableau 1.3 illustrates four ranked constraints and four candidates. Candidates b to d have an exclamation mark at the place where they are eliminated. Candidate b is eliminated by constraint A, Candidate c by constraint B, and Candidate d by constraint C. Candidate a fulfills constraints A to C and violates the lowest-ranking constraint D. It is nevertheless the optimal candidate according to (5).

Summing up the conventions for reading the tableaux introduced so far:

- Left-to-right columns correspond to the dominance ranking of the constraints.
- Fulfillment of a constraint is shown by an empty cell.

Tableau 1.3 *Four ranked constraints and four candidates*

Input	A	B	C	D
☞ a. $Cand_1$				*
b. $Cand_2$	*!			
c. $Cand_3$		*!		
d. $Cand_4$			*!	

- Violation is illustrated with an asterisk *. The exclamation mark "!" shows the position where a candidate has a fatal constraint violation. At this place it reaches non-optimality.
- The symbol ☞ points to the optimal candidate.

There are three additional cases in which two constraints A and B deliver diverging results. In the first case, shown in Tableau 1.4, constraint A is fulfilled by both candidates and constraint B is violated by the first candidate but not by the second one.

Second, in Tableau 1.5, constraint A realizes a tie, as it is violated by both candidates. In this case also, the lower-ranking constraint B decides between them. This illustrates a property of the theory: The violation of a constraint is never fatal per se. It is only fatal if there is a better candidate (i.e., a candidate that does not violate it).

In the third case, shown in Tableau 1.6, both candidates violate the same constraint, but Candidate b violates the constraint more often than Candidate a. This is a case of multiple violation of a constraint by a candidate. The number of violations is decisive in such a case.

Optionality, gradience and variation are important components in all grammars. In OT this implies the possibility of having more than a single winner.

Tableau 1.4 *Constraint A is fulfilled by both candidates*

Input	A	B
a. $Cand_1$		*!
☞ b. $Cand_2$		

Tableau 1.5 *Constraint A is violated by both candidates*

Input	A	B
a. $Cand_1$	*	*!
☞ b. $Cand_2$	*	

Tableau 1.6 *Candidate b violates constraint A more often than Candidate a*

Input	A	B
☞ a. $Cand_1$	*	*
b. $Cand^2$	**!	