1

Word order: setting the scene

Many students of foreign languages will agree that one of the first things that they discover (on their own) about their target languages concerns word order, that is, the way words are arranged in a particular linear order in sentences. Word order may be one of the most conspicuous differences between one's native language and target languages. For instance, Korean learners of English quickly realize that while the verb in Korean comes last in a sentence, the verb in English comes much earlier, that is, after the subject noun phrase (NP). In fact, one of the English grammar-translation rules often imparted to Korean students is about word order. For example, consider the following Korean sentence and its English translation:

(1) mayli-ka   kongwen-eyse thom-ul   man-ass-ta
    Mary-NOM   park-LOC   Tom-ACC   meet-PST-IND
    'Mary met Tom in the park.'

The Korean-to-English (or English-to-Korean) 'word-order rule' in question consists of two steps (setting aside morphological differences between the two languages, including core case marking): (1) translate the subject NP and put it in the first position of the sentence; and (2) go to the end of the sentence and translate the remaining constituents in reverse order (i.e. X-Y-Z $\Rightarrow$ Z-Y-X). This means that the subject NP (*Mary*) comes first, as in the original Korean sentence. From that point on, however, the Korean-to-English translation proceeds from right to left. The subject NP should thus be immediately followed by the verb *met* for *man-ass-ta*. Next comes the object NP (*Tom* for *thom(-ul)*). These three constituents should all be followed by the English word *in*, translated from the Korean (non-core) locative case marker *-eyse*. Finally, the locative NP (*the park*) for *kongwen* appears as the last constituent of the sentence. That is to say *Mary*, *met*, *Tom*, *in* and *the park*, in that order, appear one after the other in the (grammatical) English sentence. Needless to say, this is just a rule of thumb, and a rough one at that, but it does provide an important insight into the word-order

1

differences between the two languages. Thus in Korean, the verb appears last in the sentence, whereas the verb in English appears after the subject NP; the object NP appears before the verb in Korean, whereas this ordering is reversed in English, that is, the object NP after the verb; the locative case marker appears immediately after the related NP in Korean, whereas its English counterpart appears before the related NP; and the locative expression as a whole appears in between the subject NP and the object NP in Korean, whereas in English it appears last in the sentence or after the object NP. Technically speaking, Korean is a Subject-(X-)Object-Verb language, whereas English is a Subject-Verb-Object(-X) language (where X represents an adverbial phrase); Korean has NP-Postposition order, whereas English has Preposition-NP order. Word order is one of the defining grammatical properties of Korean and English, and indeed, of the world's languages (to varying degrees).

In point of fact, words or constituents have no option but to linearize themselves because they cannot all together appear in a non-linear manner, that is, articulated simultaneously. This absolute linear requirement relates ultimately to the physical or temporal constraint imposed on human language, that is, the physics of speech. Thoughts or ideas can only be transmitted linguistically from the speaker to the hearer in this temporally constrained manner. Languages do not rely on telepathic means, for instance, as attributed to alien communication in some science-fiction movies. Moreover, different languages select different options of linearization (within certain limits, as will be shown throughout the rest of the book) – for example, verb-final, verb-middle or verb-initial order. Thus Korean and English, as demonstrated above, have different (or, one might say, almost opposite) ordering rules or conventions. This observation naturally leads to a number of questions about word order.[1] For instance, is it possible to explain the kind of word-order difference between English and Korean in a revealing way? If so, how? Is the choice between prepositions and postpositions, for example, an accident or a matter of chance? Or is that choice dictated by other (structural) factors? More general, but no less intriguing, questions can also be asked – for example: what is the word-order variation in the world's languages like? Is that cross-linguistic variation random or systematic? Are some word orders more or less common in the world's languages than others? If so, why? Is it possible to place constraints on the word-order variation and also to invoke a 'deep' principle(s) to explain those constraints? As will be explored in the rest of this book, these and many other intriguing questions come to the fore, demanding a great deal of scholarly attention.

Indeed it hardly comes as a surprise that word order, for its conspicuousness, is one of the long-standing research topics in linguistics. Scholarly interest in word order goes back a long way; one need only point to some of the early classic works, such as Behaghel (1909/10, 1930) and Schmidt (1926); historians of linguistics (e.g. Graffi 2011) can no doubt refer us to many other works preceding these classics of word-order research. Perhaps the most telling, in the context of the history of linguistics, is the fact that the revitalization, in

the 1960s, of Linguistic Typology as a viable theory of language (Greenberg 1963) was carried out on the back of word-order research (e.g. Comrie 1989: 86). Moreover, the monumental project entitled *The World Atlas of Language Structures* (Haspelmath *et al.* 2005) devotes many more chapters and maps to word order than to any other structural properties (17 out of the 142 chapters, or almost 12 per cent of the Atlas).

Many current linguistic theories address word order, albeit in different ways. Word order-related questions, such as those raised above, have been of such importance to linguistic theory because, as will be shown below, word order (variation) has to be accounted for, whether at the outset or the end of theorizing. There is no avoiding it, because it has to be represented in phonetic outputs, that is, pronounced sentences. The corollary of multiple approaches to word order is a vast literature. Unfortunately, practitioners of one approach tend to be unfamiliar with, if not unaware of, the work conducted in other approaches, although there are a few notable exceptions (e.g. Siewierska 1988, 1997; Hawkins 1994, 2004; Cinque 1996, 2000, 2005, 2009, 2010). Take the present writer, for instance. Largely because of his training in linguistics, he initially had a much better understanding of Linguistic Typology than other theoretical approaches. Needless to say, this situation is most unfortunate for linguistics as a discipline, and also for both researchers and students, who do not (have an opportunity to) expose themselves to different issues of theoretical import, not to mention different perspectives, because they are not familiar with the other theories. Thus while a synthesis of different theories of word order may be beyond the reach of books like this one, it is certainly worth bringing some of the major theories of word order together in one place.

The present book is concerned, as its primary theme, with word-order patterns, at different structural levels, in the world's languages, and also various correlations attested among those patterns. Moreover, the book surveys some of the major theoretical approaches to word order by discussing and evaluating their core theoretical assumptions about, and explanatory principles of, word order, and, wherever possible, by comparing them with each other. Thus various important theoretical issues and problems that have been raised with respect to word order are addressed with an eye to securing a better understanding of word orders and their correlations. To wit, this book aims to provide a critical overview of major theories of word order and the results of their current word-order research. In so doing, the book also attempts to draw the reader's attention to some of the issues that remain open to further investigation, and the problems that remain unresolved.

The theoretical approaches selected for inclusion in the book are: Linguistic Typology, Generative Grammar (i.e. Government and Binding Theory, Principles and Parameters Theory and the Minimalist Program), Optimality Theory, and processing-based theories (e.g. Early Immediate Constituent Theory and its revised/expanded version, and Dependency Locality Theory/ Syntactic Prediction Locality Theory). There are reasons, apart from space

limitations, why these, and not other, theories have been chosen for inclusion in the book.

First, these theories occupy prominent places in linguistics (and in other cognate disciplines as well). Thus those who wish to read about current word-order research will first need to familiarize themselves with what these theories have to say about word order (before moving on to other, perhaps less well-known, ones).

Second, the status and role of word order may differ among theories. This difference seems to be better demonstrated by the selected theories than by others. For instance, in Generative Grammar, word order is not as important an issue as in Linguistic Typology or processing-based theories. Within Generative Grammar, word order may be largely a matter of phonology, that is, surface realization or linearization of (far more important) representations already generated elsewhere in the grammar (i.e. syntactic and logico-semantic components). Be that as it may, word order (or linearization), in Generative Grammar, is something that needs to be stipulated in grammars or to be derived from, or reduced to, principles, axioms or algorithms. In contrast, Linguistic Typology, because of its empirical or inductive orientation, treats surface word orders as something important to be accounted for in their own right (e.g. not to be derived from some abstract representations). Moreover, the very concept of abstract representations, let alone individual abstract representations, is not verifiable or acceptable to every linguist (and may even be antithetical to some theories, including Linguistic Typology). Far more frequently than not, such abstract representations or structures stem directly from (highly) theory-internal considerations. Theory-internal considerations or arguments lose their force or validity outside their respective theories, however. Positioned somewhere between these two 'poles' is Optimality Theory. Because of its theoretical adoption of optimal outputs (probably something inherited from its place of birth, that is, phonology), Optimality Theory pays a greater deal of attention to surface (word-order) properties than Generative Grammar does. Moreover, because of its strict adherence to language-particular constraint rankings, Optimality Theory also strives to account for the cross-linguistic variation (on word order) in a relatively robust manner. (Otherwise, language-particular ranking constraints could not be worked out.) As will be shown in Chapter 5, unfortunately, this strength of Optimality Theory seems to be somewhat compromised, primarily because of its inheritance of some of Generative Grammar's assumptions and analytical means. Thus it will be an interesting exercise to find out how these different theories attempt to handle word orders and their correlations within their theoretical confines.

Third, word order has become important for linguistic theorizing on a more conceptual level as well. Even if word order itself may not be an issue of theoretical priority or importance in a particular syntactic theory (e.g. Relational Grammar, Role and Reference Grammar), it will have to be dealt with in one way or another. Put differently, once issues of theoretical priority have been

addressed, the question of how to linearize words and constituents (that is, so as to produce phonetic outputs) will arise eventually. This, in turn, will demand serious, careful theoretical thinking, because at this point, efforts will have to be directed towards preventing the strongly upheld theoretical assumptions or constructs from being jeopardized or undermined (e.g. a universal, underlying word order). In contrast, theories that dispense with abstract representations or structures altogether, by dealing directly with surface word-order properties (e.g. Linguistic Typology), will eventually need to formulate a principle(s) to explain attested (as well as unattested) word orders and their correlations (e.g. why is X always found to co-occur with Y, but never with Z?). To wit, irrespective of which direction word order is approached from, the two issues – that is, empirical validity (= word-order patterns) and theoretical explanation (= word-order principles) – need to be addressed. With this understanding in mind, the theories to be surveyed in this book have been selected with a view to demonstrating the two directions from which to approach word order as an object of inquiry.

The foregoing considerations have a direct bearing on how the remaining chapters of this book will be ordered (no pun intended!). The Linguistic-Typological approach will be the first to be taken up because of its unequalled emphasis on empirically (or statistically) tested word-order variation and correlations. In other words, no other theories can deliver a better empirical coverage of word order than Linguistic Typology can. Moreover, Optimality Theory and processing-based theories (especially Early Immediate Constituent Theory and its expanded version) draw heavily on the data amassed by Linguistic Typology. There is the need to be familiar with cross-linguistic word-order variation and correlations before examining how they have been dealt with by Optimality Theory and processing-based theories. This will also enable the reader to understand how the word-order data may have shaped these theories themselves. For these reasons, the Linguistic-Typological approach to word order will be the topic of Chapter 2. This chapter will be followed by a kind of 'entr'acte', in Chapter 3, on the historical and conceptual background of Generative Grammar (see below for justification). The Generative Grammar approach will be the topic of Chapter 4, and the Optimality-Theoretic approach the topic of Chapter 5. Many specific formal constraints and representations utilized in Optimality Theory are imported directly from Generative Grammar. Thus it makes much sense to read the chapter on the Optimality-Theoretic approach after the chapter on the generative approach (as well as, if required, the entr'acte chapter). Processing-based approaches to word order, including Early Immediate Constituent Theory and its latest development, as well as Dependency Locality Theory/Syntactic Prediction Locality Theory, will be discussed in Chapter 6. Once again, there is a logistic reason for placing these processing-based theories after the others. The processing-based theories to be surveyed in this book also tend to assume some of the (least controversial) theoretical assumptions and representations from Generative Grammar – for example, basic (i.e. not too abstract) constituent structure, assignment of case relations and thematic roles.

More importantly, unlike the other theories, processing-based theories bring directly into the equation the speech-act participants (i.e. the hearer and, to a lesser extent, the speaker) and their processing/production needs. These factors are additional, nonetheless very important, variables in the context of word-order research. Whatever conclusions or inferences may be reached with respect to word order, all theories of word order, at the end of the day, will need to think about how their own conclusions fit in with what is understood about the verbal interaction between the speaker and the hearer, as well as what happens in real-time processing or production. After all, word order is the direct outcome of the speech-act participants' verbal activities, not vice versa. Moreover, the inclusion of the chapter on processing-based approaches will, hopefully, go some way towards nudging the other approaches out from the domain of largely language-internal explanations. The book will close in Chapter 7, with some remarks on the convergence among the theories and the prospects of word-order research.

There is one more compelling reason why the Linguistic-Typological approach needs to be placed before all the other approaches in this book. It is not incorrect to say that different theories highlight different (types of) word-order patterns. For instance, in the Generative Grammar tradition, V-initial order (that is, as theoretically understood in Generative Grammar), so-called scrambling and, to a lesser extent, non-configurationality are regarded as high on the research agenda, as attested by a considerable amount of research on them. However, these topics do not feature (prominently) in Linguistic Typology or in the processing-based theories. Most of these Generative Grammar discussions also tend to be highly theory-internal, not to mention abstract, and, typically, their focus is on how to derive V-initial word order or scrambling from the abstract, universal word order within certain theoretical strictures. Even within the same theory, different structural representations or constraints, not to mention different theoretical assumptions, may be utilized. In a book like this one, therefore, a line has to be drawn in the sand, as it were, as to how much is to be included for discussion, and common ground also needs to be sought for all the theories to be surveyed in the book. Otherwise, the book would lack a common or coherent theme, not to mention needing considerably more space than that available. To the present writer's mind, the strong empirical foundation provided by Linguistic Typology serves as a convenient and logical frame of reference insofar as what any serious theory of word order must minimally deal with empirically is concerned. Thus the theories included for discussion here will all be surveyed primarily with respect to the word orders and their correlations discussed routinely in Linguistic Typology. For this reason, the chapter on the Linguistic-Typological approach is allowed to precede those on the other approaches. This seems to be a reasonably justifiable decision.

Some words are in order to explain why information-based theories of word order (e.g. new vs old, topic vs focus) are not included for discussion in this book, although some reference will be made, in Chapter 6, to the role of information status in word/constituent ordering – for example, given (or more

thematic) information vs new (or less thematic) information. Suffice it to give three reasons for this exclusion here (see Chapter 6 for further discussion). First, information-based studies represent diverse theoretical approaches to information packaging, proposing various types of information status, such as topic, comment, focus, topicality, theme, rheme, which, more frequently than not, prove to be incompatible with, or even contradictory to, each other because they mean different things in different theories (Vallduví 1992: 28–43; also Hawkins 1994: 111–12 on this problem). Indeed "[d]ifferent theories make incompatible predictions regarding information structure and ordering, and different investigators using the same measure of topicality … arrive at conflicting conclusions about their relationship" (Wasow 2002: 65). (In point of fact, a separate book would be needed to do justice to the various information-based theories of word/constituent ordering.) Second, Hawkins (1994: 237–42; 2000: 255–7) provides evidence that information structure may not play so important a role in word/constituent ordering as is claimed; indeed he goes so far as to conclude: "pragmatics appears to play no role whatsoever" (Hawkins 1994: 240–1). Third, as Hawkins (1994: 112) rightly points out, information-based theories fail to address the issue of grammaticalized word orders and their correlations. Indeed "there is no systematic investigation of what the grammatical structure actually is of the sentences and phrases of languages whose ordering is claimed to be driven by discourse [i.e. information packaging] alone" (Hawkins 1994: 112; also Hale 1992). Even languages with word-order freedom "still have phrase structure and there may even be grammatical arguments for a conventionalized basic order" (e.g. Preposition-Noun vs Noun-Postposition, Adjective-Noun vs Noun-Adjective) (Hawkins 1994: 112). What this implies is that information-based theories have a long way to go towards explaining the irrefutable range of word-order options or choices attested among the world's languages. Thus it is very difficult to imagine what role, if any, pragmatic concepts such as topic, comment and focus play in choosing, for instance, between Preposition-Noun and Noun-Postposition, or, more generally, head-initial vs head-final word order. Similarly, word-order correlations such as Verb-Object&Preposition-Noun, Object-Verb&Noun-Postposition, Object-Verb&Verb-Auxiliary and Verb-Object&Auxiliary-Verb do not seem to be amenable to information-based theories, and to the best of the present writer's knowledge, no one has ever attempted to explain these correlations in terms of information structure or information status – for example, what is the similarity in information status between Verb and Auxiliary in the context of Object-Verb&Verb-Auxiliary and Verb-Object&Auxiliary-Verb correlations?[2] There is simply far too much cross-linguistic variation on word order that information-based theories are incapable of handling. In fact, information-based theories seem to be better suited for accounting for (non-conventionalized or non-grammaticalized) within-language word-order variation, although theories such as Hawkins's (1994, 2004), without relying on information-based concepts, are able to handle within-language variation equally well, and possibly even better. By Ockham's razor alone,

theories such as Hawkins's should probably be preferred to information-based ones, which may also need something akin to Hawkins's theory to explain conventionalized word orders and their correlations.

Finally, there remain a couple of comments to be made about why the entr'acte on Generative Grammar is needed in this book and why Optimality Theory is included as the topic of one of the chapters. First, one of the defining characteristics of formal theories such as Principles and Parameters Theory and the Minimalist Program is their decidedly deductive, as opposed to inductive, outlook on the nature or study of language. What this entails for readers is that in order to understand what these theories have to say about word order, they first need to have a reasonably good grasp of the theories' core (meta)theoretical assumptions and concepts – to a much greater extent than might be the case with inductive approaches such as Linguistic Typology or the processing-based theories. Unfortunately, it cannot be reasonably assumed that every reader commands such a theoretical understanding; without it, it will be very difficult to appreciate, let alone understand fully, the nature and content of Generative Grammar research on word order, not least because theoretical or analytical decisions on word order are driven by the theoretical assumptions and concepts themselves. Hence the entr'acte chapter on Generative Grammar. This, of course, is not to say that such a preliminary discussion is not required for Linguistic Typology. What it means is that Linguistic Typology does not need such a lengthy preliminary discussion because in inductive theories like Linguistic Typology, theory tends to arise, as it were, out of data or while data are being investigated. Once again, the downside of specialization in one theory tends to be a lack of understanding of other theories. Thus it makes logistical sense to become familiarized with the theoretical foundation of Generative Grammar before tackling the results of its current word-order research. Moreover, Optimality Theory shares many theoretical assumptions and concepts with Generative Grammar. Thus a sound understanding of Generative Grammar will go a long way in the context of this book. (In a way, this also reflects how the present writer has come to better understand the Generative and the Optimality-Theoretic approach to word order.) Needless to say, readers who are already au fait with the architecture of Generative Grammar are advised to skip the entr'acte chapter and proceed straight to Chapter 4 and/or Chapter 5. Thus while the inclusion of the entr'acte chapter may not be uncontroversial for some readers, the present writer hopes that it will be of much use to many others. For a similar reason, Chapter 5 contains a somewhat lengthy discussion on the theoretical architecture and concepts in Optimality Theory as a prelude to the main discussion. Once again, readers who are familiar with Optimality Theory and wish to learn about its approach to word order are advised to skip this preliminary discussion (i.e. §5.1, §5.2 and §5.3, pages 162–83).

The second comment responds to readers who may potentially question the present writer's wisdom of including Optimality Theory in this book. Fundamentally, Optimality Theory, just like Principles and Parameters

Theory, is a formal theory, developed in the bosom of generative phonology (read: Generative Grammar). Thus some may think there is too much focus on formal theories in the book, at the expense of other non-formal theories. Moreover, while there has so far appeared a substantial amount of research on areas other than phonology, the bulk of Optimality-Theoretic research indeed concerns phonology, not syntax. For instance, two major introductory texts on Optimality Theory – that is, Kager (1999) and McCarthy (2008) – deal almost exclusively with phonology. However, while it has its intellectual roots firmly in the Generative Grammar tradition, Optimality Theory has taken cross-linguistic variation (or between-language variation, as it is commonly referred to in that theory) more seriously than other mainstream generative theories, including Principles and Parameters Theory, may have. Equally importantly, Optimality Theory is decidedly an output-oriented theory, and recent developments in Generative Grammar, particularly in the light of the Minimalist Program, point to the increasing importance of interface or output conditions. Thus Optimality Theory has much to offer for the minimalist view (on word order or linear-ization in particular), not least because of the former's theoretical affinity with Generative Grammar. Conversely, Generative Grammar has much to benefit from Optimality-Theoretic research on output conditions. While it has so far produced a comparatively not large amount of work on syntax – word-order research in Optimality Theory is clearly in its infancy, or at least is not one of its primary research foci, as a check on the Rutgers Optimality Archive (available at http://roa.rutgers.edu) will reveal – Optimality Theory, to the present writer's mind, well deserves a place in a book like this one.

2

The Linguistic-Typological approach

Empirical validity and explanation

Linguistic Typology (LT) can be defined succinctly as the study of structural variation in human language with a view to establishing limits on this variation and seeking explanations for the limits (e.g. Mallinson and Blake 1981; Comrie 1989; Whaley 1997; Song 2001, 2011; Croft 2003). Because LT, more than any other approach, including those surveyed in this book, prioritizes the task of discovering cross-linguistic variation as attested in the world's languages, the LT approach to word order is chosen justifiably as the first of the main chapters of this book. As explained in Chapter 1, this is intended to ensure that readers have an opportunity to familiarize themselves with what kind (or range) of word-order data to investigate before tackling the remainder of this book. This way, they will be in a better position to appreciate the kind and range of data that different theoretical approaches to word order (do not) address in their respective investigations. Put differently, LT imposes the lower bounds, as it were, on what must be accounted for empirically by all theoretical approaches (to word order). Whatever theoretical imperatives or expediencies may initially have a bearing on the way a given theory is to be developed, the reality of the world's languages – that is, empirically valid cross-linguistic variation – must eventually be addressed (and also accommodated) by that theory. To wit, data, not theory, will have the final say in the matter. Moreover, LT, as practised today (or since its rebirth in the early 1960s), had its very beginning in none other than word-order research or, more accurately, Greenberg's (1963) seminal work on word order. Therefore, there is much merit in commencing discussion of word order research with the LT approach.

LT has a centuries-long tradition, with its seminal ideas dating back, not improbably, to as early as the seventeenth century (Ramat 2011; also Graffi 2011). Its very long tradition notwithstanding, LT went largely ignored, if not completely forgotten, between the 1870s and the 1950s, not only because of general dissatisfaction with the morphological classification of languages, which pre-modern (read: pre-Greenbergian) LT was best known for, but also because of

10