NATURAL MORPHOSYNTAX

 A TALE of TWO SYSTEMS and

 IMPLICIT LEARNING

 JEAN A. RONDAL, PH.D., DR. SC. LANG.

 In memory of Ernst Moerk

 FOREWORD

 Psycholinguistics has been much influenced by structural and generative linguistics. It has largely adopted the idea that linguistic rules, categories, and hierarchies, are necessary tools for explaining morphosyntactic functioning in language users.

 The first part of the book is devoted to a reanalysis of the traditional linguistic and psycholinguistic literature on this topic. It will show that a number of past claims are no longer sustainable. Recent contributions suggest the need to distinguish two separate systems in the psychological treatment of morphosyntax: a formal system independent of meaning and a semantic-relational one with a determining role in sentence organization. The two systems may substitute each other. They may also interact such that morphosyntactic functioning is optimized. It is argued that the semantic-relational system is the first one to develop based on general cognition and may remain the most important one beyond development. Both systems have to be learned as no support for a genetic basis of grammar has been found. Language users are not conscious of the processes involved in sentence treatment. Morphosyntactic learning is implicit although some aspects may reach conscious level. It can be accounted for referring to recent theoretical conceptions of implicit learning.

 I am indebted to my son Stéphane Rondal for his help with the informatization of the figures in the text.

 I would like to dedicate this work to the memory of my friend Ernst Moerk, prolific contributor to the field of developmental psycholinguistics for many years.

Contents

FOREWORD

INTRODUCTION

1. THE HERITAGE OF STRUCTURAL AND GENERATIVE LINGUISTICS

1.1. Structural linguistics

1.1.1. Formal classes

1.1.2. Functional classes

1.2. Generative linguistics

1.2.1. Chomsky’s ideas

1.2.2. The independence of syntax

1.2.3. Evolution of generative linguistics

2. PSYCHOLINGUISTICS’EARLY ORIENTATION

2.1. Experiments on transformational and generative grammar

2.2. The modularity hypothesis

3. GRAMMATICAL REPRESENTATION IN LANGUAGE USERS

3.1. Acceptability judgements

3.2. People’s grammatical knowledge

3.3. The innateness issue

3.4. The learnability question

3.5. Brain representations

4. SEQUENCES AND HIERARCHIES

5. SEMANTIC-RELATIONALTHEORY

5.1. Chafe’s and related models

5.2. Generality of semantic relations

6. TWO MORPHOSYNTACTIC SYSTEMS

6.1. Semantic-relational system

6.1.1 Semantic-syntactic patterning

6.1.2. Grammatical morphology as a self-editing process

6.2. Formal morphosyntactic system

7. IMPLICIT MORPHOSYNTACTIC LEARNING

7.1. Conceptualizing implicit learning

7.2. Implicit learning of natural morphosyntax

7.2.1. Syntactic learning

7.2.1.1. Semantic-relational framework

7.2.1.2. Formal syntactic learning

7.2.2. Learning grammatical morphology

7.3. Devoted memory systems

7.4. Developmental language disorders

7.4.1. Specific language impairment

7.4.2. Cognitive handicap

CONCLUSION

BIBLIOGRAPHY

 Introduction

There is no consensus among specialists regarding the nature of morphosyntactic representations in standard language users and the mechanisms at work in the acquisition of morphosyntax. This state of affairs may have much to do with the relationship between theoretical linguistics and psycholinguistics (the branch of psychology whose object is language functioning and acquisition). The two disciplines are rather strange bedfellows. Their rationales, purposes, and methodologies differ in significant ways.

Syntax may be defined as the principles and processes constraining the formation of sentences in a language. Grammatical (or syntactic or inflectional) morphology is concerned with two major types of morphological processes for expressing indications of number, gender, time, aspect, mode, case, or person: the affixation of a bound morpheme to a word stem (i.e., base or root) or a modification of the stem (for example, in forming the past tense of some irregular English verbs: e.g., *eat*/*ate*). Derivational or nominalization morphology is concerned with the formation of new lexemes (i.e., words in the lexicon or vocabulary of the language). Derivational morphology also proceeds mostly with affixation, modifying the word stem in some systematic way to form a new word (for example, an adjective from a noun: *forest*/*forester*). In this book, only grammatical morphology will be taken into consideration.

The basic unit of syntactic organization is the clause that can be subdivided into phrases. It is defined as a group of words minimally containing a grammatical subject (overtly expressed or implied) and a predicate (the part of the clause expressing what is said of the subject). This is the same unit whether it functions in isolation as a simple sentence or as part of a clause complex, i.e., a compound or complex sentence (Halliday, 1985). In the remaining part of the book, I will use the word *sentence* with reference to simple, compound, or complex sentences; the word *clause* for a specific reference to this unit as part of a clause complex; and the expressions *compound sentence* and *complex sentence* for designating sentences with coordinate clauses and sentences with a main clause and one or several subordinate clause(s), respectively. The term *proposition* will be used as synonymous with *clause* but in the logical (semantic) sense of being composed of a predicate and one or several arguments.

Semantics relates to meaning, or, in general terms, the properties of linguistic phenomena determined from nonlinguistic (cognitive) elements. Lexical semantics deals with the denotative meaning of the words in the language. Relational (also labelled structural) semantics (or thematic, case relations, or yet theta grids) regroups the meaning relations between words incorporated in phrases or sentences (e.g., notions of transitivity, possession, location, qualification, quantification, instrumentation, accompaniment).

 A sentence can be analyzed from a syntagmatic or a paradigmatic point of view. The syntagmatic (horizontal) axis concerns the relationships between linguistic elements (phonological, morphological, and syntactic) that co-occur in the speech chain. The paradigmatic axis refers to a dimension that can be represented as a vertical plane corresponding to a set of alternative lexemes that would be legitimate in a syntagmatic sequence. Only one of the alternatives can occupy a particular location in the sequence at a given time and each alternative may modify the meaning of the sequence.

Linguistics is a hermeneutic discipline, the specific object of which is the interpretation of the available data regarding present and past states of languages in the world. Epistemologically, it can freely appeal to all concepts it judges appropriate for describing languages and in particular their morphosyntactic component. However, it lacks the methodological tools necessary to go beyond description. Linguists, typically, have no experimental control over the situations in which language behaviors are occurring, cannot draw accurate causal inferences, and have no objective ways for validating alternative theoretical models. Psycholinguistics, or modern language psychology, is a neurocognitive discipline rooted in observation and experimentation. Its particular object is the analysis of the processes involved in the learning and treatment of language by “flesh and bone” people. As a functional discipline, psycholinguistics cannot satisfy itself with mere descriptions and intuitions. While this indication leaves linguistics untouched, it should make wary anybody wanting to derive psycholinguistic explanations from theoretical linguistics. As we will see, this cautious advice has not always been followed in psycholinguistics, which has led to unnecessary confusions (already Palermo, 1971).

 The term *functional* is used here, referring to psychology, to denote a discipline aiming at the explanation of biological phenomena through empirical investigation. The same word is sometimes used in linguistics as well, but with a different (and varying) sense. For example, Halliday (1985) labelled his grammar “functional”, meaning that it was to be ultimately explained by reference to how language is used. As we will see in Chapter 1, a long-time tradition in linguistics assigns the parts of speech to various formal and functional categories. *Functional*, in that context, refers to the role that the members of these classes play in the grammatical organization of the sentence (e.g., grammatical subject of a finite verb). *Formal* refers to an approach that is concerned with the forms and not with the meaning that these forms may convey. *Abstract* refers to something that is not directly given in perception. And the notion of rule, ubiquitous in modern linguistics, can be defined in a first approximation, as a generalization that has the potential of open-ended applications.

 The theoretical discussions in the text are concerned in the first place with language production (expression). Only passing references are made to language comprehension (reception). Lashley (1951) already noted that speech production and comprehension have too much in common to depend on wholly different mechanisms. Pinker (1994) suggests that speaking and understanding share the same grammatical data base (p.197). Pickering and Garrod (2013) argue that production and comprehension are tightly interwoven and that this interweaving underlies the ability of people to predict themselves as well as other people. Globally, comprehension can be viewed as proceeding the other way around of production. In other words, in order to understand a sentence it is necessary to start with its surface organization and figure out the underlying meaning relationships as likely intended by the speaker. One difference between language production and comprehension, however, is that the listener may often be able to bypass partially or completely the morphosyntactic level of analysis if (s)he can guess at the meaning of the sentence from previous or present knowledge of the speaker, the context of utterance, etc., or just rely on the meaning of individual words. These strategies are well-known and have been illustrated in numerous psycholinguistic texts.

In his 2004 book, Michel Paradis suggests that in order to render clearer the distinction between implicit competence and explicit learning (a central issue in the present essay), it would be preferable to restrict the use of the word *learning* to conscious phenomena and utilize the word *acquisition* for non-conscious phenomena and mechanisms. I believe that such a terminological distinction is justified indeed. However, as there exists a long tradition, still going on, of use of the expression *implicit learning* in the specialized literature (that will be thoroughly analyzed in the book), to refer to incidental or non-intentional learning, changing terminology could represent a problem to the non-expert reader. This is the reason why I have kept with the word *learning* in both expressions *implicit* and *explicit learning* throughout the book. *Implicit* designates learning that is incidental and occurs without awareness in the learner of what has been or is being learned, of the internal processes involved in learning, and even in some cases that there was something to learn. *Explicit learning* implies intention to learn and may lead to awareness of the output, product, and peripheral aspects of behavioral adaptations. Meta-considerations belong to the explicit register (for example, metalinguistics is “using language to speak about language”) but they imply a kind of reflexive attitude of mind regarding behavioral outputs.

The book is organized in the following way.

Chapter 1 recalls a number of key notions in structural and generative linguistics concerning the morphosyntax of language. Technical notations and symbols in Chapter 1 and later in the book have been kept as simple as possible. They are retaken from *Syntactic structures* (Chomsky, 1957, Appendixes I and II), except when further theoretical development has led to the introduction of new symbols. All symbols are defined on first appearance in text. It could perhaps be asked whether such a long and technical chapter is justified. I think that it is, given that generative linguistics has dominated the field of linguistics and heavily influenced psycholinguistics for a long time (and still does to some extent). No precise discussion of important issues and theories in psycholinguistics in following chapters would have been accessible to the reader without an introduction of the kind. It matters to understand where a number of theoretical ideas that are still very much with us (see Evans, 2014) have come from, even if some of them have successively been proven incorrect, and why they need to be replaced by more psychologically realistic considerations. Dik (1991) quite reasonably states that a language theory should be “compatible with the results of psycholinguistic research on the acquisition, processing, production, interpretation and memorization of linguistic expressions” (p. 248).

 In Chapter 2, major borrowings of early psycholinguistics from traditional structural and generative linguistics are discussed. They include the belief that the categories used in linguistics for describing the morphosyntax of language, are those of the regular language users. Given that these notions seem to fit relatively well in with a grammatical description of the parts of speech and their functional configurations, it was admitted by most psycholinguists from the 1950s on that these notions ipso facto have psychological reality. Second, the recourse to the notion of unconscious rule, i.e., a rule supposed to play a role in language treatment but unknown to language users. This notion is still very much with us. The technical literature in linguistics, psycholinguistics, and neurolinguistics is replete with the word *rule*. It could be the single word most frequently used in the publications of these disciplines.

Again the necessity of an historical chapter of the kind may be questioned. As for the preceding one, I believe that it will help the reader in understanding better where several of the key theoretical concepts in psycholinguistics have come from, why they were thought to be correct at a time, and why in some cases it has become clear that they have to be rejected or seriously amended.

 Chapter 3 analyses the question of the psychological reality of linguistic notions and morphosyntactic representations in regular language users. Grammatical acceptability judgements and people’s responses to grammatical questions demonstrate that their conscious formal knowledge is reduced. No genetic basis for a hypothesis of grammatical representational innateness seems to exist either. Yet, neurological data suggest that some syntactic representations exist at brain level. It would follow that they have to be learned. There may be two systems regulating morphosyntactic treatment: a formal syntactic one largely independent of meaning and a semantic-relational one.

In Chapter 4, one argues that sentence sequential organization is as fundamental to language treatment as the hierarchical dimension. It will be argued that the two dimensions are less opposed that it may seem.

 A semantic-relational theory adapted from Chafe’s book (1973) *Meaning and the structure of language*, is presented in Chapter 5, together with a brief analysis of other semantic theories. Semantic relations are abstract notions. They may be considered as having psychological reality given that they stem from an interface with general cognition (Bierwish & Schreuder, 1993). As to their basic dimensions, they appear to be universal. This is in agreement with the rejection of the Worf-Sapir linguistic relativity hypothesis.

 Chapter 6 addresses the question of the regulating forces in sentence processing. A proposal is made and illustrated, in which two systems (semantic and formal morphosyntactic) function independently of each other but may be called forth to interact.

In Chapter 7, morphosyntactic acquisition is explained with particular reference to the mechanism of implicit learning. This mechanism has been theorized with reference to the learning of artificial grammars. It will be adapted to the context of natural morphosyntax. Relevant aspects of morphosyntactic acquisition in children will be analyzed as well as the memory systems involved. Recent research on developmental language disorders of the morphosyntactic type will also be presented. They suggest that the language difficulties of these children originate in a fundamental limitation in their capacity for implicit learning and procedural memory.

The theory presented in this book is empirical. As such, it can be evaluated by the number of facts which are explained or dealt with by the theory. It is of the general constructivist type, meaning that natural morphosyntax is (re)constructed by each child from input evidence with the powerful help of devoted brain mechanisms.

Technical terminology and symbolics are borrowed from traditional structural and generative linguistics[[1]](#footnote-1) on the one hand, and from Chafe’s (1973) text regarding relational semantics.

Several of the analyses presented in some of the chapters are tentative and at times even speculative (plausibly speculative, I would like to believe). But is not science at the theoretical level modestly a question of plausibility overall (at best provisorily definitive, at worst definitively provisory, as Karl Popper (1935, 1992)) liked to say. The aim of the discussions can thus be viewed as departure points for further research. This book, particularly its later chapters, will not be easy reading. A plausible excuse is that, as Wallace Chafe (1973) wrote, “The observation of semantic facts is impossible without an expenditure of cerebral energy” (p.11).

 Chapter One

 The Heritage of Structural
 and Generative Linguistics

Generative linguistics, developed in the second half of the 20th century, is logically, conceptually, and methodologically, a branch (and a continuation of structural linguistics but with a number of peculiarities). Major characteristics of traditional structural linguistics will be recalled before moving to an analysis of the major tenets of transformational and generative grammar (as it was labelled initially) and their evolution in time.

1.1. Structural linguistics

Structural linguistics (synchronic, i.e., concerned with the state of a language at a given moment of its evolution) appears in the early years of the 20th century, partly in reaction against historical and diachronic linguistics (the study of language over time) that had dominated the field in previous decades. Structural linguistics was defined from the start as a formal discipline whose goal is to produce taxonomic grammatical descriptions of natural languages. The formal orientation was confirmed and even amplified in the second half of the 20th century in generative linguistics.

De Saussure (1916, 1955), generally considered the father of structural linguistics (although he is reported as having never used the word *structure*), defined language as a system of signs resulting from the association of a *signified* (the conceptual meaning of a given sign) and a *signifier* (the form the sign usually takes, i.e., its particular sequence of phonemes)[[2]](#footnote-2). The association is arbitrary (i.e., non-motivated) and conventional (as already noted by Leibniz (1703)). As is known, signifiers vary substantially between languages independently from the meaning expressed. This led Saussure and followers to decide that the particular object of linguistics is the form of the signs and the combination of signs and not, or only secondarily, the content expressed.

The Danese linguist Hjelmslev (1928) developed Saussure’s ideas into an articulated structural theory. He defined a “linguistic linguistics”, purged from mentalism or psychologism. According to Hjelmslev, language is articulated in several major components or subsystems (phonology, lexicon, morphosyntax), each one with its particular structure. The components are both autonomous and interdependent, corresponding to a set of synchronic relations of distinctions and oppositions. What matters is the organization of the formal structures that make up the inherent substance of the language. Hjelmslev wrote: “Grammar is a discipline that is *by definition* (his stress) indifferent with respect to any question of semantics” (1928, p. 202, my translation). The orientation is purely descriptive and taxonomical.

Linguistic structuralism is an epistemology of the absolute minimum in the comment of Milner (1989). There is no room for any psychological consideration. It has been criticized, besides its abstract and formalistic character, as conducting to a “mutilation” of the object of linguistic science; given that obviously language is concerned with meaning in the first place (Martinet, 1953). There is little doubt that this theoretical stance has markedly delayed the development of structural semantics which had to wait the second half of the 20th century to see its first systematic formulations. As will be seen also, generative linguistics has been compelled to reintroduce semantics as a component in the analysis of the grammatical system of language but in a particular way that will be analyzed later in the chapter.

In structural linguistics, priority is given to the syntagmatic axis of language analysis over the paradigmatic one. Syntagmatic must not be equated with linear[[3]](#footnote-3). Underneath the linear appearance of a sentence, one posits that there is a structure of constituents. The structural order is different from the temporal one. Tesnière (1966) pointed out: “To speak a language is transforming a structural order into a linear one, and understanding a language is transforming a linear order into a structural one” (p.19, my translation).

Syntactic analysis subdivides the utterances into smaller and smaller units until the level of the morphemes is reached (i.e., the minimal units combining meaning and form). A similar technique, under the name of analysis in immediate constituents, has been proposed about at the same time by the American linguist Bloomfield (1933). No space is assigned to a specific study of the meaning of the utterances. Meaning for the Bloomfieldians is assimilated to the reference of the word and the situation in which it is produced and reacted upon by the interlocutor (a definition close the behaviorist conception in psychology, as expressed, for example, in Skinner, 1957).

 Zellig Harris (1951), the academic adviser of Chomsky at the University of Pennsylvania, professed a purely structuralist approach to linguistics. “That is, descriptive linguistics could (and should) be limited to questions of distribution - of the freedom of occurrence of portions of an utterance relatively to each other” (1951, p. 5). A crucial component of Harris’ approach was the “autonomy (of syntax) hypothesis,” i.e., the idea that syntactic analysis should be done without recourse to semantic terms; a conception that we will find again, and prominently, in Chomsky’s writings.

 1.1.1. Formal classes

The notional taxonomy of structural linguistics distinguishes formal and functional categories of words or lexemes. Formal categories (also called lexical classes) regroup the abstract notions of articles, nouns, pronouns, verbs (including participle and infinitive forms), adjectives, adverbs, prepositions, and conjunctions). Article and demonstrative adjectives (e.g., *this*, *that*) are sometimes regrouped in one sub-category labelled “determiner.”

 It is generally thought that the major grammatical categories or classes of words were primarily defined by the philosophers of ancient Greece, Platon, Aristotle, and the Stoics, with the aim of describing parts of the discourse (*logos*). The Alexandrian philologists (g*rammatikoi*), particularly Aristarcos of Samothrace, Denys the Thracian, and Apollonius Dyscole, fixed to eight the parts of discourse: noun, verb, participle (considered a mix between noun and verb), article, pronoun, preposition, adverb, and conjunction. The word for grammar comes from the Greek *grammatikè tekhnè* (the art of letters - of the alphabet), attested for the first time in Platon’s *Cratyle*. Aristotle anticipated grammatical morphology with his notion of case (*ptosis*), i.e., for him, the noun forms differing from the nominative (subject) form, and the verbal forms differing from the present tense of the indicative mode (Lallot, 1988, 1998; Egger, 2017).

The precise definition of the formal classes (beyond simple “family resemblance”; Wittgenstein, 1961) has always been problematic. Lagarde (1988) has analyzed the perennial debates on that matter. Major difficulties are as follows. Classifications do not rest on solid theoretical foundation. A variety of criteria, formal, semantic, pragmatic, or logical have been used to define the categories with little intrinsic coherence. Some criteria are circular. For example, verbs are said to be identifiable as such because they are inflected on the model of a numerous set of words traditionally labelled as verbs. Nouns are labelled nouns because they figure typically in some kind of word combination that have been traditionally recognized as characteristic of noun complexes in opposition to other kind of complexes labelled otherwise (Martinet, 1965). There is no explanation of which criteria were used in the first place to identify the categories of words.

 An important degree of overlapping exists between partially permeable categories. For instance, the word *rest* can be a noun or a verb depending on the syntactic context. The gerund in English is a noun form able to govern like a verb, e.g., *His doing this is disputable*. The classifications heavily depend on the word units with little interest for the morphosyntactic dimension. This leads to conferring to the words properties actually belonging to the phrases or the bound morphemes.

There are a few exceptions. Following theoretical suggestions by Chomsky (1970) and Jackendoff (1977), regarding the so-callled X-bar theory, Emonds (1985) has suggested that at least for the noun, adjective, verb, and preposition phrases, the structures can be reduced to recursive specifier-head configurations. Each head has its own set of specifiers and can be defined according to its ability to combine with them to form a phrase. It is not sure, however, that these suggestions are sufficient for resolving the problem of the identity of the formal classes, as no unquestionable definition of the classes is supplied.

Some distinctions between formal classes do not seem to have universal value. They are relatively clear in English and in most Indo-European languages but less in other linguistic families. Categories such as verb and adjective do not seem to be universal (Evans & Levinson, 2009; Haspelmath, 2007). In Mandarin Chinese some words are used to designate indifferently an object or a corresponding action performed with that object (Perrot, 1953). In some American-Indian languages, like Hupa (spoken in Oregon), active or passive verbal forms at the third person singular are used as nouns (e.g., *nanya* - gloss: *It comes down[[4]](#footnote-4)* - meaning *rain*; *naxowilloie* - gloss: *It is attached around him/her* - meaning *belt*) (Benveniste, 1950).

 1.1.2. Functional classes

Linguists are rather well agreed on the fact that each clause of a natural language is correctly represented by at least one phrase marker of some kind. The functional categories concern the grammatical roles that different formal kinds of words play in sentence organization. They include the notions of grammatical subject and object (direct, indirect, or oblique) of the verb, complement, modifier (related to the function of the adjectives and adverbs), attribute, and anaphora. Several formal classes of words also act as head of corresponding phrases (nominal, adjectival, verbal, adverbial, prepositional, participial, and infinitival). Hierarchies holding between phrases account for sentence structure.

A common idea is that a hierarchical grammar of phrase structure of some sort provides the best approximation to the syntactic organization of human languages.

Actually, the question to know whether the syntactic organization of natural languages is better approximated by a hierarchical grammar of phrase structure, a finite-state grammar, or other theoretical models, is still an open one (see for diverging opinions, Sproat, 2002 ; Levelt, 2008 ; Fitch, Friederici, & Hagoort, 2012 ; and Frank, Bod, & Christiansen, 2012). The notion of finite-state grammar, often contrasted with hierarchical grammar, is not univocal. It is usually defined as involving a linear left-to-right derivation corresponding to a Markov chain (a Russian mathematician from the early 20th century who first suggested analyzing verbal utterances in this way), that is a disposal of contiguous dependencies. As a matter of fact, a Markovian scheme need not be limited to first-order concatenations. It can deal with higher order dependencies where, for example, the probability of selection of a given item in the sequence depends on the presence or the absence of one or several anterior items (see, for example, Manning & Schütze, 1999). It is possible in this way to account for relatively long series of elements linked together by complex sequential dependencies. This includes the organization of sentences according to an analysis in constituents, i.e., at a higher level of abstraction than the sequences of words at the surface of the sentences.

A hierarchical phrase structure analysis can be illustrated with the tree-like diagram[[5]](#footnote-5) (Figure 1.1) of a simple clause (declarative active affirmative): *The player hits the ball* (leaving aside the grammatical morphological marking on the verb). Other notational systems exist (for example, diverse types of bracketing).



Fig. 1.1. The clause (symbolized S for a whole sentence) is split into a noun phrase (NP) and a verb phrase (VP) at a first level of hierarchical depth. At a second level of depth, the NP is split into an article (A) and a noun (N), and the VP into a verb (V) and another NP, itself split into T and N at a further level of depth.

The diagram displayed in Figure 1.1 illustrates the hierarchical nature of phrase structure grammars. The depth of the analysis from the surface of the clause may be important; the more so if one follows Chomsky’s (1957 and continued) recommendation that a tree diagram be preferably binary, i.e., consisting of two parts under each node. The configurational definition of the functional categories can also be illustrated in Figure 1-1. The grammatical subject of the clause is the head of the leftmost phrase located immediately below the S symbol. It is most often a noun phrase. The grammatical subject therefore is the noun occupying the highest position in the tree. It can also be a pronoun, an infinitive, or even a whole clause. A simple sentence is defined as minimally consisting of a noun phrase and a verb phrase with a functional relationship.

1.2. Generative linguistics

There is no doubt that the American linguist Noam Chomsky has been the undisputed leading figure during the fifty years or so that generative grammar has dominated the international field of linguistics (for a direct testimony, see the conversation between Paul Postal, one of the early proponents of the generative movement, Geoffrey Huck, and John Goldsmith, in Huck & Goldsmith, 1995, pp. 126-142). Rosenberg (1993) notes that the history of linguistics in the second half of the 20th century is to a large extent that of the evolution of Chomsky’s grammatical theory. Expectedly the following presentation is centered on Chomsky’s writings. It is clear, however, that many other linguists have been involved in various ways in the construction of the theory and its evolution over time, even if it is not possible to acknowledge their contributions in this essay.

1.2.1. Chomsky’s ideas

According to Piaget (1968), transformational and generative grammar represents a more genuine structuralism than the previous structural linguistics, thanks to its transformational component. “From such a point of view, the linguistic ‘structure’ reaches the level of the most general structures with their laws of totality which are laws of transformation and not descriptive and static, and with their self-regulation allowed by the characters of that composition “ (Piaget, 1968, p. 68; my literal translation). Piaget was referring to an early version of generative grammar. The transformational component as such (with the optional transformations) was modified and simplified in following versions of the theory as will be seen below.

 There are important differences between traditional structural linguistics and generative linguistics. Chomsky wanted to overcome what he saw as an impoverishment in structural linguistics, i.e., its purely descriptive emphasis and its limitation to the analysis of sentences in immediate constituents. In *Syntactic structures* (1957), he suggests: “Syntactic investigation of a given language has as its goal the construction of a grammar that can be viewed as a device of some sort for producing the sentences of the language under analysis” (p. 1). Linguists need to be concerned with determining the fundamental underlying properties of successful grammars. And “The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages” (*ibidem*).

 Saussure’s and Hjelmslev’s recommendations to avoid all form of psychologizing are left aside. Chomsky, in many places, defines linguistics as a branch of cognitive psychology while at the same time proclaiming the purely formal nature of language (1969), sanctifying the linguists’grammatical intuition, and rejecting the empirical methods of psychology. In *Aspects of the theory of syntax* (1965), which together with *Syntactic structures* (1957), provided the fundamentals of the new theoretical movement, Chomsky writes: “… by a generative grammar I mean simply a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences. Obviously, every speaker of a language has mastered and internalized a generative grammar that expresses his knowledge of his language. This is not to say that he is aware of the rules of the grammar or even that he can become aware of them, or that his statements about his intuitive knowledge of the language are necessarily accurate. Any interesting generative grammar will be dealing for the most part, with mental processes that are far beyond the level of actual or even potential consciousness; furthermore, it is quite apparent that a speaker’s report and viewpoints about his behavior and his competence may be in error. Thus a generative grammar attempts to specify what the speaker actually knows not what he may report about his knowledge” (1965, p. 8).

The terms *generative* and *generate* are intended by Chomsky in the sense of logic, i.e., *to* *enumerate or formulate explicitly*. These terms have been often understood in the more usual meaning of “*bring something to existence*, *cause something to exist*”, which Chomsky regrets but sees no reason for a revision of terminology (1965, p. 9).

Actually, this is not just a casual misunderstanding. Chomsky and his followers and interpreters, over the years, are largely responsible for what is more than a lexical ambiguity. I beg the reader to bear with another rather long quotation from Chomsky with the excuse that it illuminates the debate. Chomsky (1965) goes on: “… a generative grammar is not a model for a speaker or a hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker-hearer. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use - the theory of performance. No doubt, a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer’s knowledge of the language; but the generative grammar does not, in itself, prescribe the character or functioning of a perceptual model or a model of speech production” (p.9).

There is an important potential for a conceptual confusion between linguistics and psychology in these statements that were repeated almost *verbatim* many times in the long course of generative linguistics. Chomsky has insisted that his descriptive suggestions on the inner organization of morphosyntax should not be taken as psychological models for the treatment (productive and receptive) of language. And this has been taken seriously by a number of linguists (for example, Lyons, 1970). But at the same time, his formulations are ambiguous. They may easily be interpreted (and have been) as encouraging psycholinguists to see generative formulations as psychological models. Just looking at the above quotation, it is indicated that a “reasonable” move in psycholinguistics would be to incorporate generative grammar “as a basic element” of the speaker-hearers’ knowledge of the language. And this is exactly what many psycholinguists have tried to do for decades (see Chapter 2).

There is more, however, in Chomsky’s writings that just a suggestion to make use of generative grammar in psycholinguistic analysis.

 Let’s consider the competence-performance distinction. Chomsky, throughout his career, has tried to separate out language from communication (Evans, 2014). Current language performance, Chomsky argues, i.e., what people say when communicating linguistically, is exposed to errors, memory failure, inconsistencies, false departures, interruptions, and the like. It cannot directly reflect competence. Chomsky further claims that competence is the linguists’ exclusive province relying on their language intuition. As Dennett (1994) reports, “The general attitude among cognitive scientists has been that since the linguists seemed to think they could do it all without benefit of controlled experiments (they just consulted their grammatical intuitions, which had long since been sullied by their own theories) and since they thought the brain was irrelevant, they could just go off and play their mandarin mind-games amongst themselves” (p. 10).

Disambiguating language knowledge from performance is not an easy task but there is no reason to think that it is an impossible endeavor. Language assessment, when done correctly, may reach a kind of asymptote with regard to people’s “genuine” knowledge, which does not render it trivial or useless even if it can never be considered to correspond exactly to the “real” knowledge. The same problem exists with any kind of psychological assessment.

 Chomsky (1965) further indicates that linguistic competence is the knowledge of language deposited in the mind of the speakers-hearers and that this knowledge is a generative grammar of the language; not merely a static inventory of items as in traditional structural grammars, but rather a system of processes able to account for the creativity of language.

The latter notion corresponds to the fact that native speakers of a language can produce and understand an indefinitely large number of grammatical sentences that they have never heard before or which may never have been uttered before by anyone.[[6]](#footnote-6) The central nucleus of this system is a universal grammar (Chomsky, 1966); a rationalist conviction maintained until the present times (Chomsky, 2011).

The notion of creativity is related to that of recursion (Chomsky 1957). Recursion may be defined as the existence in a system of a finite number of rules able to generate infinity of sequences from a finite set of symbols. Linguistic recursion is the repeated sequential use of a particular type of linguistic element or structure. It is also defined as the ability to insert one language component (for example, a phrase) inside another component of the same kind at the same hierarchical level.

In Chomsky’s writing (e.g., 1980, pp. 221-222), recursion is connected with the idea of infinity. However, the infinitude of human language has not been established and could probably not be (see several chapters in Van der Hulst’s edited book, 2009, for a discussion).

The idea of universal grammar[[7]](#footnote-7) is inspired by the *Grammaire générale et raisonnée*, also known as the grammar of Port-Royal, an abbey in Paris. It was published in 1660 by Arnaud and Lancelot and was influenced by the conceptions of the French philosopher Descartes and his follower Cordemoy (1668). This grammar purported to base the study of language on general logical principles declared specific to the human mind. The proposition (a concept borrowed from logic and corresponding roughly to a complete sentence) is considered to supply the various levels of analysis from which the parts of speech can be identified. It is given for certain that the propositional order of the words reflects the order of reason (which will become problematic as soon as languages other than French will be studied, but was not realized at the time). This conviction justifies the absence of a formulation regarding the rules for constructing sentences in the grammar.

Other analytical notions in Port-Royal grammar are a distinction between sentence surface and deep structure, the notion of subordination, the addition and suppression of some forms in the surface of the propositions disregarding the normal order of the words as corresponding to the general meaning of the proposition, and in a similar way the displacement in surface of some forms from the place they should occupy given the meaning of the proposition (for example, in accounting for some subtypes of relative subordinates). It is easy to see the use that the transformational and generative grammar of Chomsky will make of these notions.

Chomsky’s (1965) notion of universal grammar is that all human languages share a common basis of core principles (i.e., general types of rules). More precisely, it is defined as the “system of categories, mechanisms and constraints shared by all human languages and considered to be innate” [Chomsky, 1986](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4477053/#B39), p. 3; [2007](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4477053/#B44), p.1).

It is important to distinguish between formal and substantive universals. Substantive universals are concerned with the generalizations that can be made regarding the phonological, morphological, syntactic, and semantic units, as constitutive components of human languages. Formal universals relate to the constraints bearing on the form of the grammars.

 In a further elaboration (Chomsky, 1981), some of the core principles of the formal universals were redefined parametrically (in the mathematical sense of a quantity that may vary in different cases), i.e., as structural dimensions along which the grammar of particular languages can be allowed to vary to a limited extent.

For example, regarding the stem parameter, some languages like English allow the production of “nude” lexical stems (as in *I work*, *you work*, *we work*, *they work*) whereas other languages like Italian do not (*lavoro*, *lavori*, *lavora*, *lavoriamo*, *lavorate*, *lavorano*). An example of syntactic parameter is the null subject. Some languages like English and French do not accept sentences formulated without grammatical subject of the verb (except in imperatives where the subject may be implied, which may serve to distinguish declarative and imperative sentences) whereas other languages (like Spanish or Italian) do not present a corresponding restriction (e.g., Italian: *Parlano bene l’italiano* - gloss: *They speak Italian well*; Spanish: *Quieras decir algo?* - gloss: *You want to say something?)*. The parametric version of universal grammar may be better equipped for dealing with the important structural differences existing between languages and families of languages; an argument often levelled against the notion of universal grammar (see Evans, 2014).

However, given that two particular grammars could be obedient to the same parameter and yet differ substantially plus the fact that given languages could be exceptions in that they would not recruit particular universal, one is entitled to ask whether a parametric universal grammar has a proper content or whether all its content is not absorbed by the putative various grammatical choices made by the particular idioms (Milner, 1989, p. 236).

Many questions are without a clear answer. For example, it is not sure that the notion of movement that has replaced the optional transformations in later versions of generative grammar (see below) is appropriate in accounting for the way speakers formulate questions in a number of languages (e.g., in Mandarin Chinese). In some languages, it is possible simply to replace the item that is the object of the request in information with a general interrogative term. Hagège (1976) and Toley and Van Valin (1984) further argue that a number of notions presented as universal are only found in a limited number of languages, overall in English syntax. English, they add, is a particular language in several respects, among them a reliance on a relatively rigid word order. Another counterargument to the existence of strong language universals is the well-known presence of so many irregular forms and exceptions in all languages (Lederer, 1989).

Chomsky (1965, p.118) acknowledges that comparative work has indeed documented a great deal of diversity in the surface structure of languages (see below for a precise definition of these terms). He counteracts indicating that as the study of deep structure (idem) was not the object of these studies, they cannot be taken as supplying a valid counterargument to the notion of universal grammar. Chomsky argues that it is more relevant to hypothesize that the deep structure of natural languages is largely universal (also Postal, 1968). In a more recent work (e.g. Hauser, Chomsky, & Fitch, 2002), the human faculty of language is considered to correspond more narrowly to a computational mechanism allowing recursion (see below).

 The question of the existence of a consistent universal grammar still is the object of debate in linguistics. A number of supportive observations have been published (for example, Greenberg, 1963, regarding statistical tendencies in surface word order, strictly speaking, typological universals; Chomsky, 1965; several chapters in the book edited by Bach & Harms, 1968). Other linguists are skeptical of the existence of universal grammatical properties beyond a small set of relatively simple principles (Hagège, 1976, 1982; Newmeyer, 2005; Evans & Levinson, 2009). Evans (2014) even considers absolute language universals to be a myth.

Several psycholinguists have published severe criticisms of universal grammar as a theory underlying the acquisition of language on the ground that it is not supported by developmental facts and that it is far from clear how it could be linked concretely to language acquisition (Ibbotson & Tomasello, 2016).

 Dabrowska (2015) indicates that universal grammar is suspect as a concept and that the empirical evidence for it is very weak. There is no agreement on its content. In particular, languages differ from each other in profound ways and there are very few true universals. What needs explaining is not universality but cross-linguistic diversity.

Ambridge, Pine, and Lieven (2014) claim that universal grammar provides no realistic solution to the learnability problem of morphosyntax in natural language. It does not help to identifying syntactic categories, acquiring basic morphosyntax, understanding general principles such as subjacency or binding, and even setting various parameters postulated in universal grammar. In each of these domains, universal grammar as presented, even if it were true, would not facilitate the task of the language learner. Ambridge, Pine , and Lieven (2015) add that the fatal flaw in most UG-based approaches to acquisition is their focus on adult language described in terms of a particular linguistic formalism. As a consequence, such accounts typically neglect to link acquisition to the input to which the learner is actually exposed.

 Lin (2017) has produced probably the most radical criticism of the Chomskyan notion of universal grammar to date, ending up in a flat refutation of the theory as methodologically flawed and deeply problematic regarding its contents. Important methodological reservations are already found in Milner (1989) indicating that generative linguistics is devoid of an objective “observatory.”

Lin (2017) notes that although Chomsky has pretended in various publications that he is employing the standard method of the sciences or a “Galilean style” of inquiry, it is clear that this is not the case[[8]](#footnote-8). His method is idiosyncratic and does not measure up to the requirements of empirical science. In a nutshell, the heuristic goes as follows: start from some grammatical data (most often from English), find some rules which can explain the data and then try to obtain some more abstract principles; and if recalcitrant data are found, the rules and principles are subjected to revision.

The methodological caveat can be illustrated with one important aspect of the theory of principle and parameter (Chomsky (1981), i.e., the principle of subjacency. Other principles could be discussed in the same way. Subjacency specifies restrictions placed on movement of syntactic elements. It is based on English data. Other languages as Italian, Russian, and Swedish appear to violate this principle (Newmeyer, 2004). In order to decide that the subjacency principle is universal, one should logically test it against all existing languages on the planete (approximately 6,000 languages). Given that the number of possible human languages is potentially infinite there may actually be no way to ascertain the claim that all human languages must satisfy subjacency or even some parametric version of it.

There is another logical problem with the theory, this time concerning parameters. When two or more languages differ in a certain basic grammatical phenomenon, the difference is attributed to parametric variation. The more languages that will be studied, the more likely they will be found to differ from the known ones, and more parameters (possibly an infinite set) will have to be posited. The idea expressed in *Government and binding* (1981) according to which in human languages there are a finite set of grammatical principles and a finite set of parameters may be preposterous as the set of parameters for all human language variations cannot be determined by looking only at some limited number of existing languages.

Nevertheless, as indicated above, for Chomsky and followers, basic principles as those ascribed to universal grammar and the human faculty of language do not need to be learned or derived from language exposure in any way. They are said to be innately given[[9]](#footnote-9) corresponding to the genetic component of the human language faculty, which, it is argued, supplies the clearest differences between mankind and the rest of the animal kingdom. The notion of a language faculty specific to mankind comes from Descartes (1646, November 23, in a letter to the Marquis of Newcastle). In Chomsky’s writings, it has become a protean concept, sometimes to be taken in a completely abstract way, sometimes implying a neural organization underlying a mental content corresponding to a grammar.

 Chomsky (1966, 1968, 1997) sees no reason to put in doubt the Cartesian idea that the capacity to use linguistic signs to express freely conceived thoughts provides the true separation between man and beast.

Barsky (1997) has reproduced quotes from Chomsky’s interviews in the *New York Times* (September 25, 1975; March 31, 1995) where it is said that the efforts by psychologists to teach animals to speak (sic) are considered “an absurdity by every serious biologist;” and “have not the remotest relation to science, though people in white coats and with equipment sometimes do it.”

The point is clear given the Cartesian premise but why such a strong dislike? As is well known, the modern (successful) attempts to teach some nontrivial language to apes, dolphins or sea lions have not relied on speech but rather on language modalities more in line with the biological equipment of these species. Although this is not the topic of the present essay, the interested reader may want to have a look at my books on the topic (Rondal, 2000, 2016), as well as at the more general discussion of Francois (2014) on the emergence and evolution of human language from a neuroscientific point of view. Starting with general abstract ideas like universal rules of grammar, as Chomsky does, one is doomed to fail finding much nonhuman equivalence. Instead, from a more open functional approach, it has become clear in recent decades that the human language organization including its syntax is prepared in nontrivial ways through the abilities of several higher animal species. The bottlenose dolphins and the sea lions, for instance, are able to use simple but genuine syntactic means to understand sophisticated instructions given by humans in sign or ultrasonic language modalities. Given the empirical data now at disposal, it may be considered likely that there has been an evolutionary construction of the language ability including in its basic syntactic organization.

This, of course, is contrary to Chomsky’s radical variety of formal rationalism, which may at least partially explain his gratuitous bad faith in the matter. As he has decided that human language is a unique phenomenon without any equivalent in the animal world, it must follow that it is meaningless, even foolish, trying to address the problem of the evolution of human language from more primitive communication abilities.

As will be seen in the next pages, Chomsky knows of Lashley’s (1951) text on *The problem of serial order in behavior.* He uses it in his 1959 criticism of the Skinnerian approach to language (see below). The classical text of the Harvard psychophysiologist terminates on an interesting note: “I am coming more and more to the conviction that the rudiments of every human behavioral mechanism will be found far down in the evolutionary scale and also represented even in primitive articles of the nervous system. If there exists, in human cerebral action, processes which seem fundamentally different or inexplicable in terms of our present construct of the elementary physiology of integration, then it is probable that that the construct is incomplete or mistaken, even for the levels of behavior to which it is applied” (pp. 134-135).

The Cambridge linguist must have read these lines too but appreciated them less than other passages in Lashley’s essay (see below).

Regarding epistemology and methodology, as said above, Chomsky makes it clear in various sources that the investigation of the grammatical system of generative processes, including its innate schema (Chomsky, 1968), is the specific task of linguists. Psycholinguists should concentrate on language performance and testing the grammatical intuitions of the linguists (a form of introspection). Curiously, his proposed division of work, incoherent with the suggestion that linguistics is a branch of cognitive psychology, has been readily accepted by many psycholinguists in the 1950s and 1960s, and even in later years (see, for example, Pinker, 1994, and Hauser & Bever, 2008).

 If it were really applied, this suggested division of work would determine a major epistemological conundrum in the study of language. Indeed, if language competence cannot be ascertained through psychological studies of language performance because the latter is corroded by extraneous factors, that would mean that language competence, left to intuitive reconstruction, would be *de facto* out of reach of empirical science.

 Chomskyan linguistics also carries a strong distaste for learning and associationism in general, a probable consequence of Chomsky’s Platonic doubt regarding even the possibility of inductive learning and knowledge (which he claims to have learned from the philosopher and logician Nelson Goodman, another one of his teachers at the University of Pennsylvania).

 Chomsky’s (1959) harsh criticism of Skinner’s (1957) *Verbal behavior* is often cited at face value. Several criticisms of Chomsky’s review have been published rightly denouncing his lack of knowledge of the basic learning principles and theories and his rejection of a functional analysis of language in favor of a purely formal one (e.g., MacCorquodale, 1970; Richelle, 1976). However, these critics addressed only a part of the problem.

 Chomsky rejects the Skinnerian notion of verbal operant as purely pragmatic and entirely dependent on reinforcement contingencies. Skinner indeed defines verbal behavior as behavior reinforced through the mediation of other persons. Reinforced in this context means that the probability of occurrence of a given verbal response will be influenced by the reactions it may have provoked on preceding occurrences. Chomsky also rejects the oversimplified scheme proposed by Skinner for classifying verbal behaviors (technically labelled verbal operants because they were theorized with reference to the principles of operant conditioning, a form of learning particularly studied by Skinner; cf. Skinner, 1938).

Skinner’s scheme for analyzing language includes the notions of mands (requests), echoics (imitations), tacts (utterances whose probability is increased in the presence of a given stimulus; Skinner, 1957, p. 82), intraverbals (word associations), and autoclitics (descriptive propositional devices that can be utilized for asserting, emphasizing, negating, qualifying, or quantifying one’s own verbal production; relational and manipulative autoclitics serving to regulate sentences morphosyntactically act through grammatical “frames”; and autoclitics of composition organizing sentences in paragraphs and texts).

Chomsky (1959) rightly points out that more is implied in sentence structures than the insertion of lexical items into grammatical frames. He insists that no approach of language refusing to take into account the existence of deeper processes with respect to surface structure, can deal with the grammatical organization of human language. Chomsky is keen to recall a communication of Lashley (1951). In that address, Lashley warns his fellow scientists that conditioned response chaining (a prevalent behaviorist technical concept) cannot account for the sequencing of complex behaviors. He argues instead for a cognitive account in which serial order behavior is controlled with a central plan. Regarding language, Lashley suggests the need to reckon with a particular mechanism imposed on the speech act along its production but not corresponding to simple associations between successive words. He mentioned numerous cases where the wording of early portions of a sentence is determined by the character intended in later portions but confesses that he could not come with a precise theorization.

Chomsky (1959) assimilates the regulatory mechanism suggested by Lashley to “a grammar of the language correctly formulated” (p.56). In a further elaboration regarding behavior planning, Miller, Galanter, and Pribram (1960) refer to Chomsky’s first grammatical model (1957) as an example of a cognitive plan in the organization of complex behaviors. It would seem indeed that behaviors beyond the reflex level need to be planned. The question is to identify the processes involved. More recent interpretation of Lashley’s insight insists on the hierarchical dimension of the planning involved in sequencing complex behaviors such as language, motor sequences, or music (Rosenbaum et al., 2007; Fitch & Martins, 2014).

Actually, Chomsky and followers have overinterpreted Lashey’s suggestion in favor of their particular conceptual view. A careful reading of Lashley’s original text reveals two things relevant to the present discussion.

First, Lashey uses the term “hierarchy” (in its plural form) only one time in the text (on page 121). Thereby he means the multiplicity of integrative processes taking place from top to bottom in the central nervous system in the temporal organization of a speech act: “…the order of the vocal movements in pronouncing the word, the order of words in the sentence, the order of sentences in the paragraph, the rational order of paragraphs in discourse”(ibidem); not a lot to do, it seems, with hierarchical levels of analysis from deep to surface sentence structure in generative grammar (see next section)).

Second, when Lashley writes of plan and integration needed to account for the organization of sentences (as well as other complex sequential behaviors), he is not referring to a particular formal device. Lashley leaves open the question of the precise nature of the integrative mechanism, indicating solely (but interestingly) that some data (e.g., words misplaced in sentence order, slips and interferences, spoonerisms) suggest that the mechanism in question is relatively independent both from the motor units and the general thought structure. Venturing just one step further in closing his contribution, he speculates that the mechanism in question could proceed as follows: one neural subsystem is held in a state of partial excitation prior or during the execution of a sequential task while it is being scanned by another subsystem; the first subsystem then resumes a full excitation state and transforms the memory traces into a succession (a good definition of a buffer mechanism).

In ending this section, I believe of interest recalling two particular points in Skinner’s (1957) theory of language. Skinner argues that the morphosyntactic treatment of language is sequential (an idea found also in behavioral scientists from a different tradition; e.g., Osgood, 1971), consisting mostly of self-editing, i.e., the speaker’s online monitoring of his own production. “(verbal) Responses cannot be grouped or ordered until they have occurred or at least are about to occur (1957, p. 332). Skinner goes too far, in my view, suggesting that morphosyntactic treatment is linear (although he does not use this word), which is not the same thing as sequential - see Chapter 4); thus claiming that a sentence may be viewed as a behavioral chain with no structure or organization beyond what is implied by conditioned associative connections between elementary items.

Skinner also suggests that inflectional morphology operates online in a self-editing way. The grammatical morphological markings correspond to particular semantic characteristics of the entities referred to and to proximal or distal associations between words in utterances.

 1.2.2. The independence of syntax

In *Syntactic structures*, Chomsky intended to confirm the study of syntax as a formal discipline autonomous and independent of meaning and on which probabilistic models cannot give relevant insight.

This double objective was reached, arguably, contrasting a small series of sentences such as:

1. *Colorless green ideas sleep furiously.*
2. *Furiously sleep ideas green colorless.*
3. *Have you a book on modern music?*
4. *The book seems interesting.*
5. *Read you a book on modern music?*

*6. The child seems sleeping*.

 Chomsky said that any speaker of English would readily agree that sentences (1) and (2) are equally nonsensical but that only sentence (1) is grammatical, i.e., corresponds to the morphosyntactic standards of the English language. Similarly, among the other sentences, only (3) and (4) are grammatical sentences. Chomsky concluded: “Such examples suggest that any search for a semantically based definition of ‘grammaticalness’ will be futile” (1957, p. 15), and: “Grammar is best formulated as a self-contained study independent of semantics” (ibidem, p. 106).

In the same way, he argued that the notion of “grammatical” in English should not be equated with that of statistical approximation. Very likely, neither sentence (1) nor (2), nor any part of these sentences, has ever occurred in an English discourse. Yet, though both nonsensical and statistically irrelevant, (1) is grammatical while (2) is not.

 Viewed from the level of grammaticality judgement, Chomsky’s analysis is unquestionable. But is it all there is to these sentences?

Regarding statistics, first.

Numerous analyses have documented statistical properties of human languages. Zipf (1949), for example, showed that in language use there are a few very common words, a middling number of medium frequency words, and many low frequency items (see Mandelbrot, 1954, 1968, for a more general relationship between rank and frequency). On this basis and other observations, Zipf conjectured his principle of least effort (maximal efficiency for a minimal energy investment): both the speaker and the hearer trying to keep their effort at a functional minimum. Their effort is lessened by having a smaller vocabulary to choose or retrieve from. The same could probably be said of the partially formulaic or idiomatic[[10]](#footnote-10) character of many language exchanges (Bolinger, 1975). In stochastic terms, formulaic refers to the mathematical models which describe the temporal evolution of a phenomenon within a probability space. In so-called fixed formulaic expression (like casual greetings, remarks about the weather, etc.), the transitional probability from one item to the next one is 1 or close to 1 on a scale from 0 to 1. Formulaic expressions are frequent in common phrases and sentences. They contribute to facilitating and speeding up combinatorial language. Kuiper (1996), for example, has illustrated the high proportion of fixed and slightly variable formulaic expressions in the professional speech of auctioneers and sportcasters.

 Numerous sequential preferences have been documented in languages. Hagège (1986), for instance, proposed « a principle of second heavy » for characterizing a common phenomenon in languages. That is a tendency in fixed locutions to place in second position the word that is “heavier” in terms of phonetic envelope, either because it aligns an acoustically lower vowel or a longer series of syllables, than the first word in the binomial. For example, in French, *ici et là* (gloss: *here and there*), *plus ou moins* (*more or less*); in Spanish: *tarde o temprano* (*later or sooner*); in Russian, *tam i sjam* (*there and here*); in English, *flip-flop*, *by guess and by gosh*).[[11]](#footnote-11)

Costermans (1980) has documented sequential preferences in French: for example, the orders animate-inanimate, singular-plural, and definite-indefinite.

A mathematical technique, labelled conditional entropy[[12]](#footnote-12) (from Shannon, 1948), can be utilized for measuring the flexibility of the sequential order in a language. Rao et al. (2008) compared the stochastic structure of sequences of signs randomly sampled from several languages: English, written Sumerian, Old Tamil (a Dravidian language spoken in India), ancient written Indus, Fortran (an imperative programming language originally developed in the 1950s); as well as two biological systems; the first millions of bases on human chromosome 2, and the whole set of sequences of amino acids forming the proteins of bacterium Escherichia coli.

Results show that the linguistic systems have a conditional entropy intermediate between rigid nonlinguistic systems such as the sequences of amino acids in the proteins of Escherichia coli, and flexible ones as the sequences of bases in the human DNA (deoxyribonucleic acid). In English, conditional entropy is around 3 nats on a scale of 0 to 6, increasing slightly with the size of the language sample.

 Chomsky (1957, p.17, note 4) recognizes that one might seek to develop a more sophisticated relation between statistical aspects of language and syntactic structure than the simple approximation model that he rejected in commenting on the sentences above. And indeed there have been many more studies of the statistical dependencies between units in languages (Costermans, 1980; Manning & Schütze, 1999; Chaudhari, Damani, & Laxman, 2011; for reviews and methodological considerations). Chomsky argues, however, that the relation between syntactic and statistical aspects of language can only be properly studied after the syntactic structure has been determined on specific grounds.

 Chomsky (1957) extends the same argument to the relation between semantics and syntax. Let’s take another look at the first two sentences in the examples given above, however:

1. *Colorless green ideas sleep furiously.*
2. *Furiously sleep ideas green colorless.*

It seems justified indeed, here and in general, to separate syntax and semantics. They are distinct dimensions of the language system. It is correct also to remark that sentences such as the above are deprived of conventional global meaning (but not of individual lexical meaning). But it is incorrect to suggest that there could be no semantic relational meaning implied. *Colourless* and *green* are qualities, *ideas* are entities, *sleep* is a state, and *furiously* a characteristics of a state (or a process, or an action). Their semantic relations are perfectly expectable in whichever way they may be expressed (compare the two examples above). In semantic combinations, it can always be expected that an entity receives a qualification, is set in relation with a state, a process, or an action, and that the state, the process, or the action, are characterized one way or the other. It follows that Chomsky’s statement “…grammar is autonomous and independent of meaning …” (1957, p. 17) is correct but does nor exhaust the analysis. The two language components are distinct but semantic relations can be involved as soon as there are associations of words even in nonsensical sequences.

Discontinuity, contiguity, sequential proximity or distance, have no direct relevancy in semantic relational matters. The same relations may hold in whichever order they are expressed providing they can be associated in mental representations. Sequencing, proximity, linear distance, are syntactic notions, not genuine semantic ones. Semantic relations are mental constructions that are not naturally sequenced. For example, the semantic binomials possessor-entity possessed or entity possessed-possessor, agent-action or action-agent, location-entity located or entity located-location, carry the same relational meaning whichever the order of expression. Of course, one is obliged to align these notions when speaking or writing of them, but this is pure convention of expression.

The theoretical treatment of semantics has long caused problems to structural and generative linguistics. In line with traditional structural linguistics, Chomsky in *Syntactic structures* (1957) proposed a formal structural and transformational theory of grammar with practically no semantics at all. Rules in a purely formal system, it was judged, cannot make reference to the meaning. Chomsky even declared the notion of structural meaning “quite suspect” (p.108) and stated: “…only imperfect correspondences hold between formal and semantic features in language” (p.101)[[13]](#footnote-13). However, given that correspondences exist between formal and semantic features and cannot be ignored, he judged reasonable the question to know whether meaning could be assigned directly to the grammatical devices in a more general theory; which announced a further reflection on the topic.

In *Aspects* (1965), indeed, Chomsky (somewhat reluctantly, according to Postal’s recollection; in Huck & Goldsmith, 1995, pp. 126-142; and Lakoff’s, ibidem, pp. 107-119), presented a new conception, known as interpretive semantics (as part of the so-called standard theory of generative grammar). In this theory, semantic relations or roles (or thematic roles, which will be labelled theta roles in further work) interpret (“read”) deep syntactic structures; the latter being considered as determining meaning. Further theorization has modified this point of view somehow (see below), but rather in a cosmetic way according to Postal (in Huck & Goldsmith, 1995).

Postal says, speaking about Chomsky’s *Government and binding* theory (1981) and supportive work in generative grammar: “They have no notion of semantics whatever. They have adopted some notations that *look* (stress Postal) like the predicate calculus and use terms like ‘logical form,’ but in reality none of it has anything to do with the representation of meaning. They’re pretty explicit about that, that there is no requirement that sentences with the same meaning have the same logical form, for example. And their so-called logical forms are filled with all sorts of things that have nothing to do with logic, like actual words from particular languages, and so forth. I said this contentiously some years ago, that it was just a new level of syntactic representation. But Chomsky now says that explicitly” (p. 138).

1.2.3. Evolution of generative linguistics

Generative grammar has evolved substantially between 1957 (the year of the publication of *Syntactic structures*) and 1995 (the *Minimalist program*). And yet the same basic interpretation of the relations between syntax and semantics has been maintained.

The conceptual endeavor starts with the publication of the book *Syntactic structures* in 1957, based in part on Chomsky’s doctoral thesis (1955, but published only in 1975), at the University of Pennsylvania under the direction of Zeilig Harris, an important figure of structural linguistics (Harris, 1951).

In *Syntactic structures*, a grammar is viewed as having a tripartite structure: (1) a sequence of phrase structure rules corresponding to a tree diagram as illustrated in Figure 1 above, regarding the analysis in immediate constituents (the previous structural label that Chomsky changed into phrase structure grammar). This first device operates with rewriting rules of the type “rewrite X as Y”, where X and Y are strings (i.e., linear sequences) of symbols. For example, in “rewrite Sentence as NP (noun phrase) + VP (verb phrase);”[[14]](#footnote-14) (2) a sequence of transformational rules that transforms a string with a given phrase structure into new strings. There are obligatory and optional transformations (see below). When the obligatory transformations have been applied, the derivation yields a kernel of basic sentences (simple active declarative affirmative) with no complex verb or noun phrase; and (3)a sequence of morphophonemic rules converting strings of morphemes into strings of phonemes.

A simple phrase structure grammar has limitations, however. It runs into difficulty when applied to a series of common structures in English, like inserting a constituent inside another one (for example, *people* in *Money does not render people happy*), joining clauses to form compound sentences, the auxiliary (Aux) verbs (e.g., *is* + *being* + *taken*, *will* + *take*, etc.), and the active-passive relationship. New rules are needed to account for these syntactic phenomena. They will increase the depth of the structural description.

The new rules led to a new conception of syntactic structure; a transformational grammar viewed as more powerful than a simple description in terms of phrase structure. Chomsky distinguishes between obligatory and optional transformations. Obligatory transformations (for example, those regarding the correct positioning of the affixes (*en*, *ing*, etc., with regard to a verb stem) must apply to every derivation or the end product will not be a grammatical sentence. Optional transformation (like the passive one) may or may not be applied in particular cases, but either way the derived sequence is a grammatical sentence.

 A transformation is defined structurally by an analysis of the strings to which it applies and the change that it produces on these strings. For example, the passive transformation applies to strings of the form NP-Aux-V (verb stem)-NP, and has the effect of interchanging the two NPs, adding *be* + *en* to Aux, and inserting the preposition *by* in front of the final NP. The negative transformation introduces the negative element (e.g., *not*) after the second morpheme of the verb phrase if this phrase contains at least two morphemes or after the first morpheme if the phrase contains only one morpheme. The interrogative transformation yielding yes/no questions operates on strings of the type NP-V… (e.g., *They arrive.*) and has the effect of interchanging the first and second segments of these strings (e.g., *Arrive they?\**[[15]](#footnote-15) that becomes *Do they arrive*? with the application of additional obligatory rules). The invention of the transformational level establishes a difference between deep and surface levels in the generative analysis of sentences that will be worked out in more detail in *Aspects* (1965).

There has been a discussion in the literature regarding the paternity of the idea of transformation in linguistics. Several commentators have mentioned the proximity of Chomsky’s transformational version and the notion of kernel sentence to Harris’ theories (for example, Katz, 1981). However, there is an important difference in the two versions. For Harris (1951, 1970), a transformation relates to surface structure forms. For Chomsky, as seen before, it is a device mapping one phrase-structure tree to another (Barsky, 1997, 2010). Harris thought that it was not necessary to increase the complexity of the hierarchical structure as the dependencies between words in the set of sentences were sufficient for determining transformations. This is a passing indication of the resistance that Chomsky’s move to increase the hierarchical depth of syntactic analysis has met in some linguistic circles.

Fig. 1.2 illustrates schematically how the grammatical components outlined in *Syntactic structures* are organized.

 In *Aspects of the theory of syntax* (1965), considered for several years the standard generative theory, Chomsky puts forward a more comprehensive theory of transformations. A base component to the grammar is distinguished that is roughly comparable to the phrase-structure component of the previous theory. It opens on the semantic options that give access to the meaning of the sentence. Meaning is seen as depending upon the syntactic base of the language, which it is supposed to interpret. The rationale is to submit the meaning of the sentences to the same kind of precise formal analysis as their syntactic structure.

 A transformational component is maintained in the theory but it is restricted to the obligatory transformations. The optional ones identified in *Syntactic structures* are dropped. The difference between a declarative and an interrogative sentence, or between a passive and an active one, for example, is now described in terms of a choice made in the base component. For example, in the derivation of a passive sentence, there is a base rule of the form: NP + VP + NP (+ agentive). The element “agentive” distinguishes the strings underlying passive sentences from those underlying active ones. Then, there is an obligatory transformation operating under the specific instruction of the base that leads to the derivation of the passive form.

The base rules are responsible for generating a large set of phrase markers which yield the deep structure of the system. Once these markers have been converted into derived phrase markers, they give the surface structure (grammatical level) of the sentences (which has still to be turned into a real sentence by morphophonemic rules and phonemic realization).

 In *Aspects*, thus, the semantically relevant grammatical notions are explicitly defined in terms of deep structure relations. As said, the idea is that there is a syntactic deep structure that fully determines meaning. The semantic component is “projected” on the phrase markers and “reads” the grammatically relevant properties of the syntactic elements. The paternity of the projection rules is attributed to Katz and Fodor (1963). The rules apply from bottom to top in the tree of the base structure amalgamating the lexical elements dominated by the same node and attributing meaning to this constituent. The condition for the amalgam is that the elements satisfy a mutual requirement for compatibility in terms of syntactic and semantic features. Projection goes on upward until reaching the symbol S. At that level, one obtains the meaning attributed to the highest node in the tree, i.e., the whole sentence.

 In Fig. 1.3, the major components of the *Standard theory* as defined in *Aspects* and their relationships are sketched.



 Interpretive semantics, as formulated in *Aspects*, left a number of people unsatisfied including among Chomsky’s close colleagues. Some of them initiated a theoretical challenge that came to be known as *generative semantics*. They identified a double contradiction in the standard theory: some synonymous sentences with different syntactic structures have differing deep structures; and some non-synonymous ones with similar syntactic structure are considered as corresponding to the same deep structure. This motivated a major revision of Chomsky’s theory of deep structure.

 The new idea was that the role of derivational constraints and transformations is to relate semantic representations directly to the surface structures of the sentences, hence the name of generative semantics given to this theory. It represents a kind of bottom-up approach contrary to the more deductive kind of interpretive semantics. Major protagonists in this movement were McCawley (1968, 1971), Lakoff (1971a; 1989), and Postal (1970). A detailed review of their proposals and the differences between them is beyond the scope of the present text (see Maclay, 1971; Brame, 1976; and Huck & Goldsmith, 1995, for detailed presentations and discussions).

 As indicated, the standard semantic theory in transformational grammar is interpretive. This left syntax at least partially autonomous with respects to semantics. In opposition, generative semanticists argued that semantic interpretation is generated in deep structure (which complicated the so far simple design of this grammatical component) and that lexical insertion could itself be the object of transformational rules, before being moved to sentence surface. Although technically more complex, the new theory had a more intuitive appeal in that the form of the sentences was considered to be derived from their meaning.

 More precisely, the projection rules of the generative model in *Aspects* were abandoned. Instead, semantic representations were described with notations inspired from symbolic logic. The base rule was replaced by a scheme involving a content element (verb, adverb, or preposition) and one or several nouns phrases. The logical theory was that a sequence in deep structure corresponds to a predicative structure with n-places and an argument for each of the n-locations.

Ruwet (1972) argued that generative semantics is based on a questionable postulate, i.e., the conviction that it should possible to build a universal semantics. He quoted extracts from Chomsky’s writings expressing doubt on this possibility. Yet, such an undertaking has been on the way for some time. Semantics emerges from an interface with general cognition which few people would say is not universal. Then, *ipso facto*, semantics ought to be universal to a large extent. I will return to this issue in Chapter 5 and the related question of the so-called linguistic relativity hypothesis.

Chomsky’s (1971) reaction to the proposals of the generative semanticists was very negative. Ross (in Huck & Goldsmith, 1995, p. 15) remembers: “For Chomsky, semantics could not be generative.” Chomsky pretended that in a grammar as the one proposed in *Aspects,* there is no directionality involved, which is clearly not true. He also stated that the differences in the proposals of the generative semanticists with regard to the standard theory were simple “notational variants;” his own version being more parsimonious in that it did not necessitate making particular hypotheses concerning the nature of the semantic representations.

 Even if officially rejected, the proposals of the semantic generativists left traces in the evolution of generative linguistics. Chomsky (1971, 1972) conceded a more important role to surface structure in the interpretation of semantic structures (already discreetly mentioned in the final part of *Aspects*, p.224, note 9). For example, it was recognized that the order of quantifiers in surface plays a role in meaning assignment. The sentences (from Chomsky, 1965) *Everyone in the room knows at least two languages* and *At least two languages are known by everyone in the room*, are not synonymous. The suggestion was that the two possible interpretations are latent in deep structure. The order of the quantifiers in surface eliminates one of them.

 In the so-called *Extended standard theory* (1979), Chomsky clarified his new position. He suggested that the only (though essential) semantic contribution of deep structure to the meaning of a sentence, is in representing thematic relations between verbs and noun phrases. Other semantic properties are determined by rules that apply in surface structure (1979, p. 163). Among these rules, besides the order of quantifiers, there are pronominal coreference (i.e., the relationship between a pronoun and the noun to which it refers), presupposition (i.e., the “old” information linking the sentence to nonlinguistic context or to a previously formulated information), and the focus (the “new” information provided in the sentence).

Later in the same opus, semantic interpretation is pushed one step further towards sentence surface. This is presented as a theoretical progress in generative theory. Semantic interpretation is now viewed as proceeding step by step with a derivational course like syntax. A new concept is introduced, that of *trace* (inspired from generative semantics). A trace (symbolized *t*) is an empty (phonologically-null) element that allows keeping track of the position of a constituent in the structure when this constituent undergoes movement. For example, in the production of a wh-question, a sequence such as *You like drinking what?* is transformed into *What do you like drinking?* The wh-word *what* being moved to the front of the sentence*.*The position from which the wh-word is moved (in this case, the position of the direct object of *drinking*), is considered to be occupied by a trace. The second sentence above may thus be considered as corresponding to the structure What *do you like drinking t?* In the example,the *t* formulation maintains a formal link between the direct object of the verb in surface structure and the thematic patient of the action attributed to the agent in deep structure. The thematic relation is considered to be transferred from deep to surface structure at the same time as the trace. This allows the semantic representation to figure also in surface (under the name of “logical form” of the sentence; also, a clear reminder of generative semantics).

 Other modifications in the *Extended standard theory* concern some simplifications of the transformational component as well as the categorical sub-component of the base according to a system of context-free rules (i.e., rules applying regardless of the syntagmatic environment) replacing the context-dependent rules of traditional phrase structure analysis. The new rules are formulated according to the X-bar syntax introduced by Jackendoff (1977); a notational system more stratified than the previous analyses in simple constituents. It allows deriving the phrases from the formal properties of their head (node), accounting more clearly for their composition.

Fig; 1.4 schematizes the major links between the sub-components in the *Extended standard theory*. It incorporates the level of “Initial syntagmatic marker” introduced by Chomsky (1976) instead of deep structure.

 Jackendoff (1978) pleads in favor of an extended version of the semantic treatment in generative theory, stressing that beyond projection rules semantic structures have their own conditions of good formation in the same way as syntactic structures. The origin of the former is found in general cognition. However, Jackendoff, in good Chomskyan orthodoxy, maintains that the identification of the semantic structures must proceed from the syntactic ones without altering the integrity of the syntax component in the generative system. In a further work (Jackendoff, 1983), semantic structures disappear altogether. They are replaced by conceptual structures declared to be common to language, visual judgements, complex motor behaviors, etc. In Jackendoff (1990), then, various categories of concepts are identified that cut across conceptual structures. He distinguishes between E-concepts, or concepts referring to “…something out there in the world “(p.7) and I-concepts, which are “within one’s head” as private entities (ibidem). This distinction parallels that of Chomsky (1986) between E-language and I-language. I-language is the inner language, the one represented in people’s brain. E-language is the external language used in human communication (a distinction seemingly reminiscent of the previous competence-performance distinction).

 In another theoretical trend, Bresnan (1978) and Kaplan and Bresnan (1985) have suggested to attribute a more important role to semantics in generative grammar, this time through the lexical component. Their proposal known as the “lexical-functional grammar” establishes a correspondence between lexical entries and surface structure. Syntactic rules identify grammatical functions with the help of morphological marks and particular constituents but without transformation rules. Two levels of syntactic description are proposed for each sentence. The syntagmatic configurations are represented in a structure of constituents (the C-structure) that has the form of a traditional syntagmatic tree-like analyzer, on which the phonological representation operates. The grammatical functions correspond to the other level of description, labelled functional structure (F-structure), which supplies the only input to the semantic component. The F-structure contains the information regarding the relations between predicate and argument(s) as well as the indications concerning time, aspect, number, gender, definitiveness, etc. The linking of predicates to arguments is accomplished by the mediation of grammatical functions (subject, object, etc.).

One is still in an orthodox Chomskyan framework as semantic interpretation remains subordinated to the syntactic dimension (actually Chomsky, 1970, had suggested a lexicalist hypothesis oriented in the direction of what became lexical-functional grammar). There are some differences, however, with previous treatments. On the one hand, the rules in the C and F structures must be able to generate all surface structures as there is no transformational intermediate. On the other hand, and concomitantly, the lexical entries specify for each term in the sequences of the base its formal category (noun, verb, etc.), the predicate-argument(s) structure in which it can be inserted, and within this structure which grammatical function it may have.

 A further elaboration of generative grammar is presented in Chomsky (1981, 1982), under the name of *Government and binding theory*. It involves the revocation of the notion of simple rule in favor of a conception in terms of principles (general types of rules) and parameters that may vary between languages. Following the suggestions in lexical-functional grammar, Chomsky indicates that there is no reason actually to have in the base both rules of syntagmatic structure and grammatical definition of the lexemes; hence the former may be eliminated (as the later cannot).

Transformations are limited to move-alpha and modify-alpha. The distinction between formal and substantive universals is dropped. It is replaced by a contrast between formal (that may vary parametrically) and absolute universals (that may not).

 Another important theoretical consideration expressed in *Government and Binding* is language modularity. Chomsky (1984, 1988) formulates the hypothesis that the structure of mind is modular, constituted by a system of relatively autonomous sub-systems each one with its own characteristics. Besides this general modularity, Chomsky distinguishes modularity internal to grammar. Several sub-theories within generative grammar are presented as being autonomous (for a criticism, see, for example, Russel, 1987): the X-bar theory, the limit theory (dealing with restrictions on the possible movements of elements in sentences), the government sub-theory (dealing with the hierarchical relations between syntactic constituents), the case theory (regulating the attribution of grammatical functions to noun phrases), the theta-theory (assigning thematic roles to syntactic configurations), the binding theory (dealing with anaphora and pronominal coreference), and the control theory (dealing with the antecedent of the empty symbol PRO, i.e., the null element occupying the subject position in embedded infinitive clauses (for example, *John wants PRO go to the country*. *PRO go to the country bothers Peter)*)*.*

 A schematic representation of the components of the generative model of *Government and binding* is given in Fig. 1.5.



 The base component, following the lexical hypothesis impelled by Bresnan (1978), incorporates the lexical semantic elements with their grammatical properties in a feature system that classifies them differentially. It goes in pair with the abandon of the idea that only deep structure determines meaning. Structure D is a representation of relations of syntactic dependence and sub-categorization in sentences (denoting the necessity for a lexical item of a given category to require or allow the presence of the argument-s with which it co-occurs). Structure S is a representation of the hierarchical relations between constituents. The logical form is retaken from the preceding theoretical formulation including the traces that keep in “memory” for the logical form the original position of the elements displaced by the transformation move-alpha.

 The theta-criterion is relevant to the relationship between syntax and semantics. It states that each element of the logical form must receive one and only one theta-role and that each theta-role, determined by the lexical properties of one element of the head of a phrase, must correspond to one argument of the sentence. Every sentence containing a main verb (noncopular therefore) is grammatical providing that the theta-criterion is satisfied. Copular verbs (e.g., *be*, *appear*, *look*, *sound*, *get*, etc.) do not assign thematic roles. To the theta-criterion, it is necessary to add the projection principle (reminiscent of *Aspects*). This principle stipulates that at each syntactic level, the representations are projected from the lexicon, meaning that they must be consistent with the sub-categorical and theta-characteristics of the lexical items.

The theory presented in *Government and binding* may give the impression that semantic information plays a more important role in sentence treatment than in previous theoretical exercises; in particular with the incorporation of the lexicalist hypothesis. This is only an impression, however. In *Government and binding*, like previously in generative grammar, the semantic relational dimension remains subordinated to the syntactic one; this time through the grammatical role attributed to the lexicon.

 In 1995, Chomsky published what appears to be the last major revision of his generative theory, under the name of *Minimalist program*. This version retains the major theoretical properties of the principles and parameters model. It is said to be simply a continuation of what has been undertaken from the earliest years (Chomsky, 2011). However, an important effort of parsimoniousness is made. Various analytical procedures are eliminated. New generalizations appear that subsume distinct principles in previous theory. Grammar is said to rests on two pilasters: a mental lexicon and a computational component. The former is defined in a way that is reminiscent of the lexical-grammatical account in *Government and binding*. Mental lexicon is made of a large set of items. Each member of the set is a complex of phonological, semantic, and morphosyntactic properties that can be represented by a whole series of features (e.g., occlusive, noun, number, etc.). The computational component operates with a double mechanism: “move” (displacement) and “merge” (combination), and there is an attempt to reduce the former to a special case of the latter.

Merge is an operation that takes a pair of syntactic objects (either lexical items or sets that have been themselves formed by merge) and replaces them by a binary structure of a higher rank (for example, combining an article and a noun in a NP; a subject and a VP to create a sentence). Chomsky (2001) distinguishes further between external merge (whenever the two lexical items are separate objects and internal merge when either one of them is part of the other).

The merge component contains rules that analyze sentences into hierarchical tree-like structures, the units of which are phrases and lexical formal categories. Whether or not any two such entities can be merged is determined by a “feature-checking” device, verifying whether the properties and dependencies of the elements to be merged match one another. Congruence between these sources works to ensure sentence well-formedness at a basic, “canonical,” level. The move procedure has the capacity to create recursively an infinite series of syntactic structures.

The *Minimalist program* reinterprets also the notion of level of representation from the previous theory. One dispenses with structures D and S, as autonomous realities, in favor of a more direct interface between logical and phonological forms.

 Chomsky (1995) also makes his a suggestion of Kaynes (1994) that the universal linear order in sentences is SVO (subject-verb-object), which happens to be modified secondarily into other linear orders in approximately 60% of the languages on the planet.

In a collective paper published in 2002, partially inspired from Chomsky (1995), Hauser, Chomsky, and Fitch define the language faculty as made of three major components: (1) a sensory-motor system; (2) a conceptual-intentional system; and (3) a computational mechanism allowing recursion. Only the last one is specifically human. It constitutes entirely and solely the faculty of language in a narrow sense, i.e. the part of language that is uniquely human and uniquely linguistic (as opposed to a “broad faculty of language” regrouping parts of language not unique to humans or not uniquely involved in language).

 This is a curious claim. First, previous extracts from not too remote Chomsky’s writings (for example, in Chomsky, 1995, p. 167), suggest the need to distinguish the language faculty both from a conceptual system and a system of pragmatic competence. Second, the title of the collective 2002 paper mentions: “How did it evolve (i.e., the language faculty)? As is known, and has been documented above, Chomsky is a convinced “discontinuist.” He has a long history of repelling the evolution question as irrelevant to generative grammar. Chomsky seems to have rallied lately the exaptation position of Hauser, i.e., the view that the human capacity for language and the language faculty are exceptions with regard to standard Darwinism, due to a particular brain development in humans that came over for other reasons than language (see also the notion of exaptation in Gould, 1991) (for an opposite Darwinian view, see Pinker, 1994, and Pinker & Bloom, 1990).

Third, numerous aspects of human languages are not recursive, in particular: phonology, morphology, case markers, grammatical concord (Pinker & Jackendoff, 2005). And fourth, recursion, though perhaps absent from animal communication (see Evans, 2014, for a contrasting opinion), is found in visual cognition. Hence, it cannot be the sole evolutionary development that would have brought the language faculty to humans (Pinker & Jackendoff, ibidem).

 Fitch, Hauser, and Chomsky (2005) straightfully reject the criticism of Pinker and Jackendoff as based on a misunderstanding of their position, for a part, and irrelevant for the rest. Interestingly, at the end of the paper, Fitch et al. (2005) suggest that progress in understanding the evolution of language will require more empirical research grounded in comparative biology and more interdisciplinary collaboration, which for coauthor Chomsky may have sounded like a painful retreat from his previous authoritative anathema on the “so-called studies of language” in animals.

 Everett (2005, 2008) has analyzed the language of a tribe in the Amazonian region of Brasil. He reports that the language of the Pirahan Indians is completely deprived of clause or phrase recursion. There is no clause inserted within another one in a clause complex, no phrase inserted within another one, no coordination or disjunction of clauses; whereas the ideas expressed testify to conceptual relationships inserted into one another (in narratives, for example).

Everett (2008) supplies long corpus extracts (pp. 91-93, 119-121, 123-126, 135-136, 204) documenting a linguistic expression typically made of sentences that are relatively simple in terms of syntactic structures (frequent declaratives and imperatives, mostly active ones) following each other paratactically. The sentences produced are communicatively functional even if leaving some interpretation to the context. On this basis, Everett (2008) rejects the posited universal character of linguistic recursion (and the hypothesis of a faculty of language in the “narrow sense”) while insisting that recursion is part of standard human intelligence. Everett (2008) is convinced that other languages, many in fact, or families of languages not studied or not yet systematically studied among the numerous languages and dialects in New Guinea and Africa, for example, may have characteristics similar to the language of the Pirahan Indians.

 Everett (2008, 2012) proposed also that the absence of recursion in the language of the Pirahan Indians originates in a specific cultural feature, which he labels “the immediacy of experience principle.” The Pirahan Indians are only linguistically interested in what can be asserted (that is, logically, a declarative proposition presented as true) and refers to a scene to which the speaker has been a direct witness. In linguistic recursion (for example, a centrally embedded relative), the recursive element is not assertive. Instead, the Pirahan Indians will formulate independent clauses in succession (without marking coordination explicitly either).

Everett’s results and interpretation have been harshly called into question (e.g., Nevins, Pesetsky, & Rodrigues, 2007, 2009a; see also Everett’s replies, 2007, 2009; and Nevins et al.’ s reply to Everett’s reply, 2009b). Everett’s contradictors argued that he had misanalysed his data and that Pirahan language does have recursion. Everett countered in declaring Nevins et al.’s criticisms wrong on just about every single point; adding that they had conducted no independent fieldwork, and did not suggest any experiment to verify the accuracy of his (Everett’s) claims.

 A relevant question discussed in Bickerton (2009), is the theoretical status of recursion as a language process. Bickerton insists on a distinction between recursion and iteration. Iteration means repeating something an arbitrary number of times whereas recursion implies embedding something within another instance of itself. Bickerton (2009) suggests to dispense with the notion of recursion altogether and replace it with iteration. He considers that Chomsky is responsible both for the introduction of the recursion concept in linguistic theory (1957) and for its elimination (according to Bickerton’s analysis of the *Minimalist program*, 1995).

 Bickerton finds it curious that Chomsky having eliminated recursion from his latest theory, signs a paper in which recursion is presented as the sole component of the language faculty in a narrow sense. Bickerton speculates that Chomsky cannot jettison recursion and replace it by iteration, though his (Chomsky’s) recent theory shows that even the generation of the most complex sentences does not require it, because he is viscerally attached to the idea that syntax is the central part of language. Iteration could not be said to be required only by language, then recursion had to be maintained; hence the compromise that may have served as a basis for the collaboration with Hauser.

 Bickerton (2009) is convinced that merge is an iterative process and that syntax, no matter how complex it can be, is simply a function of lexicon + merge.

 Regarding the question of the evolution of language, given that animals have merge-like (iterative) processes in their cognitive abilities, it might be asked why they do not use it in communication. The answer of Hauser et al. (2002) is that this is because this process in other species is a modular system designed for noncommunicative functions (e.g., navigation) and “impenetrable” with respect to other systems, including communication.

 Bickerton (2009) has another explanation. Other species but the human one cannot apply merge to their communication because “…the units of their communication, in contrast with words, are holistic, symbolic, and non-referential…Since they are the equivalent of sentences rather than words, and since each unit is situation-specific and designed less to communicate than to improve the caller’s fitness, no earthly purpose would be served by concatenating them via merge or anything else; …in principle, no syntactic precursor can exist in the absence of words or word-like units” (p. 543-544).

 This is an interesting analysis but it is based on a very imprecise knowledge of animal communication. There are today a large number of controlled and replicated observations demonstrating that referential communication does exist in a number of species of mammals (e.g., spotted hyaenas, California ground squirrels, diverse subspecies of monkeys, prairie dogs; East & Hofer, 1991; Owings & Leger, 1980; Seyfarth & Cheney, 1993; Slobodchikoff et al., 1991). Even more to the point, carefully controlled experimental studies have shown that large repertoires of word-like units can be learned by a variety of higher species (apes, bottlenose dolphins, sea lions). Their learning and use of these entities appear to correspond rather well to those of humans when analyzed in terms of semantic features (Gardner & Gardner, 1975; Fouts, 1973; Fouts, , Fouts, & Van Cantfort, 1989; Miles, 1990; Patterson, 1980; Herman, Richards, & Wolz, 1984; Schusterman & Gisiner, 1988)[[16]](#footnote-16). It is not known whether bottlenose dolphins and sea lions use similar or corresponding units in their natural communication.

 Such a lexical ability even if only virtual (i.e., documented exclusively in laboratory studies) may constitute a kind of reserve for language evolution conceived either in a Darwinian sense or according to a theory like punctuated equilibria (Eldredge, 1985).

 If Bickerton’s analysis (2009) can be corrected for his uninformed assessment of word capacity in higher nonhuman species, his definition of syntax as a function of lexicon + merge, opens a door in the generative camp for the proviso that primitive forms of syntax are possible in infrahuman species. As Bickerton (2009) notes, merge is a distributional process not a representational one. This means that it has no preconceived structures but instead build them from scratch. It might be thought that it is accessible to organisms with lower cognitive capacities. As indicated, real word-like units exist in infrahuman species. This means, transposing from Bickerton’s own reasoning, that elementary genuine forms of syntactic ability are possible. And indeed they have been observed experimentally in the language comprehension of bottlenose dolphins and sea lions (Herman et al., 1984; Schusterman & Gisiner, 1988).

 In a recent book written with Berwick (Berwick & Chomsky, 2016; see also Berwick & Chomsky, 2017), Chomsky clarifies his position regarding the evolution of language. Berwick and Chomsky repeat that there is a faculty of language interacting with two other mental organs, a sensory-motor system for externalization, and an intentional-conceptual system for interpretation, planning, etc. The core of the language faculty is the merge principle. Its reiterative application allows the generation of an unlimited number of hierarchically organized and potentially recursive strings of words.

 The key question is how merge came into existence. Human have it and nonhumans don’t as shown (allegedly - my addition) by the research on animal language. Regarding theories of evolution, Berwick and Chomsky (2016) favor a nongradualist account. Modern language could have appeared some 60,000- 80,000 years ago[[17]](#footnote-17). Merge was the result of a rewiring of a particular region of the brain (Brodmann areas 44 and 45). Lastly, they argue that the primary function of language is for thought (as internal mental tool) rather than for communication that emerged only as a secondary process, which (in passing, Chomsky, 2011) suggests that the speculation and the research about the origin of language as a communication system is on the wrong track.

Leaving aside Chomsky’s usual contempt for communication and psychobiological research on this matter, the statement regarding the prevalence of thought in language function is ambiguous. Historically, in the evolution of the hominids thought may have preceded elaborated language. Phylogenetically, language abilities in animal species are based on particular cognitive abilities. Ontogenetically, nonverbal thought invariably precedes language. However, verbalized thinking processes in children need to await the development of speech to become operational and this takes five or six years, as already documented in the classical studies of Vygotsky (1962), Luria (1961), and Sokolov (1971).

 But all that does not prove that the primary function of language is internal. Berwick and Chomsky’s (2016) pronunciamento is another manifestation of their exclusive interest for I-language (internal language) at the expense of E-language (i.e., external, communication language). Everett (2017) rejects this conception as preposterous. He argues that language is a tool for communication and not primarily a means of thought expression. He adds that hierarchical, recursive grammars are later “embellishments” that are neither necessary nor sufficient to define human language.

 CHAPTER TWO

 PSYCHOLINGUISTICS’ EARLY ORIENTATION

 Moderrn psycholinguistics or language psychology holds a birth certificate. It was officially created at a summer seminar held at Cornell University (New York state), in 1951, with the participation of noted psychologists (C. Osgood, J.B. Caroll, G. Miller) and linguists (T. Sebeok, F. Lounsbury). It was followed by another seminar in 1953, from which Osgood and Sebeok edited a book entitled *Psycholinguistics, a survey of theory and research problems* (1954), containing an ambitious research program inspired by a wish to reach a synthesis between learning psychology, information theory, and linguistics.

The research program proposed in Osgood and Sebeok (1954) was relatively equilibrated and even perhaps leaning on the behavioral side.

Under the name of sequential psycholinguistics, there was a research program particularly concerned with transitional probabilities in language functioning. The aim was to relate the transitions between units in the message with the activities of the speakers during the production or the comprehension of the message.

 Rapidly, however, the field took a more cognitive orientation (i.e., one more concerned with mental entities and processes and less with traditional learning analyses). George Miller, a noted Harvard psychologist, had a major role in this reorientation of the field of language (as a part of more general reorientation of psychology itself). Miller published extensively and his 1956 paper on short-term memory (*The magic number seven plus or minus two: Some limits on our capacity for processing information*), together with his other contributions*,* were very influential.

 Miller and Selfridge (1950) demonstrated an effect of language structure on the recall of strings of words. They found that recall improved in proportion with the nearness of such strings to the statistical regularities of English. Miller and Selfridge interpreted their data in terms of the conception that sentence structure is determined by left-to-right associative dependencies between successive elements, which was the prevalent view in psychology at that time.

 However, a few years later, the interpretation changed radically. In Miller and Isard (1963) and in Marks and Miller (1964), the authors used corresponding stimulus materials to illustrate the specific effect of syntactic factors in the perception and recall of sentences. Fodor, Bever, and Garret (1974), in their *Introduction to psycholinguistics and generative grammar*, speak of “the discovery of syntax” to characterize this conceptual change, while acknowledging that these experiments fell short of demonstrating a fundamental difference between syntactic and semantic sentence treatments. However, they attracted the attention on syntax as a potentially important variable in language psychology.

 Miller came to be convinced that the notion of linguistic competence that was developed by Chomsky during these years had to be integrated into a cognitive conception of language, and particularly morphosyntax. He kept this view during his entire career. See, for example, his contribution to the edited book by Halle, Bresnan, and Miller (1978), where Miller reformulates his belief that linguistic competence has to be a part of a psychologically real language ability, while acknowledging that he does not know how exactly that should be theorized.

 In 1965 (reprinted in Oldfield & Marshall, 1968), Miller published a short but very influential paper entitled *Some preliminaries to psycholinguistics.* Itcontains an explicit definition of the new research agenda for the discipline. In this paper, which could have been quoted directly from Chomsky’s writing, Miller summarized several aspects of human language that, he wrote, should be understood by any psychologist wanting to venture in psycholinguistics.

 As it is impossible to describe English by listing all its grammatical sentences, it means that the sentence of English must be described in terms of rules that can generate them (p. 205). And “Only on the assumption that a language user knows a generative system of rules for producing and interpreting sentences can we hope to account for the unlimited combinatorial productivity of natural languages” (p.206). From the fact that we are usually capable of recognizing that we have made a mistake, it can be inferred that the rules are known implicitly even though they cannot be stated explicitly. It is important to distinguish between knowledge and performance. The psycholinguist’s task is to propose and test performance models for the language user, but “he must rely on the linguist to give him a precise specification of what it is a language user is trying to use” (p. 206). And finally, it is important to keep in mind that there is a large innate component to our language ability. All languages have features in common that are called language universals (ibidem).

 In his book *The language instinct* (1994), Steven Pinker, another Harvard psychologist, reproduces almost word for word Miller’s (1965) suggestions as to what psycholinguistics is about and how language should be studied. Pinker is extremely clear and confident regarding the psychological reality of linguistic categories, generative grammar, and, more generally, Chomsky’s ideas as to the language faculty (except regarding its evolution). He writes “When people learn a language, they are learning how to put words in order, but not by recording which word follows which other word. They do it by recording which word *category* (Pinker’s stress)- noun, verb, and so on - follows which other category” (p.94). “A sentence is not a chain but a tree” (p. 97). And there are formal hierarchies and sets of rules. “The invisible superstructure holding the words in place is a powerful invention (of the language faculty, my addition) that eliminates the problems of word-chain devices” (p.99).

 I invite those who would be tempted to think that these are outdated opinions in the field of psycholinguistics, to read Yang’s (2016) essay forcefully reaffirming the psychological reality of the formal linguistic machinery (see Chapter 7 here).

 **2.1. Experiments in transformational and generative grammar**

 A whole series of experimental works were conducted on various aspects of generative theory following Miller’s advice. The first ones were devoted to assessing the psychological reality of intersentential relations in the family IPN (interrogatives-passives-negatives) as specified in *Syntactic structures* (1957). Other works, conducted later, undertook to investigate several psychological aspects potentially linked to further elaborations of generative grammar. What follows is simply a brief account of these researches. For more details and more extensive discussions, one may see, for example, Peterfalvi (1970), Fodor et al. (1974), and Costermans (1980).

 Miller, McKean, and Slobin (1962; reported in Miller, 1962) proposed native adult speakers of English to match a sentence-stimulus with another one in a determined syntactic form that they had to select from a list containing various syntactic types of sentences (kernels, passives, negatives, and passives-negatives), equivalent for lexical complexity, length, and frequency of appearance in the language. The two lists of sentences were presented simultaneously to the subjects. The dependent variable was the time needed to complete the task. Results showed that the time necessary for matching two sentences was longer when the sentences differed by more than one transformation than for those differing by one single transformation. This was taken as an empirical confirmation of Miller’s (1962) hypothesis that passing from a sentence to another one corresponds to a mental operation more or less complex (and therefore more or less time consuming) depending on the number of syntactic transformations involved.

 Based on a first experiment conducted by Slobin (and reported in Miller, 1962), Miller and McKean (1964) measured the time required by native English speakers to match sentences without transformation (actives-affirmatives) compared to sentences with transformation(s) (actives-negatives, passives-affirmatives and passives-negatives). Subjects were requested to identify a given sentence in a first list with a meaning-related but syntactically different sentence in a second list (constructed with the same vocabulary items as the sentences in the first list). At variance with the preceding experience, the two lists of sentences were presented separately and in succession (first then second list). The subjects were instructed to transform a given sentence from the first list into another syntactic form and then to identify the transformed sentence in the second list. Results showed that the matches involving active-affirmative sentences always required less additional time than other matches. Results also showed that an active-passive difference required more additional time than an affirmative-negative difference. When both transformations are involved, the additional time required is approximately the sum of the two subtypes of sentences separately. However, taken together the experimental observations suggested (strongly, according to the authors) that the additional times were spent on semantic rather than syntactic operations. The only main effect that could be safely attributed to a syntactic operation was the one observed for the active-passive difference in the matching task.

 Marshall (1964, reviewed in Costermans, 1980) assessed the relative structural complexity of sentences from the IPN family in a task where the participants manipulated only one sentence at the time. The word order in each sentence had been changed randomly and the task of the subject was to reestablish the correct order as soon as possible (dependent variable). The time needed to perform the task increased in relation to transformational complexity. The errors corresponded generally to a simplification of the syntactic form of the sentence, i.e., the participants reconstructing a sentence that was not the one presented in a scrambled order but one that tended towards a kernel form.

 Mehler (1963) calculated the number of semantic and syntactic errors made by English speakers in recalling eight different subtypes of sentences of the IPN family to which they had been exposed (kernels, negatives, interrogatives, passives, negatives-interrogatives, negatives-passives, interrogatives-passives, and interrogatives-passives-negatives). More syntactic errors were registered than semantic ones. On the whole, the kernel sentences were better recalled. More generally, syntactic errors were simplifications from a transformational point of view, i.e., the recalls of the participants tended to favor simpler forms (from IPN, to IP, NP, NI, etc., and in the final analysis towards the kernel sentences). Mehler’s interpretive hypothesis, named “coding hypothesis,” was that when one has to memorize a sentence, it is decomposed into a kernel structure plus the transformation(s) involved.

After Chomsky modified his first theory (in *Aspects*, 1965), Mehler (in Mehler & Boysson-Bardies, 1971) reformulated his coding hypothesis suggesting that a sentence is memorized under the form represented in deep structure plus the information necessary for deriving a given surface structure. However, in a similar experiment to that of Mehler (1963), Martin and Roberts (1966) found a significant effect of sentence type attributable in major way to a lower recall of the kernel sentences compared with the other types tested (negatives, full passives, truncated passives, and passives-negatives).

 Savin and Perchonock (1965) studied the treatment in short-term memory of sentences of the IPN family and sentences realized by the so-called emphatic transformation (placing at the beginning of a sentence the units that are stressed to the attention of the interlocutor) and the wh-transformation. Sentences were proposed to the participants with the instruction to memorize them. Following the presentation of each sentence, one or more isolated words with no semantic or syntactic relationship with the sentence were added and the participants had to memorize them as well. The experiment was based on the notion (developed by Miller) of a limited capacity of short-term memory. The more the participants can restitute additional words to the sentence, the less the sentence is supposed to occupy space in short-term memory span. Results showed that sentences produced by two transformations occupy more space in short-term memory that those produced by one transformation.

Jakubowicz (1970) experimented on the IPN family of sentences in French. Subjects were requested to turn a kernel sentence into one of the other forms of the IPN family. The dependent variable was the time needed to proceed. Results did not correspond to what was expected in terms of relative transformational complexity. For example, Interrogative sentences took more time to proceed than interrogative-passive and interrogative-negative ones; passive-negatives took more time than interrogatives-passives-negatives, which is contrary to what can be predicted from transformational theory.

 Slobin (1966) proposed to a sample of native adult English speakers to judge whether a sentence can be considered to be true or false with respect to a picture presented simultaneously. Other independent variables involved in the experiment were the semantic reversibility or not of the sentence and the syntactic form (kernel, passive, negative, and passive-negative). Results showed that nongrammatical factors interact with grammatical ones. The transformational order appears to be respected for the reversible sentences but not for the nonreversible ones, where there is no clear difference between sentences depending on their syntactic form. Kernel sentences true with respect to the picture are easier to understand whereas negatives that are true with respect to the pictorial stimulus are more difficult to treat judging from response time and the errors committed.

Globally, the outcomes of this trend of empirical research on the model presented in *Syntactic structures* supply at best a weak support for the psychological reality of the transformational model; at best, because the results obtained could be explained with reference to other variables. For example, Mehler’s coding hypothesis (1963) can be replaced by an analysis in terms of sentence surface characteristics (e.g., passive sentences are more complicated to process in their surface form than corresponding actives).

Bever and Mehler (1967; reported in Fodor et al., 1974) took advantage of a peculiarity of the *Standard theory* (1965) stipulating that so-called stylistic transformations like adverb movement are not marked in base structure and therefore are not affected by the syntactic revisions that had taken place in generative theory since *Syntactic structures*. They utilized the same recall methodology as Mehler (1963) but with sentences containing “sentence adverbs” (i.e., adverbs attached to the symbol S in deep structure, as *probably* or *surprisingly*) and sentences containing verb-phrase adverbs (attached to the verb phrase in deep structure, as *slowly* or *carefully*). Both types of adverbs were presented in initial- and within-sentence positions in the learning phase. Results illustrated a tendency to locate adverbs in their deep structure position; sentence adverbs tending to be recalled in sentence initial position and verb-phrase adverbs in a position adjacent to the verb. However, the effect was found only in late trials, which fact the authors could not explain.

Blumenthal (1967) studied the memorization of sentences differing in deep structure but similar in surface structure. For example, the sentences: *Gloves were made by tailors*; and *Gloves were made by hand*. In the first sentence, tailor is the logical subject of the sentence whereas in the second sentence it is an adverbial modifier. Blumenthal had his participants memorizing the sentences and then gave them the last word of each sentence as a cue for recall. Results showed that the last word was more facilitative to the recall when it was the logical subject of the sentence. But was it really and only a question of deep versus surface structure? No experimental control was effectuated on alternative interpretation hypotheses.

 Clifton and Odom (1966) have verified whether the subjective distances in deep structure between sentence members of the IPN family corresponded to the suggestions of Katz and Postal (1964). Various experimental procedures were used. In one of them, as an example, the participants were first presented with a series of sentences that were different syntactic forms of the same kernel sentence. Then in a series of sentences transformed from the first ones, they had to recognize which sentences they had already seen albeit in another form. Results indicated that the participants’ responses corresponded relatively well to a three-dimensional space occupying positions akin to those predicted in Katz and Postal’s model.

Mehler and Carey (1967) examined the role of deep structure in sentence perception. They utilized sentences with a similar surface structure but a different deep structure (for example, sentences like: *They are easy to please; They are eager to please*) and sentences differing in deep and surface structures (e.g., *They are forecasting cyclones*; *They are conflicting desires*) A series of sentences with the same syntactic structure was presented first; then another (single) sentence with a differing syntactic structure was presented without informing the participants of the change. The idea was that the participants would perseverate with the receptive strategy used to decipher the first series of sentences. In conditions of perturbed perception (e.g., a noise produced experimentally in concomitance with the sentence to treat), this hypothesis was confirmed and it was shown that the loss of intelligibility was more important with the sentences differing also in surface structure.

Segui and Kail (1971) have conducted a corresponding research utilizing a production task. Participants were presented with a series of (grammatical) subject or object nominalizations. According to the theory in *Aspects* (1965), a nominalization is a transformation applied the deep structure of a declarative sentence whereby the verb becomes a noun while the subject or the object of that verb become noun complements (for example, in deep structure: *piece of* *art (artist fabricates it);* in surface structure, becomes: *the fabrication of a piece of art,* whereby the first noun becomes an object complement of the previous verb transformed into a noun). Following the presentation of the nominalizations, the participants were invited to complete a phrase introduced by a noun. Results showed that among the participants who had been exposed to a noun subject, a statistically significant part completed the phrases with a noun subject, and conversely for the participants who had been exposed to noun objects. It would seem therefore that according to the theoretical framework formulated in *Aspects*, the speakers are sensitive to deep structure differences between phrases with the same surface structure.

 Reading the preceding paragraphs, one might perhaps come to the conclusion that all in all the indications from transformational and generative grammar could be considered as having at least a part of psychological reality, in that they would correspond more or less to mental operations performed by language users when producing and understanding sentences. It should be clear, however, that these observations cannot be considered as supplying an unquestionable support for the generative theories.

 Several important problems can be mentioned that reduce the significance of the above trend of works (for more detailed criticisms, see, for example, Watt, 1970, and Costermans, 1980). Briefly mentioned, in the first experiments reported (Miller, 1962; Miller & McKean, 1964), the participants were openly requested to transform sentences. It is difficult to generalize from such particular situations to the spontaneous use of regular language speakers. Regarding other experiments (Mehler, 1963; Marshall, 1964; Savin & Perchonock , 1965), even if the data gathered suggest that sentences characterized in transformational grammar as more complex, may be more difficult to recognize, recall, or reconstruct, that does not mean that in the normal practice of language complex sentences are inevitably produced or interpreted through the mediation of simpler syntactic forms.

 Perhaps one could be willing to conclude minimally that what has been proven is that there exists some sort of particular sensitivity in language users to the structural properties of sentences.

 But even that is not guaranteed because there is a fundamental question that has to do with the generative theory itself. As in *Aspects* and following versions of the theory, the transformations operating on the phrase markers in deep structure are said to go hand in hand with semantic differences, it becomes virtually impossible to decide whether the similarities or dissimilarities identified by the participants in the experiments (e.g., in Clifton & Odom, 1966) must be attributed to syntax or to semantics.

 In another series of experiments (for example, Sachs, 1967; Begg, 1971; Honeck, 1971; Bransford, Barclay, & Franks, 1972; reviewed and analyzed by Costermans, 1980), where more attention was paid to the linguistic context of the sentences, the researchers tried to determine the respective importance of the syntactic and semantic dimensions in sentence treatment. The results suggested that actual recall may be neutral between the sentences heard and some of their logical consequences. The available data do not present a uniform picture, however, although the semantic variables seem often to play an important role .The subjects keep in short-term memory, only for a short period of time, the information on syntactic and lexical information. After a while, they appear to store sentences in longer-term memory mostly in semantic form taking into particular account logical and/or pragmatic consequences; but not at any particular level of grammatical description.

 Later attempts to assess the psychological reality of generative grammar have not met with more success.

 In relation with *Government and binding* (1981) theory, Hyams (1986, 1987) undertook to study the setting of the null-subject parameter in children’s grammatical development. She observed that the non-production of the grammatical subject seems to be a universal characteristic of early child language, even in cases where maternal language is non-pro-drop. Development therefore appears to begin with a default pro-drop parameter (+ morphosyntactic uniformity) accounting for the existence of utterances with null subjects in children’s speech (an indication considered to be arbitrary by other authors like Berwick, 1985, who believe that this is the non-pro-drop level that is “naturally” nonmarked).

According to Hyams (1986), development in English (and supposedly in other non-pro-drop languages) proceeds in the following way. Child observes the frequent presence of bare stems in language input and makes the hypothesis that her/his language allows the production of phonologically null subjects. The period of null-subject continues until the child notices the presence of verbal inflexions in some morphological contexts and their absence in others. On this basis, (s)he modifies the initial parameter. As a consequence, the grammatical subjects become obligatory and the auxiliaries are acquired.

Hyams (1986, 1987) provides developmental analyses from Italian and English. She reports that Italian-speaking children use the null-subject strategy whereas English-speaking children do not. The Italian children also produce verbal inflexions and auxiliaries early in development. And the end of the null-subject period for the English-speaking children is contemporaneous with the acquisition of verbal inflexions and auxiliaries. It could be coincidental, however.

French is considered a non-pro-drop language although it has a rich subsystem of verbal inflexions. These marks are not expressed phonologically in many cases, however (they appear in writing). One would expect to find a similar developmental trend for French as for English. Rizzi (1990) has documented the linguistic development of a monolingual French-speaking child between 22 and 26 months. His data do not corroborate the parametric hypothesis. At 26 months, 53% of the utterances are still without a grammatical subject. However, since 22 months of age, the child expressed, most of the time correctly, the perfect tense (involving the production of the auxiliaries *be* and *have*). Pierce (1987) has reported 40% of missing obligatory grammatical subjects in the productions of young French-speaking children together with the regular use of auxiliaries (e.g., *ai fait*; gloss: *have done*).

Weissenborn (1992) provides corresponding indications for French and German. Valian (1990) has studied the early syntactic development of 21 English-speaking and 5 Italian-speaking children. She reports that from the beginning the English-speaking children produce more grammatical subjects than their Italian-speaking peers (69% vs. 30%, respectively) in non-imitated utterances, and more pronouns in function of grammatical subject (75% vs. 35%).

Valian (1990) rejects Hyams’ theory and proposes that young English-speaking children simply learn that production of grammatical subject is required in English in all non-imperative sentences.

 A slightly different conception of the null-subject phenomenon has been proposed by Jaeggli and Safir (1987). They suggested that the structural property favoring the existence of grammatical subjects in sentences is morphological uniformity. Within a morphologically uniform system, one may find simple forms, like in Chinese, or complex forms like in Italian. In a non-uniform system such as English, complex forms (e.g., *He runs*, *I am running*) coexist with simple ones (e.g., *I run*, *you run*, *we run*, *they run*). The mixed systems tend to be non-pro-drop whereas the uniform systems are pro-drop.

 This suggestion is also problematic, however, as preceding examples from French indicate. Bouchard (1988; referring to Morin, 1985) suggested that the French constructions with *voici*/*voilà* can be considered as productions expressing a time indication at the indicative mode. These constructions do not lexicalize the grammatical subject and are the only cases in French where the grammatical subject may be omitted. The Chomskyan analysis stipulates that what determines the non-expression of the grammatical subject in a sentence is the relative morphological richness at the verbal level and the concord in persons and number between subject and verb. The case of *voici*/*voilà* in French appears to contradict this suggestion as they are grammatical elements that can be assimilated to verbs, allow the non-expression of the subject, and yet do not control any agreement between subject and verb.

Extrapolated to language development and use, the theoretical suggestions regarding the null-subject parameter lead to complicated solutions that do not seem to correspond to the available data. It would seem instead that defining linguistic expression more as an informational transaction can have more explanatory interest. In such a perspective, what has to be realized phonologically is what cannot be recovered from the linguistic context (morphological marks on the verbs in relationship with the expressed/unexpressed subject or the theme of the sentence), depending on the grammatical organization of the language, or from the extralinguistic context.

 More recently and more generally, Ambridge et al. (2014) have indicated that if parameter-setting approaches may provide computationally tractable accounts of syntax, they rely on the assumption that the learner knows word syntactic categories and grammatical roles in advance. The problem is that it is not clear how this knowledge can be obtained outside of innate prefigurations (see Chapter 7 for theoretical suggestions).

 **2.2. The modularity hypothesis**

 As indicated in Chapter 1, Chomsky’s notion of language modularity has two extensions. There is internal modularity in the sense that the major components of the language system are considered to be distinct, each one with its particular principles and rules. There is also modularity in a second sense (external grammatical), that is, the proviso that only conceptual aspects of language have deductive relations with the more general conceptual system of the mind. Chomsky (e.g., 1984) also proposes that the structure of mind is modular. He writes: “The human mind is just like other complex biological systems: it is composed of interacting sub-systems with their specific properties and character and with specific modes of interaction among the various parts” (p.16).

Fodor (1983, 1985) also defends a modular approach to the study of mind. It differs from that of Chomsky in several respects, however. According to Fodor, a basic functional taxonomy of psychological processes can be established that distinguishes between transducers, input systems, and central processors. The transducers are the sensory organs. The role of the input analyzers is to characterize “the arrangement of things in the world” (1983, p. 42). They are inference-performing systems that are “informationally encapsulated” or automata composed of “subroutines” serving special objectives. They are domain specific, fast in processing, and associated with fixed neural architectures. By contrast, cognitive processors are viewed as holistic or non-modular (“horizontal faculties”). Fodor lists language among the input analyzers, clearly giving prevalence to its grammatical organization.

 Fodor’s modularity conception has been criticized. Gardner (1985), for example, suggests that fully encapsulated modules in Fodor’s sense are “ideal structures” possibly observable early in development and later in special cases (e.g., some autistic children). In normal development, Gardner argues, encapsulation gradually dissolves because the highest human capabilities depend on the ability to integrate information from various sources. However, some degree of encapsulation may always be necessary in the input-output systems to account for their automatic-mandatory character and their procedural rapidity. Also the first developments (encapsulation then relative dissolution) may be followed in some cases by a secondary modularization phase and/or a functional automatization, which may come as a consequence of relatively large amount of practice and experience (also, Sternberg, 1985).

The matter is complex and it is likely that several types of organization might have to be distinguished; some having more to do with fixed neural architectures of the brain and others with learned modifications of some parts of this architecture (Caramazza, 1991). Paradis (2004) suggests that modules are isolable but interactive. However, their interactions do not modify the nature of their computational procedures. In other words, they may receive inputs and provide outputs but their internal structure is not influenced by that of other modules.

 Specific abilities, unrelated to intelligence, exist in cognitive handicap (mental deficiency or retardation as it used to be labelled; developmental intellectual disability as currently named), autism, and other pathological entities. The cases of a small number of language exceptional (“hyperlinguistic”) persons with a marked cognitive handicap have been documented in the literature. They are briefly presented and discussed below. The terms “exceptional” or “hyperlinguistic” must be understood as referring to ability in some aspects of language that is particularly advanced for people with an important cognitive handicap. Outside the area of language, exceptional capacities in various domains (music, drawing, perception and recognition memory for shapes, calendrical calculation and numerical ability) have been reported (see, for example, O’Connor & Hermelin, 1988, 1989; for partial syntheses).

 As to language, the most striking cases reported involved the study of Paul, a person with standard trisomy 21, by Seagoe (1965); Christopher, by O’Connor and Hermelin (1991), successively pursued and extended by Smith and Tsimply (1995); several cases studied by Curtiss (1982, 1988); the study of Laura by Yamada (1990); FF, an adolescent with Down syndrome studied by Vallar and Papagno (1993) and Papagno and Vallar, 2001); and the case study and follow up of Françoise, another person with standard trisomy 21, that I had the privilege of documenting (Rondal, 1994a, 1995).

 These cases are summarized and discussed in Rondal (2003). They demonstrate that advanced morphosyntactic development (in some cases, near normal productive and receptive functioning) is possible in spite of marked intellectual limitation, thus raising a difficult question for the theories that make morphosyntactic development narrowly dependent on cognitive development. Strong cognition hypotheses have been proposed by Bever (1970), Slobin (1973), Sinclair (1971; in the Piagetian tradition of Piaget, 1963), Bates and MacWhinney (1987), and more recently by Tomasello (2003).

The question is how to interpret correctly the above observations.

 A first consideration is that advanced cognitive development is not necessary for advanced syntactic development. That could have been deducted already from typical language development. Non-retarded children acquire most of the grammar of their maternal language during their first four or five years[[18]](#footnote-18), i.e., at times at which they are still quite a way from having completed cognitive development. This means that normal grammatical functioning is perfectly possible with an intellectual development akin to 4-5 years mental age, which is the mental age reached by the language exceptional subjects in the studies mentioned. This is also the average mental age reached by “standard linguistic” persons with moderate or severe cognitive handicap. So, why is it that these standard persons do not develop a grammatical ability of the same type as the one documented in the language exceptional persons?

 In my 1995 opus, I valued the hypothesis that the language-exceptional persons with a marked cognitive handicap have preserved morphosyntatic abilities, for yet unknown genetic reasons (individual variation of some kind), allowing them to make the correct categorizations of grammatical elements in their speech. That was also the bulk of Smith and Tsimply’s (1995) interpretation of Christopher’s linguistic talents. I now believe such a hypothesis to be of a low degree of plausibility. Representational innateness can be and has been challenged on logical, mathematical, and neurobiological grounds (see next chapter). It is becoming clearer that linguistic representations are constructed on the basis of people’s experience with language.

 More likely, the language-exceptional individuals with cognitive handicap have preserved brain macrostructures devoted to language treatment (Rondal, 2003), i.e., cytoarchitecture, transmitter types, number of layers of neurons, packing density, basal cortical circuitry, and connections between brain regions; Elman et al;, 1997; Gazzaniga, 2008). This refers to the innate structuring of the brain information processing system devoted to the acquisition and use of language representations.

 As will be elaborated further in the following chapter, the difference between a conception of the above type and representational innateness does not concern the belief that humans are biologically prepared for language. They certainly are. It is in the idea that linguistic knowledge not being given beforehand is constructed with the concourse of particular brain structures following regular epigenetic sequences.

 Morphosyntactic development (as will be argued in Chapter 7) may be thought to depend largely on implicit learning. As is known, implicit learning and procedural (implicit) memory are supported by particular neurological structures diverse from those involved in explicit learning and memory. Regarding the exceptional cases of morphosyntactic development in persons with a marked cognitive handicap, I would add that their particular abilities also depend on a (at least partially) preserved working memory capability (a modern form of short-term memory). Baddeley’s model of working memory (Baddeley, 1990; for a more recent version, see Baddeley, 2012) attributes an important role to attention. Recent theoretical accounts of implicit learning also give a particular place to attentional focus in the process of learning from the input.

The suggestion regarding the exceptional cases reviewed above is that they may also have been relatively gifted in this respect in comparison with standard persons with a cognitive handicap. Françoise (Rondal, 1995) had an auditory-vocal working memory span of 4 items (digits, words, nonwords). She had near-normal auditory-vocal working memory processes (attested according to the assessment techniques explained in Baddeley, 1990), and she was able to use spontaneously rehearsal strategies relying on semi-private speech. Curtiss’ Anthony and Rick had an auditory-vocal working memory span at the 6-7 years old level. Vallar and Papagno’s FF had a corresponding average span of 4.75 units. This is below normal functioning but markedly better than what is currently observed in standard persons with a marked cognitive handicap. Françoise’s mean length of utterance - MLU - (calculated in number of words + obligatory grammatical morphemes divided by the number of utterances) was 12.34. Sentence span was 14 words. At times, she could repeat correctly sentences containing up to 20 words. This is near-normal functioning (Miller & Selfridge, 1950; Craik & Massani, 1969).

A final note to acknowledge sadly that Françoise developed Alzheimer disease and died at 48 years. We could reassess the state of her language and cognitive abilities when she was 47 years, that is, 11-12 years after the first study (Rondal et al., 2003). The Day center that she frequented had informed us that her state of mind had deteriorated and asked if we could reassess her abilities. Her language capacity, as we found at the time of reassessment, had considerably lowered with respect to her previous performance. Phoneme articulation remained correct in spontaneous speech but there were large number of dysfluencies; something that never happened before. The rhythm of speech had lowered from 3.3 words per second previously to one word per second. MLU had dropped to 6.91 and she appeared to have lost almost entirely her previous ability to produce and understand compound and complex sentences. Her episodic memory and visual-spatial processes had also deteriorated markedly, as well as her auditory-vocal working memory span (down to 2 items).

 CHAPTER THREE

 GRAMMATICAL REPRESENTATION IN LANGUAGE USERS

Assessing the psychological reality of the formal and functional categories, hierarchies, rules, and derivational devices proposed in linguistic theories, is not an easy endeavor. Some linguists suggest that these notions are simple descriptive tools, hypotheses, or idealizations that have proved useful in theoretical accounts of grammar. They should not to be taken as precise models of the speakers-hearers functioning. Numerous publications in linguistics, particularly in generative linguistics, are ambiguous in this respect, however. No matter of what linguists may think, the fact is that many psycholinguists have come to believe that flesh and bone people process morphosyntax according to the indications in linguistic theories. There are several ways to approach the psychological reality of linguistic notions regarding the morphosyntax of language. They will be analyzed in what follows;

 **3.1. Grammatical acceptability judgements**

 Grammaticality judgements have been widely used with the objective to approach the conscious knowledge that language users may have of the morphosyntax of their language. One should speak of judgement of grammatical acceptability (Pinker, 2007) or well-formedness. As stressed by Chomsky (1957), acceptability is about the performance of speakers. Acceptability does not mean grammaticality. While an acceptable sentence must be grammatical, a grammatical sentence does not need to be acceptable. For a sentence to be judged acceptable, it must also appear natural and appropriate to a given context, be easily understood, and be conventionalized to a certain extent (Chomsky, 1988). For example, a string like *Colorless green ideas sleep furiously* while being grammatical is not acceptable. We are sent back to the distinction between competence and performance.

 For some linguists, grammaticality is categorical or binary. A string in a given language is grammatical or not grammatical. For others, grammaticality must be conceptualized as a scale. There are grammatical strings, strings that clearly are not grammatical, and an all manner of ranges of partial acceptability in between; other variables but grammaticality playing a role.

 Chomsky (1965) has proposed a hierarchy of formal rules that can be utilized to attribute a degree of grammaticality to sentences varying in grammatical correctness. These rules are not of the same kind as normative rules (defined in normative grammars, that is, grammars based on norms or statistical tendencies of grammatical use in a given language community). According to Chomsky (1965), three types of grammatical violations can be distinguished corresponding to three levels of rules in grammatical theory. At a first level, there are the violations of categorization rules. These rules are related to the correct choice of lexical items according to their grammatical function in the sentence (e.g., *Sincerity can virtue a child*\* - where a noun erroneously replaces a verb - vs. *Sincerity can frighten a child*). A second level concerns the rules of sub-categorization. They concern the obligatory relation between lexical subcategories and grammatical functions (for instance, the incorrect use of an intransitive verb instead of a transitive one (e.g., *John smiled Joshua to leave the room*\* vs. *John forced Joshua to leave the room*). A third category of grammatical violations is related to the selection rules. Violations of that kind lead to the erroneous selection of certain lexical items with regard to the semantic constraints induced by the presence of other lexical items in the sentence (e.g., *Wine loves James*\* instead of *James loves wine*)[[19]](#footnote-19).

 Moore (1972) has tested the hierarchical hypothesis regarding degrees of grammaticality in sentences. On a grammatical acceptability scale, participants’ assessment did not correspond to Chomsky’s suggestions. For example, English sentences with incorrect SVO (subject-verb-object) sequence (related to the third case of grammatical violation in the above scheme), were judged less acceptable than sentences with other types of grammatical violation.

 A series of other studies have been devoted to testing acceptability judgement of strings of words and sentences (e.g., Bolinger, 1968; Mittins et al., 1970; Greenbaum, 1973, 1976; Greenbaum & Quirk, 1970; see Fetzer, 2004, for a review). What is judged to be grammatical in a large majority of cases reflects more the fact that a given sentence has a current meaning, is semantically and pragmatically plausible, socially acceptable, or even stylistically appropriate, than the conformity to the prescriptions of a grammar of the language (Lakoff, 1971b; Fetzer, 2004). Lakoff (interviewed in Huck & Goldsmith, 1995) already said: “In 1969, there was a virtual explosion in pragmatics research. It became clear that grammatical well-formedness was relative to context …” (p.112). More intriguingly, Nagata (1988) has observed that the acceptability judgements are influenced by frequency of use. A string that may sound unacceptable at first heard or seen may be judged acceptable if it is presented several times in various contexts. Tomasello (2008) argues that the phenomenon of grammaticality is “…actually just another instantiation of social norms for everyday behaviors… - but reinforced by the fact that commonplace grammatical utterances are heard dozens or even hundreds of times every day so that their pattern is quite entrenched in our communicative activities” (p. 291). Accordingly, less frequent sequential patterns do not sound as unacceptable when violated as more frequently heard ones (Brooks et al. (1999).

 Several works have tried to approach syntactic consciousness in children (for example, Gleitman, Gleitman, & Shipley, 1972; De Villiers & De Villiers, 1972, 1974; Moore, 1975; Gowie & Powers, 1977; Valian & Stojak Caplan, 1979; see Hakes, 1980; Bredart & Rondal, 1982; Tunmer & Grieve, 1984; and Gombert, 1990; for reviews and analyses).

 In Gleitman et al. (1972) study, at 30 months of age, roughly 50% of the incorrectly ordered imperative sentences (e.g., *Box the open*\* instead of *Open the box*) are judged acceptable. Only 15% of the wrong sentences are corrected for word order. In 85% of these cases, children’suggestions for correction concern a change in meaning (for example, *Get in the box* instead of *Box the open*\*). De Villiers and De Villiers (1972, have reported data similar to those of Gleitman et al. (1972) with children of the same ages (around three years). With 4 year- olds (De Villiers and De Villiers (1974), there was a greater sensitivity to word order in active and reversible passive imperatives. However, undue suggestions for semantic change were still numerous.

 Gowie and Powers (1977) have reported data showing that children and adolescents in great majority reject grammatical but semantically abnormal sentences as not acceptable; failing therefore to exhibit a clear standard of grammatical acceptability.

 Valian and Stojak Caplan (1978) have proposed children aged from 6 to 10 years to repeat grammatical sentences read by a first experimenter. Half of the sentences were “clear” in the sense of being easier to understand (for example, *The games that he bought were fun*). The other half was slightly less easy to understand (for example,*The games he bought were fun*; with non-expression of the relative pronoun) while being semantically equivalent. Following the presentation of each sentence, a second experimenter, sat at the other end of the room, said: *What*? In reply to this request for clarification, most “clear” sentences were repeated verbatim or close to. Still only 25 % of the “less easy” sentences were modified at 10 years of age; the modifications consisting in half of the cases in inserting the relative pronoun at the right place in the sentence and in the other half in various irrelevant and sometimes erroneous changes (semantic, syntactic, grammatical morphological) made to the original sentence.

 In a systematic study with pre-adolescent children, aged around 12 years, Gaux and Gombert (1999) compared several techniques for eliciting acceptability judgements: repetition of grammatical and ungrammatical sentences (assuming that subjects spontaneously tend to normalize strings of words that do not correspond to the grammar of their language); acceptability judgement; explicitly requested correction of ungrammatical sentences; localization (locate the grammatical incorrectness in an ungrammatical sentence); completion of a morphosyntactically incomplete sentence; and an original and more complicated technique, labelled replication. It consisted in asking the participants to reproduce in each one of three grammatically correct sentences a grammatical incorrectness figuring in a first sentence. In the replication condition, sentences were presented both orally and in written form in a plausible attempt to reduce the load on short-term memory. The ungrammaticality of the sentences concerned word order within phrases, phrase order in clauses, as well as morphosyntactic aspects such as grammatical concord between subject and verb.

Participants’’ performance proved excellent with average scores close to the ceiling level: 99, 10% of correct responses for the judgements; 96, 84 for the corrections; 93, 37% for the localizations; and 92, 32 for the repetitions. Expectedly lower scores were obtained with the replication technique (77%).

The question is to identify the determinants of the participants’ responses.

 As indicated by the authors, the data suggests that the tasks proposed were effectuated by the participants without reference to normative grammatical rules. They judged the linguistic material proposed and acted on it as requested, from what appears to be an implicit analysis of the semantic content and not from a reflection on the syntactic nature of the sentence components. Gaux and Gombert (1999) concluded that given the important number of variables involved in grammaticality inquiries of the kind, it is extremely difficult and perhaps impossible to determine with accuracy which processes the subjects actually use in their judgements, given that other processes but intentional reflection on the morphosyntactic form itself, allow to comply successfully with the task most of the time.

 In conclusion, this trend of research on grammatical intuition fails to yield clear information on the possible formal knowledge of children, adolescents, and adults, regarding morphosyntactic aspects of language. What comes to the foreground in most cases is the important role of semantic and pragmatic variables in matters of grammatical acceptability.

 **3.2. People’s grammatical knowledge**

 Regular language users while perfectly able to understand and produce language, have little to no conscious knowledge of the linguistic categories and rules that most psycholinguists consider necessary for morphosyntactic treatment. This observation was already made times ago. The German philosopher Immanuel Kant (18th century) wrote: “One speaks without even knowing the grammar; and he who speaks without knowing it actually has a grammar and speaks following rules of which he is not conscious” (1800, p.9; my translation). Kant who commented little on language in his work, did not specify what he meant by grammar and by rule. But he may probably be taken to signify that grammatical knowledge in people, whatever it may be, is implicit.

 In psycholinguistics, there is a curious paucity of inquiries regarding conscious grammatical knowledge in language users.

 Paradis (2000) has documented the case of Michelle, a French-speaking adult with normal intelligence but little school instruction. She is totally unable to segment sentences into nouns, verbs, article, prepositions, and other formal classes while speaking perfect French. Paradis (2004) acknowledges that formal linguistic notions can emerge from explicit learning and metalinguistic reflection. He argues, however, that they represent a different register foreign to ordinary language use. I completely agree with Paradis’ indication. Metalinguistic knowledge may certainly help talking about language but there is no clear indication that it is instrumental in determining actual language treatment.

 Primary and secondary school programs in all the countries that I know seem to share the opposite conviction that metagrammatical training is a necessary condition for correct language functioning. But there is no proof of that. It even runs contrary to the obvious fact that children and adolescents at school ages already implicitly know the morphosyntax of their maternal language. School is in the paradoxical position of constraining pupils to learn explicitly notions that they already know implicitly. The results, it seems, are not terribly encouraging as many school teachers would tell you. It is common knowledge, additionally, that the huge majority of standard language users quickly forget (assuming they ever really knew it) the limited and artificial arsenal of metalinguistic notions painfully learned, once they have left school.

But even during school years, things are not that clear either. In her Master’s thesis, Thewis (1991) observed that when primary school children from 4th and 6th grades are interrogated informally, they mostly provide semantic definitions of the grammatical categories that they are supposed to have learned in class. Nouns are names of objects and persons; verbs correspond to actions, states, or events; modifiers add a meaning indication to a word. Grammatical subjects are animate entities responsible for the actions expressed by the verbs. Direct objects are directly affected by the action expressed by the verbs, etc. Sentences are defined as the expression of a complete idea.

 I have casually interrogated dozens of native adult speakers with no particular training in language sciences in several countries and languages, and obtained basically the same kind of responses as those reported by Thewis (1991) for primary school children. Standard language users have no conscious knowledge of the rules, hierarchies, formal and functional categories defined in linguistics and postulated by many psycholinguists to be necessary for normal language functioning. What they appear to be aware of in terms of morphosyntax are mostly the normative preferences regarding word order (even in languages relatively flexible in this respect) and the correct ways of marking nouns, pronouns, verbs, and adjectives (the case being) for expressing number, gender (the case being), tense, and mood, and concord between various elements in the sentence[[20]](#footnote-20).

 In another way, the lack of conscious knowledge of morphosyntax and the related lack of explanatory power of the linguists’ intuition is demonstrated by the failure of linguistics to come up with a consensual theory of morphosyntax (cf. the many reconceptualizations of generative linguistics as illustrated in Chapter 1 to consider only this variety of linguistic theorizing). This raises the delicate question of the epistemological status of linguistic theory. As long as the question is merely to describe language and morphosyntactic facts as they appear in langage use, there is nothing to object. The caveat is in the potentially infinite number of descriptions (logical, well-made, descriptively adequate, correctly covering the data, etc.), with no way of objectively selecting between them the “correct” one, as in this sense there is no theory that would be more correct than any other one. This of course should not be taken to mean that linguistic theory has nothing relevant to propose to psycholinguistic research. The entire contrary. Many notions emerging from linguistic descriptions may be taken as departure hypotheses for psychological investigation. However, there is a huge gap between an “open hypothesis and suggestion” and a dogmatic implication bypassing psychological research in the name of a grammatical competence declared to be immune to an empirical approach.

The ubiquitous notion of rule (or principle, or parameter) provides another case in point. It is agreed that the rules are unknown to ordinary speakers-hearers. They are considered as belonging to an innate language faculty stored in the brain. It is relevant to note that in so doing one replaces the normative rules of traditional grammars with “true” rules the specification of which is the object of theoretical linguistics. It is no longer a matter of “artisanal” (Milner, 1989, p.251) conventions of representation or normative rules for fabricating sentences (mixing syntactic and semantic justifications with as many exceptions as regular cases). Rather, it is a fundamental hypothesis on the nature of language.

The problem differs, however, when taking for granted that bone and flesh language users do rely on rules of which they have no conscious knowledge, to actually produce and understand sentences. Rules unknown to language users are instances of *petitio principis*. They are metaphorical at best, plainly irrelevant at worse. Language philosophers (Baker & Hacker, 1984; Milner, 1989) have warned against that way of theorizing in language psychology. According to Kripke (1982), the necessary explicit character of a rule is also mentioned in Wittgenstein’s *Philosophical investigations* (1961)[[21]](#footnote-21).

 Searle (1975) indicates that the concept of rule must be tied to a particular knowledge. Speaking of rule must imply a mental operation that “knows itself” as activity and that knows the rule: its existence and content (p. 65). This “common notion” of rule that fits most of the ordinary usages, seems to be alien to generative linguistics and psycholinguistics.

 Searle (1975) warns against a possible confusion: it can always be that a goal-oriented behavior gives the impression that a rule is followed when actually no such rule - at least in the common sense - is applied. From the fact that human actions are regulated, one cannot, and one should not, conclude that they follow particular rules (p.65). Paradis (2004) adds: “Speakers may behave ‘as though’ they have internalized a particular explicit rule, but this only means that, through practice, they have internalized computational procedures that allow them to produce (and understand) sentences than *can* (stress Paradis) be described in terms of a particular explicit rule” (p.33-34).

 In agreement with the above statements, I suggest that the word “rule” be avoided when dealing with the psychological organization of morphosyntax. One should prefer the expression “regulation” (which refers to a mental process), even if the two words, rule and regulation, share the same etymology (from the latin *regula*, for *set square*).

As Milner (1989) questions: what is in the final analysis the justification for claiming that speakers-hearers customarily rely on rules to organize their language production and understanding? None, as it would seem. The same question may be asked regarding the notion of linguistic competence. What is a linguistic knowledge of which the language holder is not aware? Knowing here is not knowing that one knows. But then is it really knowledge?

 **3.4. The innateness issue**

Many linguists and psycholinguists believe that morphosyntactic knowledge, as part of a human language faculty, is based on specific genetic predispositions (for a recent reaffirmation, see Yang, 2016).

Pinker (1994) has advanced the hypothesis that human language is an instinct (an idea whose paternity he attributes to Darwin, 1871, in his book *The descent of man*) and this is the title of Pinker’s book. This indication has motivated a strong rebuttal from Tomasello (1995; also Evans, 2014). Actually, Pinker’s suggestion is correct. Instinct, at least in taxonomically higher species (but already in birds), involves an innate biological mechanism, already mature congenitally, and the concomitant learning of a particular category of objects on which the mechanism locks up. A simple example is imprinting in various species of birds following birth, as long studied in biology (Lorenz, 1937; Tinbergen, 1951; for classical investigations). The little bird (for example, a duck), just out of the nest, has a few hours (defined as the imprinting critical period) for recognizing the first object moving in the immediate environment, start following it, and imprint on it, which commits the animal’s entire future social and sexual life.

 Obviously, human language is an instinct in the sense that it is underlain by particular organic structures. Pinker’s (1994) book title therefore is correct. It is even trivial. What is less trivial is the suggestion that the object of the language instinct (i.e, the language code itself based on universal grammar) is also given genetically (that is the core of the representational innateness hypothesis, which will be discussed shortly). This does not correspond to the biological notion of instinct whereby the object has to be learned.

 Evans (2014) equates language instinct (and faculty of language) with universal grammar. On this basis, she rejects Pinker’s theory arguing that it is not based on actual findings. Evans calls it “the language myth” (which is the title of her book). But the equation language instinct-universal grammar is not a necessary one. Language as a general instinctive organization is grounded in particular brain mechanisms that are innately given but probably not specific to language (an issue still very much discussed; see next section).

 Pinker (1994, 1999), nevertheless, presents arguments in favor of the language instinct hypothesis. He echoes Bickerton’s (1981, 1984) indication regarding the human ability to create language in certain circumstances starting from a pidgin. Based on historical reconstruction, Bickerton argued that children separated from their parents and exposed to a pidgin at the ages at which they should normally acquire their mother tongue, will inject grammatical complexity in a pidgin and transform it into a creole. Bickerton’s suggestion is controversial, however. He may have underestimated the role of existing linguistic substrates (for example, colonial French, English, or Spanish, and the original African languages, in the case of the West Indies; Muysken, 1988; Youssef, 1988).

However, cases of genuine morphosyntactic invention have been documented in language impoverished social contexts (Goldin-Meadow, 1984, 2005). Reports of deaf children and adolescents illustrate the development of new gestural systems with no particular external influence, only the need or the willingness to establish one’s own language. In Nicaragua, deaf students started to use signs and gestures during school recess and after school, alternatively to the imposed oral training and lip-reading in Spanish. Their invention was taken up by successive classes of pupils. It is reported that what became Nicaraguan Sign Language is now the preferred idiom of deaf people in Nicaragua (Kegl, Senghas, & Coppola, 1999; Senghas, Kitas, & Özyûrek., 2004; Senghas, 2005). Corresponding observations have been made with deaf people from the Al-Sayyed tribe in the southern Neguev desert of Israel. Their particular sign language is approximately 70 years old now (Goldin-Meadow & Mylander, 1998; Sandler et al., 2005; Hopkin, 2005; Senghas, 2005).

 In these reports, it is argued that core aspects of morphosyntax have been invented. Pinker (1994) suggests that this is a manifestation of the power of the language instinct equipped with basic dispositions of universal grammar.

Pinker (1994) also evokes a language pathology known as specific language impairment (SLI) in which children of normal intelligence present an important deficiency in morphosyntactic development (see Chapter 7). SLI may be hereditary. A pedigree study has pointed to the pathological effect of a single dominant gene (Gopnik & Crago, 1991). Pinker (1994) argues that this gene belongs to the substrate of the language instinct and that its mutation is responsible for a pathological modification in the genetic basis of grammar (an interpretation maintained in Van der Lely & Pinker, 2014). This is unlikely as will be seen below.

In opposite way, Pinker (1994) suggested that people with a rare genetic disorder, labelled Williams’ syndrome, illustrate the preservation of that same genetic basis in spite of moderate intellectual impairment. Williams’ syndrome is due to the absence of at least 17 genes on one of the two chromosomes 7 (Korenberg et al., 2000). On the basis of observations reported by Bellugi et al. (1988), Pinker argued that language development in Williams’ syndrome is close to normal and that this has to be attributed to a preservation of the grammatical information in the genetic equipment of these persons. However, further studies have shown that language development in children and adolescents with William’s syndrome is not normal even if it can be considered more favorable than what is usually observed in corresponding genetic syndromes of cognitive handicap. Important limitations and deficiencies have been documented in morphosyntactic development in this syndrome (Thomas, et al., 2001; Grant, Valian, & Karmiloff-Smith, 2002; Karmiloff-Smith et al., 2003; Mervis et al., 2003). The speech of these people is often of good quality and lexical development may be relatively preserved (with important interindividual differences, however; Pezzini et al., 1999), which may easily give the impression of a better language functioning.

The hypothesis of representational innateness is that basic grammatical notions corresponding to universal grammar do not need to be learned as they are genetically coded as a property of species. The move from genes to specialized brain areas is viewed as instantaneous, being the effect of a releasing mechanism, as soon as a contact has been established with a particular language (Pinker, 1985, 1994). This implies a continuity hypothesis in the sense that basic linguistic representations are available to the learner in continuity from birth on even if performance factors prevent them to be expressed in early ages. Some authors, however, for example Borer and Wexler (1987), think on the contrary that the genetic syntactic propensity undergoes a maturational process. It takes a few years to become fully operational.

The representational innateness hypothesis is not a clear-cut concept. It has been formulated in differing ways. For Chomsky, as said, it implies a universal grammar that may be defined parametrically. Other authors have concentrated on particular grammatical aspects. Bickerton (1984) indicates that the universal abstract notions of sentence, noun, verb, determiner, rewriting and movement rules, have to be prefigured genetically. Pinker (1994) argues that at least three properties present in all languages have to be represented in the innate device: (1) the principles directing the displacement of elements in sentences; (2) the abstract prefiguration of grammatical morphemes related to the expression of time, aspect, case, mood, and positive and negative polarity in clauses; and (3) the abstract prefiguration of the formal categories of noun and verb. Radford (1990) suggests that the whole category of formal notions is innately given whereas the notions pertaining to the functional categories can be learned.

 In a more recent paper, Fitch and Martins (2014 have contributed a varying version of the representational innateness hypothesis under the fancy name of dendrophilia. They argue that the human species is genetically endowed with a particular sensitivity for temporally ordered hierarchical structures of the kind represented in tree-like diagrams of sentences. The human aptitude for syntax is based on that propensity. Importantly, they abandon the idea of the specificity of syntax in cognitive organization. The same aptitude is also at work in music and planning motor acts beyond simple reflex actions. Lashley (1951) whom Fitch and Martins say to have borrowed their idea from, already suggested that the human aptitude for treating temporally ordered hierarchies of which he thought that language provided a good example, originated in the basic organization of complex motor acts.

 Fitch and Martins (2014) argue that the major role of the hierarchical dimension of language is to supply a buffer mechanism for momentarily storing morphosyntactic information while some other operation(s) is (are) performed on the string of words. According to the same authors, the premotor cerebral cortex plays an important role in the hierarchical aptitude as demonstrated by anatomical, neurophysiological, and neuroimaging data. They call it “supra-regularity, » meaning that it allows going beyond “regularity” or the simple linear treatment of stimuli.

 Fitch and Martins’s (2014) hypothesis relies on the same grammatical notions (within hierarchical devices) as those postulated by the representational innateness hypothesis.

 In a recent essay, Hauser and Watumull (2017) also appear to give up the notion of language specificity (or rather uniqueness). They argue that language, mathematics, music, and even morality, all implement generative computational and abstract representations deriving from a general (i.e., contentless) computational system, which they label “universal generative faculty.” This nonspecific system interfaces with the different domains of knowledge. Domain specificity is preserved, however, in a kind of secondary way, that is, restricted to the primitive representations in each particular domain of knowledge and their interfaces with the universal generative faculty.

Whatever its exact definition and the extent of its specificity, grammatical representational innateness is a strong claim. Its genetic underpinning should be carefully scrutinized.

 The fact is that in spite of impressive progresses in molecular genetics over the last decades, no gene coding for specific grammatical representations has been identified. Elman et al. (1997) are convinced that the human genome does not have the resources for coding the enormous number of decisions (binary, for example) necessary for representing a universal grammar. Protein-coding genes in humans number a little less than 2.2 x 104. Only 20 to 30% of these genes play a role in the construction of the nervous system (Willis, 1991). It could perhaps be counter argued that non-coding DNA that constitutes 98, 70% of DNA (deoxyribonucleic acid, the substance of the genes) could supply the necessary biological substrate for the innateness hypothesis. The function of non-coding DNA, however, is one of regulating gene expression. It is difficult to imagine how they could encode grammatical information.

 Two “language” genes have been identified. One is called FOXP2. It is an autosomal dominant gene located on chromosome 7. FOXP2 belongs to a family of genes coding the synthesis of proteins that have a so-called forkhead-box (FOX) domain (referring to how they the bind to specific areas of DNA). FOX proteins are a type of transcription factor regulating the expression level of a series of genes important in early development (Lai et al., 2001; Vargha-Khadem et al., 2005). A mutation of FOXP2 causes an important volumetric reduction of the white matter of the striatum (one of the nuclei of the basal ganglia of the forebrain). The striatum is connected to several cerebral cortical and neocerebellar regions involved in the sequential treatment of information.

Abnormal morphological development of several regions of the left brain (the Broca area, the perisylvian area, the basal ganglia nuclei putamen and pallidum) has been observed in the familial cases studied. Clinically, the FOXP2 mutation is associated with verbal dyspraxic, dysphasic, and dysgraphic disorders where difficulties in regulating combinatory expression in articulatory, semantic, and syntactic tasks, and in procedural learning, are dominant. These difficulties are associated with an underactivation of the above brain areas as observed in neuroimaging studies (Watkins et al., 2002; Liégeois et al., 2003). FOXP2 is the gene that was hypothesized in the study of Gopnik and Crago (1991) mentioned previously.

 FOXP1 is another member of the FOXP gene family of transcriptions factors involved in the development of the central nervous system (Sin, Li, & Crawford, 2015). It is also expressed at the level of the basal ganglia of the brain in songbirds (Teramitsu et al., 2004). These birds use a specialized subcortical-cortical circuit for perceiving, learning, and producing the song characteristic of the species (Brenowitz & Beecher, 2005). A lesion in this neural device hinders the acquisition of the song in disturbing the ability to perceive and produce the typical sequences of musical notes. Given the colocalization of FOXP1 and FOXP2 found in several structures of the bird and human brains, Teramitsu et al. (2004) suggested that mutations in FOXP1 can also be related to speech disorders.

 The communicative nature of FOXP2 is further documented in a series of studies in genetic engineering reported by Päâbo (2014). The protein coded by FOXP2 is present in a broad range of mammals. In human, it differs from the same protein in apes and almost all mammals in the position of two amino acids. Pääbo and colleagues introduced these two changes in the genome of mice. At two weeks of age, the pups born to these mice demonstrated noticeable differences in vocal communication with the control group; differences compatible with the hypothesis of a significant role of the protein coded by FOXP2 in vocal communication. Further work in Pääbo’s laboratory showed that this protein is involved in the process of neuronal outgrowth (a key factor in communication and signaling between neurons) in brain areas devoted to motor learning.

A second gene identified as “language” gene, called CNTNAP2, is also located on chromosome 7 (Vernes et al., 2008). This gene codes the synthesis of a protein of the neurexine family involved in neuron communication and the interconnexion of cerebral areas. It is particularly expressed in the frontal and temporal lobes. A mutation of CNTNAP2 is associated with a set of disorders including the autism spectrum, the syndrome of Gilles de la Tourette (characterized by involuntary motor - including vocal - movements), and delays and disorders of speech and language development.

 Numerous studies of twin, adoption, and genetic linkage concerning language have been conducted (Stromsworld, 2001, for a review). Genetic linkage refers to the fact that two alleles of two different genes closely located tend to be hereditarily transmitted together. Clinical works on genetic linkage in language have compared the genotype of persons affected with a language disorder with those of parents, also affected with the same disorder or not, as a way of establishing whether the same genes are associated with the same pathologies across generations and whether they occupy the same place (locus) on the same chromosome(s). Results suggest that familial dysphasia and dyslexia are genetically heterogeneous, corresponding to two loci on chromosomes 6 (at q31) and 7 (at p21.3). These loci are related to other syndromes affecting speech, language, and communication, like autism, Gilles de la Tourette syndrome, schizophrenia, and the syndrome of attention-deficit hyperactivity-disorder.

Quite clearly, there is a genetic basis for language. However, no demonstration of the existence of particular genes or family of genes coding for the grammatical notions postulated by the representational innateness hypothesis, has been made. This hypothesis cannot be completely excluded (proving the non-existence of an entity or a phenomenon is logically impossible in inductive sciences), but it looks quite unlikely.

Instead the genetic substrate evidenced seems to be primarily related to brain mechanisms involved in the organization of sequences, motor or otherwise. These mechanisms involve cerebral structures and processes devoted but not specific to language. Major devoted language territories in the brain involve the left-perisylvian areas, (Broca area in the prefrontal region for language production; the left-posterior temporal region for language comprehension) and the right neocerebellum overall for sequential treatment (Fabbro, 1999; Paquier & Mariën, 2005).

Gazzaniga (2008) indicates that the left-cerebral hemisphere differs from the right one in its microscopic architecture. The cerebral cortex has six layers of neurons stacked on top of each other and lined up with those in the layer above and below. This forms minicolumns of pyramidal cells that cross the layers perpendicularly. It is accepted that the neuronal columns are the fundamental processing units within the cerebral cortex and that assemblies of columns create complex functional circuits. The minicolumns are wider and the space between them greater on the left-hand side of the brain, which suggests a smaller number of minicolumns interconnected than in the right hemisphere. This could indicate that the local processing architecture in the left hemisphere is more elaborate and less redundant. There is also an augmented number of extra-large pyramidal cells in the most superficial (supragranular) layers of the anterior and posterior (language) areas of the left cortex in comparison with the corresponding areas of the right cortex. This is an additional indication of connectional asymmetries between the two hemispheres. They may play an important role in the temporal processing of stimuli. It is interesting to add, in this respect, that Plante (1991), in a study using a technique of magnetic resonance imaging (see Section 3.5 below) with members of four families having a child with specific language impairment and who themselves had presented language difficulties in childhood, has documented an abnormal volumetric relationship between the two cerebral hemispheres in the four children as well as in their brothers and sisters. The perisylvian volumes were equivalent on the left and right sides of the brain or more important on the right side, in opposition to the typical anatomical situation whereby these volumes are more important on the left side of the brain (Wada, Clarke, & Hamm, 1975), already from the fetal stage of brain development (Chi, Doaling, & Gilles, 1977).

 These language-devoted brain areas are functional early in life. Dehaene-Lambertz, Dehaene, and Hertz-Pannier (2002) have observed that the perceptive abilities of neonates for speech sounds are lateralized on the left-hand side of the brain and that they rely on the same language territories as documented in adult subjects. Dehaene-Lambertz (2000) has also shown that two-month old babies already present a mismatch negativity wave when exposed to a changing phonetic sequence such as *babagaga*. This is a wave originating in the frontal and temporal regions of the brain and associated with the detection of a change in a sequence of acoustic stimuli.

 Kimura and Watson (1989) have suggested that the left-cerebral hemisphere has specialized evolutively in the coordination of sequential motor events. This lateralized monitoring may have attracted in the left hemisphere the mechanisms devoted to sequential aspects of cognitive functioning and in particular the combinative organization of language. Along this line of reasoning, Wilkins & Wakefield (1995) have proposed that the anatomical-physiological adaptations that have given existence in humans to the Broca area and the parietal-temporal-occipital crossroads, occurred by exaptation (Tattersall, 1998), i.e., the appropriation of brain territories initially devoted (in this case) to the regulation of motor activities (Arbib, 2002).

 The cases of spontaneous invention of grammar mentioned before should probably be reinterpreted as demonstrating the power of brain language processes much more than the intervention of innate representational grammatical information. The question is to define on what kind of information these processes may act in the course of morphosyntactic development.

 **3.4. The learnability question**

 Traditional linguists (e.g., Hjelmslev, 1928; Tesnière, 1966) were keen to suggest that formal grammatical notions could not be learned because they cannot be readily identified in sentence surface. More recently, Bickerton (2009) confirmed that the rules determining the references of empty categories in generative grammar (e.g., trace elements) cannot be learned because they have no physical-perceptive expression. The same remark can be done for other aspects of generative theory.

 “Classical” learnability theorists (Gold, 1967; Wexler & Culicover, 1980) established that certain types of grammar, including hierarchical grammars of phrase structure, are learnable only if the learner is exposed to appropriate input (argument of the richness of stimulus), has information on what is ungrammatical in a given language (negative evidence), and can rely on a mechanism capable of selecting the grammar to which the input corresponds best among a series of possible grammars.

These conditions are not met in natural language learning. Children are exposed to little or no negative evidence regarding the grammaticality of their utterances (Rondal, 2006). No selection mechanism is provided by the environment as to possible alternative grammars. And the psycholinguists with a generative orientation (e.g., Pinker, 1994; Haegeman, 1994; Yang, 2016) sustain that natural language input to children is never diversified nor saturated enough with the relevant grammatical structures to allow the induction of a formal grammar out of scratch (argument of the poverty of the stimulus).

 Chomsky (1965) adds: “… much of the actual speech observed consists of fragments and deviant expressions of a variety of sorts” (p.201), which contributes to make of natural input a poor resource for grammatical learning. This would complicate the grammatical learning task even further. No study of the grammaticality of everyday speech has been published to the best of my knowledge except for the input to language-learning children which has been found to be grammatical (see Chapter 7).

 Contrasting with the above indications, one finds a series of experimental works in the connectionist orientation showing that an artificial network of neurons can isolate in a given language input the distributional equivalents and analogues of morphosyntactic categories. Artificial networks of neurons are devices used in artificial intelligence and cognitive sciences. They proceed from a metaphorical model involving formal neurons (i.e., statistical algorisms) connected to each other. So-called simple recurrent networks (Elman, 1990, 1991) have several layers of neuronal units allowing for a « working memory. » This enables the system to assess new entries with respect to previously analyzed ones and treat information in a cumulative way.

 In a pioneering study, Rumelhart and McClelland (1986) have fed an artificial network with 420 pairs of verbal stem-English past tense and tested acquisition with regular and irregular pairs not presented in the learning phase. The machine could learn the task and its learning curve reproduced the typical U development observed in children (i.e., first, correct indiscriminate marking; then, over-regularization of irregular forms; and last, correct marking of regular and irregular forms; see Chapter 7). One may want to see Pinker and Prince (1988) and MacWhinney and Leinbach (1991) for technical and conceptual criticisms on Rumelhart and McClelland’s study. Other works in the same vein and with corresponding methodologies have been conducted (e.g., Grossberg, 1987; Marchman, 1993; Plunkett & Juola, 1999) confirming the preceding findings and extending them to other affixes.

 Several works in the connectionist orientation have been concerned with syntactic learning (e.g., Elman & Zipser, 1988; McClelland, St. John, & Taraban, 1989). See Elman et al. (1997), Steedman (1999), Rohde & Plaut, 1999), and Dell, Chang, and Griffin, (1999), for additional data, reviews, and discussions.

 Taken together these studies suggest that a recurrent network has the capacity of identifying in a given input the distributional equivalents of several syntactic categories (for instance, distinguishing verbs from nouns, indirect from direct objects, transitive from intransitive verbs), and even dealing with the productive and receptive treatment of sentences involving centrally embedded clauses, solely exploiting transitional probabilities between items with no *a priori* grammatical information and no corrective feedback. Marcus (1996) has criticized these works and the conclusions reached. He argued that recurrent neural networks cannot generalize outside the sample of language provided because they have no explicit knowledge of the formal categories of noun, verb, etc. Such networks do not really partition grammatical and ungrammatical sentences. They simply exploit the probabilities among successive items. See Rohde and Plaut’s (1999) reply to Marcus, insisting that analogies between items and sequences of items are strong enough to sustain the learning of a number of formal characteristics of sentences.

 Perfors, Tenenbaum, and Regier (2011) have applied Bayesian statistics for inductive inference to two corpora of adult speech addressed, one to Adam and the other to Eve (two “Harvard” children studied by Brown, 1973)[[22]](#footnote-22). The corpora were extracted from the world data base in child language (Childes; MacWhinney, 2000). Bayesian statistics does not demand, as traditional inferential statistics, that one decides beforehand on a precise hypothesis regarding a possible experimental or observational outcome, to be confirmed or invalidated later relying on a predetermined error probability level. It permits starting with a preliminary explanatory model which will be confirmed, refined, or modified, depending on subsequent data analysis. Language-wise, this methodology allows, for example, comparing a series of sentences of a given syntactic type in a corpus with the structural representations that a given grammar can legitimately generate.

 In Perfors et al.’s study (2011), an ideal learner was supposed equipped with two cognitive abilities: (1) a representation of the forms that a hierarchical grammar of phrase structure may take; and (2) access to an inferential Bayesian device capable of calculating the relative probabilities of a series of grammatical representations from the data.

 Results show that a Bayesian device can detect the structural organization of yes/no interrogative sentences in English, with the fronting of the auxiliary verb with respect to the pronoun-subject (for example, in *Are you coming*?). More generally, the Bayesian mechanic applied to the two corpora of language suggests that a hierarchical grammar of phrase structure provides a better approximation to the sentence statistical distributions than a series of nonhierarchical grammars. The implication appears to contradict Gold’s (1967) statement, as mentioned earlier, that it is not logically possible to apprehend statistically a hierarchical phrase structure from positive evidence.

The methodology utilized by Perfors et al. (2011) demands that each lexical item in the corpus be coded according to its formal category (noun, verb, preposition, etc.) and that each sentence receives a structural definition in such a way as to be able to compare its distribution to the theoretical prescriptions of a grammar. Nothing suggests, however, that in natural conditions learners, particularly younger children, can rely on a coding ability, memory capacity, and sophisticated computational techniques, as those involved in Perfors et al.’s brand of experiment. There is no guarantee either that ordinary language learners are ideal learners in neither any particular respect, nor that they necessarily learn to front the auxiliary verb with respect to the pronoun-subject in yes/no interrogative sentences according to the specification of a hierarchical phrase structure grammar. Other studies suggest that the same operation can be realized more simply in noting the relative positioning of relevant elements in sentences (Fitz, 2010).

Abend et al. (2017) have also made use of a Bayesian device to assess the possibility of syntactic bootstrapping in language acquisition. The notion of syntactic bootstrapping, first suggested by Gleiman (1990), proposes that children acquire their native syntax in associating structured representations of word meaning in sentences with their syntactic categories (e.g., noun, verb, etc.). In the original proposal, children are supposed to be endowed with an innate knowledge of the syntactic categories. Abend et al. (2017) report that a robot learner equipped with a Bayesian mechanic exposed to the Eve (Harvard child) corpus of the Childes data base as input, exhibited a syntactic bootstrapping effect in which previously learned constructions facilitated the learning of novel words, including the early bias favoring the learning of nouns over verbs (as predicted by the syntactic bootstrapping theory on the dubious ground that verbs are learned with a delay with regard to other syntactic parts of speech because the information allowing their acquisition would not be available during the early stages of language acquisition).

 Hsu, Chater, and Vitányi (2011) generalizing in relation with the Bayesian approach, indicate that “…*any* (their stress) language generated from any computable generative probability distribution (including any grammars proposed in linguistics) can be precisely identified given a sufficiently large identical independently distributed sample” (p.388). It is not at all obvious, however, that this possibility may be simply transposed to the natural conditions of language acquisition. And yet, as will be documented in the next section, neurological data suggest that the distributional equivalents of at least some grammatical categories and syntactic constituents are represented in the brain.

 **3.5. Brain representations**

The efforts to map language in the brain have taken several forms since the work of the French neurologist Paul Broca in the second half of the 19thcentury (see the book edited by Grodzinsky & Amunts, 2006, for a detailed historical account). The first brain maps located expressive language mechanisms in Broca’s region (Brodmann areas 44 or pars triangularis and 45 or pars opercularis in the third-frontal gyrus, just anterior to the face area of the motor cortex in the left-cerebral hemisphere in most right-handed individuals). The receptive and semantic mechanisms were located in Wernicke’s region (Brodmann area 22 in the posterior section of the superior-temporal gyrus area also of the left-cerebral hemisphere in most right handed individuals). The two regions are connected by a large neuronal pathway, the *arcuate fasciculus*. Since then new observations and experimental methods have allowed probing the cerebral correlates of language functioning in more precise terms, taking into account the linguistic levels of representation and particularly the morphosyntactic component.

Early clinical neuropsychological studies of brain-injured patients had suggested that specific brain areas respond to different linguistic categories, in particular nouns and verbs. Broca’s aphasics typically have marked difficulties with verb naming whereas Wernicke’s aphasics experience difficulties in dealing with nouns (e.g., Caramazza & Hillis, 1991).

Shapiro, Shelton, and Caramazza (2000) and Shapiro and Caramazza (2003a) have documented the case of a patient (RC) with a left-hemisphere lesion demonstrating more marked difficulties when requested to repeat or inflect words or pseudo-words presented as verbs (e.g., *He judges*; *He wugs*) than the same forms presented as nouns (e.g., *the judges*; *the wugs*). Another patient (JR) with a left-cerebral lesion more posterior than RC, presented an opposite difficulty, that is, a more marked difficulty with nouns than with verbs. Shapiro and Caramazza (2003b) concluded that nouns and verbs are indeed represented in different areas of the left brain.

 However, other studies have placed the accent more on the conceptual and semantic dimensions associated with the formal categories of nouns and verbs in the difficulties exhibited by aphasic patients (for example, Miceli et al., 1984; Berndt et al., 1997).

 Attempts at identifying the brain basis of grammatical functioning using neuroimaging have been made from the 1990ies on (see Grodzinsky, 2003, for a summary of early works, and Grimaldi, 2012, for a review of more recent investigations). Various non-invasive brain study techniques have been developed together with progresses in electrophysiology, informatics, and neuroimaging. They include: electroencephalography (EEG), event-related potentials (ERPs), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG). The techniques vary somewhat in spatial and temporal resolutions. For example, ERPs has millisecond accuracy, fMRI has a resolution of a few thousand neurons per pixel, and MEG has a finer temporal resolution than fMRI. They all share a reductionist approach considering that components or processes of the linguistic system are potentially realized in the brain and can be functionally localized.

Paradis (2004) acknowledges that the promises of cognitive neuroimaging are enormous, but he lists a series of conceptual and technical problems with current neuroimaging studies: e.g., different imaging methods; lack of sophisticated linguistic theory underlying the experimental designs (less so in more recent times, however; see below); use of an extended variety of tasks and stimuli rendering difficult the comparison of data between research works; generalizations to the language system as a whole when results are based on tasks relying on declarative memory (see below); little or no taking into account of the difference between implicit and explicit learning and knowledge. Additionally, it is almost impossible most of the time to ascertain with confidence which portion of the language network is activated as the “where something is represented” depends on how it is processed, which depends on the part of the memory system supporting it.

In spite of these limitations and caveats, it may be worth considering, albeit with caution, the major data emerging from this domain of research regarding brain representation of linguistic categories and functions.

PET studies have confirmed that nouns and verbs elicit responses from different regions of the brain: verbs or action names activate the prefrontal and frontal-temporal regions and the supramarginal gyrus - the inferior portion of the parietal lobe (Damasio & Tranel, 1993; Tranel et al., 2001; Cappa & Perani, 2003). Preissl et al. (1995) found ERPs differences between nouns and verbs in the frontal- and the posterior-brain regions 200milliseconds after stimulus onset. They suggested that neuronal activity in the motor cortices is due to a strong association elicited by the verbs while the activity in the visual cortices is due to a strong visual association elicited by nouns. Pulvermüller et al. (1996) have reported similar findings based on an EEG study.

Kupperberg et al. (2006; see also Kupperberg et al., 2003) have reported neurological observations based on ERPs in a comprehension task of sentences varying in semantic transitivity (a notion concerning the transfer of the property of a physical entity onto another one; Hopper & Thompson, 1980)[[23]](#footnote-23). Results suggest that receptive treatment is largely influenced by the semantic relations between verbs and their obligatory arguments. These authors suggest that sentence semantic and syntactic aspects are activated in parallel during the comprehension task in a kind of dynamic interaction: one assessing the plausibility of the semantic relationships and the other testing the syntactic correctness of the strings of words.

One interpretive problem with these studies is that action words in verbal tasks can correspond to nouns referring to actions, movement, or tools. Pulvermüller (1999) and Perani et al. (1999) have reported PET data showing that such nouns may yield activation results similar to verbs. Berlingeri et al. (2007) have observed a shared verb-related activation of a dorsal premotor and parietal network in a lexical fMRI task, which they interpreted as pointing to a strong relationship between verb representation and action-oriented spatial knowledge.

 Reservations regarding the brain representation of linguistic categories have been formulated. For example, Warburton et al. (1996) have reported PET data showing that all processes associated with word retrieval produced corresponding levels of activation for nouns and verbs in prefrontal, temporal and parietal brain regions. Tyler et al. (2001) have controlled nouns and verbs for length, frequency, and imageabilit in a PET study based on a task of lexical decision. They have reported that both nouns and verbs elicit corresponding activation in the left-prefrontal and inferior-temporal regions and suggested that linguistic categories are less differentiated at brain level than previously thought.

 Shapiro et al. (2001) have published a work providing at least a partial control of the semantic dimension in word processing at brain level. They used repetitive transcranial magnetic stimulation (rTMS) to suppress the excitability of a portion of the left-prefrontal cortex and assess its role in dealing with nouns and verbs. In one experiment, real words were used as nouns and verbs. In a second experiment, pseudowords were utilized. In both experiments, response latencies increased for verbs but were unaffected for nouns following rTMS. According to the authors, these results confirm that grammatical categories have a neuroanatomical basis and that the left-prefrontal cortex is selectively engaged in processing verbs as grammatical entities.

 The question remains, however, to establish on which basis the brain may classify lexical items in grammatical categories. Stamatakis, Randall, and Komisarjevsky Tyler (2007) have showed that differential brain representation depends on morphosyntactic markers. They compared nouns and verbs in a classification task with fMRI. Participants were presented lexical stems (e.g., *snail*, *hear*)) and corresponding inflected forms (e.g., *snails*, *hears*). They observed greater activation in the left-frontal-temporal areas for inflected verbs than for inflected nouns while there was no difference in activation for noun and verb stems.

 In another fMRI experiment, Tyler, Randall, and Stamatakis (2008) presented their participants with homophones, that is, words which function grammatically as nouns or verbs but have the same form and meaning. Words were proposed as single stems and in phrases. In the latter condition, each homophone was preceded by an article in order to create a noun phrase (NP) or a pronoun to create a verb phrase (VP). Results showed that the brain activation elicited by single-word homophones was not modulated by their frequency of usage as a noun or verb in the language. However, homophones syntactically marked as verbs elicited greater activity in the left-posterior-middle-temporal gyrus compared to homophones marked as nouns. These results suggest that differential cortical activation for nouns and verbs also depends on syntactic function.

 Relevant differences may exist between languages in linguistic treatment. Ping, Zhen, and Li (2004) have utilized an fMRI technique to study the brain representation of nouns and verbs of native Mandarin Chinese speakers in a lexical decision task. Chinese grammar is particular in that it mostly lacks inflectional morphology on nouns and verbs. Also the distinction between nouns and verbs is less marked than in many Indo-European languages. An important portion of monosyllabic and disyllabic words in Chinese can be used indifferently as nouns or verbs. Many Chinese verbs can occur freely as grammatical subjects and many nouns as predicate with no morphological change. Ping et al. observed that nouns and verbs in Chinese activate a wide range of brain areas: frontal, temporal, parietal, and occipital, in the left and the right hemispheres, as well as the cerebellum. The involvement of the right hemisphere may be due to the orthographic features of Chinese characters in the reading task and/or to the role that lexical tones play in processing Chinese words. The caudate nucleus (dorsal striatum) was the only brain area where there was an activation difference (in favor of nouns). The authors have no explanatory hypothesis for this observation. They concluded that the neural representation of nouns and verbs is tied to the various stimulus and processing characteristics of these words in the languages.

 In a more syntactic (rather than typological) perspective, Pallier, Devauchelle, and Dehaene (2011) have conducted a neuroimaging (fMRI) study that deserves to be explained in some detail. Participants were exposed to sequences of written words affording the construction of constituent structures of variable sizes. All sequences had a fixed length of 12 items. There were several conditions. Condition (C) 1 consisted in a list of unrelated words. In C2, the sequences could be parsed as a series of constituents of size 2 (e.g., *Mary’s father*). In C3, the successive constituents could be parsed in series of 3 words (e.g., *Mary’s father’s car*). And so on until C12, where a full sentence of 12 words could be identified (e.g., *I believe that you should accept the proposal of your new associate*). An additional experimental manipulation was introduced (with a different group of participants), consisting in 12 so-called jabberwocky conditions (organized in the same way as the 12 conditions with real words) where content words were replaced with pseudowords while maintaining the morphological markers and function or closed-class words (i.e., pronouns, determiners, prepositions, and conjunctions) needed for parsing the sequence. An example of such a 12 word sequence is: *I tosieve that you should begept the tropufal of your tew viroate*. In such sequences, the lexical-semantic content is drastically reduced while the syntactic constituent structure is maintained allowing separating the effects of syntactic constituency and relational semantics. This jabberwocky experiment also allowed controlling for transition probabilities between words. In the material with real words, constituency is confounded with transition probability between words. In the jabberwocky conditions, transition probability was near zero. Therefore, any systematic variation in brain activation along the 12 jabberwocky conditions would suggest a sheerer sensitivity to syntactic constituent structure.

Results showed an increase in fMRI activation amplitude with constituent size and a corresponding increase in activation delay. Results also allowed isolating two regions in the left-cerebral hemisphere with different operational agendas. Inferior-frontal and posterior-temporal regions showed constituent size effects (i.e., increased activation from condition 1 to 12) regardless of whether real content words or pseudowords (jabberwocky) were presented. On the other hand, regions in the temporal lobe (anterior-superior-temporal sulcus and temporal-parietal junction) showed a corresponding effect only in presence of lexical-semantic information. These results suggest that the brain regions dealing with syntactic constituents of sentences may operate autonomously from the regions devoted to semantic analysis. The former apparently can extract syntactic frames based on function words and morphological information alone. Expectedly, in both regions, activation was delayed in response to the largest sequences confirming that more elaborated processing takes more neurological time.

Ding et al. (2016) have conducted a series of experiments using magnetoencephalography (MEG) in patients being monitored for neurosurgery. Participants listened to sentences in both English and Mandarin Chinese in which the hierarchical structure between words, phrases, and sentences had been dissociated from intonation as well as statistical word cues. The sentences were presented with an identical timing between words. Participants were exposed to predictable sentences (e.g., *New York never sleeps*), grammatically correct but less predictable sentences (e.g., *Pink toys hurt girls*), and word strings (e.g., *eggs jelly pink awake*). Results showed that the participants’ brain tracked three sentence components in a temporal sequence: words are identified first, and then phrases and sentences are constructed. The authors interpret their data as evidence for hierarchical syntactic structure in sentence parsing arguing that tracking probabilistic relationships between words is not sufficient to explain the cortical encoding of linguistic constituent structure.

A review and discussion paper by Grodzinsky and Friederici (2006) allows going one step further in linking syntactic processing and brain representation. These authors propose that components of human syntactic knowledge are located in distinct brain loci. They base their “syntactic-topic conjecture” on later versions of generative grammar (*Government and binding* and *Minimalist program*), which they feel supply a coherent conception of the nature of syntax representation. The Chomskyan idea that semantics interprets basic syntactic structure and not the other way around, is central in Grodzinsky and Friederici’s language model.

 This model assumes three processing phases: an initial phase during which local phrase structure is built on the basis of lexical information (lexical classes also annotated for syntax-relevant characteristics); a second phase during which syntactic and thematic (semantic) dependency relationships are established; and a third phase where the final integration of the information takes place.

 Grodzinsky and Friederici (2006) mention a series of research works with fMRI, MEG, and ERPs, including their own work, proposing that the left-frontal operculum (an area in the inferior-frontal gyrus) is involved in the computation of phrase structures relying on the merge operation (in particular, noun, verb, and prepositional phrases). The anterior-superior-temporal gyrus appears to be involved also in sentence processing.

 Computing dependency relations (in phase 2) would appear to be supported by Broca’s areas (Brodmann 44 and 45). Particular activation seems to be linked to the treatment of sentences in which word order does not match canonical expectations (for example, object-first sentences). In such cases, hierarchical syntactic structures must be constructed, which implies moving elements and a more important demand on working memory.

 Processes of syntactic integration in phases 3 would appear to be subserved by the left-posterior-superior-temporal gyrus. This area is particularly activated in cases where the parser encounters ungrammatical strings or when processing complex sentences.

In conclusion to this section, although much remains to be done and many suggestions and conjectures still need to be substantiated, it would appear that at least a sketch of a brain map for syntax representation and syntactic processing is emerging. In particular, it seems that the distributional equivalents of several formal classes of words may be considered as having brain representation distinct from their meaning. It would also seem that some syntactic processes predicted by linguistic theory might be tied to the activation of particular brain centers in the left-cerebral hemisphere.

 It might seem odd in absolute value that, as proposed in the present essay, there could exist two separate systems, one semantic and the other formal, each theoretically with the capacity to sustain morphosyntactic functioning. This may appear contradictory to a kind of economical or Occam’s razor’s basic explanation of natural morphosyntax. And yet, as Paradis (2004) recalls “…it has been shown that the brain favors redundancy rather than economy in its functioning” (p. 33).

 However, it will be necessary in future work to move from simple brain-mapping research strategies toward a search for dynamic neurocognitive linguistic structures. Grimaldi (2012) also recommends that in order to avoid “detrimental reductionism,” that is, simply associating linguistic computation hypothesized at theoretical level with particular neurobiological computation, it is necessary to provide a real interdisciplinary background between neurolinguistics, psycholinguistics, and linguistics.

 As no genetic basis has been found for the representational innateness hypothesis, it follows that the conjectured elements of syntax must end up in the brain as a consequence of distributional learning following relatively longer periods of exposure to language input, given that human learners cannot rely on computation and memory means comparable to those of computationally faster artificial networks and that, possibly (as argued in the generative tradition), this input may never be fully complete or saturated enough regarding particular syntactic structures. There is no guarantee either that formal learning will ever be complete in proportion to the repertoire of formal and functional linguistic categories.

This learning is implicit as regular standard users do not have conscious knowledge of the morphosyntax of their language. It will be dealt with in Chapter 7.

 CHAPTER FOUR

 SEQUENCES AND HIERARCHIES

 As seen in preceding chapters, hierarchical phrase structure plays a major role in theoretical accounts of morphosyntax, being often taken for granted (see Bybee, 2002; and Langacker, 2010, however). Frank, Bod, and Christiansen (2012) argue in opposite way that sequential sentence structure has more explanatory power for language use. This dimension, they indicate, has been largely neglected in cognitive sciences. At a superficial (possible trivial) level, it is undisputable that sentences exhibit linear structure while hierarchical structure is only revealed through linguistic analysis. There are other arguments, however, in favor of a fundamental role of sequential structure in language use.

 As documented in the preceding chapter, syntax processing may rely on domain-general constraints including neural substrates devoted to sequencing skills, anticipating the emergence of human language, that were subsequently recruited for language treatment. This implies that the basic operations in syntactic functioning and acquisition may be fundamentally sequential in nature even if hierarchical mechanisms can also be involved.

 Reviewing neurological evidence, also in the preceding chapter, it was concluded that there exists a support for the view that some grammatical categories and processes (as well as their semantic counterpart) are represented at brain level. However, other neuroimaging studies with artificial grammars (see chapter 7 for information on this experimental technique) suggest that the left-inferior-frontal gyrus is also involved in sequential learning (Petersson, Folia, & Hagoort, 2012). Moreover, the degree of integrity of white matter in the vicinity of Broca’s area is correlated with quality of sequence learning (Flöel et al., 2009). In a further study, Udden et al. (2017) had participants implicitly acquiring an artificial language with non-adjacent dependencies. Successively, they could discriminate between grammatical and non-grammatical sequences. However, subsequent to transcranial magnetic stimulation (TMS) of Broca’s region, discrimination was impaired compared to the same stimulation of another brain region irrelevant to language (the cranial vertex). This study shows that Broca’s region is causally engaged in the processing of non-adjacent dependencies and that the implicit processing of non-adjacent dependencies is related to Broca’s region.

 Correspondingly, in agrammatic-aphasic patients exhibiting important difficulties in natural language both in production and in comprehension, there is evidence of a marked deficit in sequence learning (Christiansen et al., 2010). A corresponding difficulty in processing musical sequences has been documented by Patel et al. (2008) in patients with agrammatism, suggesting that sequential ability is not specific to language treatment.

 Frank et al. (2012) have analyzed various research data in psycholinguistics and computational linguistics underlining the importance of sequential treatment in sentence processing. It has been demonstrated, for example, that in speech production studies the rate of errors in subject-verb (number) agreement depends on the sequential rather than the hierarchical distance between words. These authors give the following example: *The coat with the ripped cuffs by the orange balls were\**… People accustomed to processing texts on computers are familiar with this type of erroneous corrective suggestion proposed by spelling corrector automata.

 Frank et al. (2012) quote from works on reading sentences showing that local information can have precedence over distal information. For example, it is known that in a reading task more predictable words are read relatively faster. It was observed that in a sentence like *The coach smiled at the player tossed a frisbee*, the word *tossed* (embedded verb) is read more slowly that *thrown* in the corresponding sentence *The coach smiled at the player thrown a frisbee*, suggesting, according to the authors, that the segment *the player tossed* was erroneously (temporarily) identified by the participants as a possible phrase.

In a similar way, several syntactic structures, for instance auxiliary fronting, generally thought to be unlearnable without taking hierarchical dependencies into account, can be learned by tracking the relative sequential positions of words in sentences (e.g., Fitz, 2010, as mentioned in Chapter 3).

Frank et al. (2012) also argue that language comprehension most of the time does not need relying on the hierarchical structure of sentences. The reason is that the meaning of a sentence uttered in a functional context is not just an integration of its constituent parts. It generally depends on a series of factors such as information on the speaker, general world knowledge, pragmatic factors, prior verbal exchange, etc. This is why in linguistic communication, as is well known, the importance of sentence structure may be reduced.

In concluding their analysis, Frank et al. (2012) suggest that although it is always possible to view sentences as hierarchically structured, this way of conceiving of language syntax is mostly a byproduct of analyzing sentences by dividing them up into parts, subparts, and so on until the ultimate constituents. They insist that psychologically such hierarchical decomposition is not a necessity. Frank et al. (2012) indicate, however, “This is not to say that hierarchical operations are non-existent, and we do not want to exclude their possible role in language comprehension or production. However, we expect that evidence for hierarchical operations will only be found when the language user is particularly attentive, when it is important for the task at hand (e.g., in metalinguistic tasks) and where there is little relevant information from extra-sentential/linguistic context” (p. 7). Most importantly, they add: “… even if some hierarchical grouping occurs in particular cases or circumstances, this does not imply that the operation can be applied recursively, yielding hierarchies of theoretically unbounded depth, as is traditionally assumed in theoretical linguistics. It is very likely that hierarchical combination is cognitively too demanding to be applied recursively” (ibidem).

 This is an important point and it seems to be conceptually related to my incompleteness suggestion (in the preceding chapter) regarding regular language users’ knowledge of linguistic notions.

Frank et al. (2012) seem to assimilate sequential and linear treatment. Actually, sequential treatment can be but does not have to be linear (in the same sense that a Markovian scheme does not need to be limited to first-order concatenations; Manning & Schütze, 1999). Etymologically, the terms *sequence*, *sequential* (from the latin *sequor*, *sequitur*, *sequential*, i.e., *follow*, *following*, but also *go with*, *accompany*) does not have the same restriction of immediacy and juxtaposition as the terms *line* and *linear*. A sequential process may have some important degree of flexibility. It needs to be globally oriented left to right and the morphosyntactic operations must be realized one after the other. However, an anterior operation may bear upon an ulterior even remote segment of the sentence and there may be movements back and forth in surface depending on the speaker’s self-editing activity, which involves working memory (buffer and momentary storing of lexical entities and relevant information). I will illustrate this dynamic in Chapter 6.

There is a temporal (hence sequential) dimension in tree-like representations of sentences. As reported, they are labelled “temporally ordered hierarchical structures” (Fitch & Martins, 2014). A tree-like diagram must be read from left to right even if the final product can by be diversely organized as a consequence of the transformations intervening from base to surface structure. Merge and move, it should be noted, are also temporally oriented operations.

In a converse way, and as will be analyzed in the following chapter, semantic relations are also abstract notions. Their sequential combinations may also reflect a hierarchical organization. For example, underlying a sentence like *John is cutting small wood in the forest*, one finds a structure whereby the relation agent-action-process-patient (*John cuts wood*) is nested in a locative indication (*in the forest*). One additional indication is embedded in the main relation (wood is *small*). And the episode is given as being in progress at the time of expression. The sequential-hierarchical issue in comparing syntax and relational semantics is not one of a clear opposition between two supposedly orthogonal dimensions.

As Frank et al. (2012) also remark, psychologically the opposition between the two dimensions has much to do with the complicated, removed from sentence surface, and cognitively little accessible use that generative linguistics (and psycholinguistics) have made of the notion of hierarchy.

 In the next chapter (5), the semantic-relational system, which I take to be one of the two systems regulating morphosyntactic functioning, will be explained. In Chapter 6, the relationship between this basic system and the formal morphosyntactic one will be examined.

 CHAPTER FIVE

 SEMANTIC-RELATIONAL THEORY

 Relational semantics is concerned with the relationships between words in multiword expressions. Semantic relations are built from lexical items but their combinations involve a hierarchical dimension as illustrated in the preceding chapter. Lexical meaning can be decomposed into bundles of semantic features (Katz & Fodor, 1963), which may be ordered in various ways but this is intrinsic to the individual lexemes. Lexemes and families of lexemes can be located in hierarchical schemes (e.g., hyponyms and hypernyms), but this is a property of the lexicon not of particular lexemes (see Lyons, 1977, and Murphy, 2002, for extensive analyses). Also lexemes are endowed with semantic properties specifying which roles they can have in semantic combinations. For instance, an action verb has to have an agent as argument and an action-process verb an agent and patient argument (see below).

There is no question that semantic relations are not abstract entities. They are obviously. For example, the semantic relations that link entity possessed and possessor, entity and beneficiary, entity and direction or location, or agent-action-process-patient, and so forth (see below), are quite abstract, i.e., they correspond to schemata purporting to represent economically a number of corresponding concrete instances. These semantic abstractions, however, may be thought to be more accessible to regular language users than formal linguistic categories as they are closer to corresponding conceptual notions.

 **5.1. Chafe’s and related models**

 A number of semantic-relational theories have been proposed over the years since the “rediscovery” of semantics in the second half of the 19thcentury (Fillmore, 1968, 1977; Chafe, 1973; Hutchins, 1973; Langacker, 1987, 2008; Schlesinger, 1988, 1995; Van Valin, 1999, 2003) following its quasi exclusion from structural linguistics.

These theories have in common the idea that relational semantics is a key variable in language use and that it determines morphosyntactic organization to a large extent. They differ from generative semantics, an offshoot of generative grammar as illustrated in Chapter 1, not in the idea of associating deep structure with semantics but in the conception of semantics itself, judged to close still to Chomsky’s interpretive semantics.

Fillmore (1968, 1977) proposes that a sentence is composed of a proposition and a modality constituent specifying mode, time, aspect, and polarity. The proposition is made of a verb and one or several noun phrase(s) associated with the verb in particular case relationships. It can be modified optionally by an adverbial constituent.

 Table 5.1 summarizes the case model of Fillmore (1968).

 1. Sentence = proposition + modality constituent

 2. Proposition = case(s) + (adverb)

 3. Cases = 1. Agentive: entity typically but not invariably animate that

 instigates the action expressed by the verb.

 2. Instrumental: inanimate force or physical object causally

 involved in the action or the state expressed by the verb.

 3. Dative: animate entity affected by the action or the state

 expressed by the verb.

 4. Factitive: entity resulting from the action expressed by the

 verb.

 5.Locative: localization, spatial orientation, origin, or direction

 of the action or the state expressed by the verb.

 6.Objective: any entity that can be expressed by a noun whose

 role in the action or the state expressed by the verb is

 identified in the semantic interpretation of the verb itself.

According to Fillmore, the semantic feature necessary for a noun to be included in a case framework are specified by obligatory rules of the type N = [animate] ad, i.e., a noun in an agentive or dative framework must contain the feature animate. Verb insertion is only possible in a corresponding case framework. For example, the verb *run* needs to be inserted in a noun agentive context.

Several operations intervene on the way between deep and surface structure. Prepositions (or postpositions in some languages) are added, affixation, displacement of elements, inflections, agreement between grammatical subject and verb, pronominalization, and possibly passivization, are performed. Whenever there is an agent in semantic structure, it electively becomes the grammatical subject in sentence surface. If not, this is the instrumental or the objective case. Some passivizable verbs make exception like, for example, *give* which authorizes either an objective or a dative case to be grammatical subject (for example, in *The books have been given to my brother by John*). Fillmore (1968) also argues that there is no purely syntactic level. Chomskyan-like syntactic deep and intermediate structures are arbitrary levels with no real usefulness.

 A particular difficulty with casual grammar is reaching agreement on a nomenclature of cases. Alternative classifications to that of Fillmore have been proposed (see, for example, several contributions in Anderson & Dubois-Charlier, 1975). Godin (1976) suggested to extent the notion of case to the modality component in distinguishing modal cases in addition to the propositional cases of Fillmore.

 Hutchins (1973) has proposed a model in which meaning is organized in sememic formulae (combinations of semantic features or semes - semons in Hutchins’s vocabulary). Each sememic formula is converted into a linear string of lexemes containing information on syntactic structure. Some sememes are realized as lexemes and others as syntactic function or grammatical categories. The sememic formulae are similar to the casual structures in Fillmore’s grammar. The model is generative but not transformational. There are no rules to proceed from one syntactic structure to another. Parenthood between syntactic structures is accounted for by parenthood in the semantic structures from which the syntactic structures are derived.

 Langacker (1987, 2008) defined a model in which syntax is entirely regulated by the semantic dimension and meaning is equated with conceptualization. Three considerations are central: (1) semantic structures are not universal, they are specific to each language; more precisely, semantic structures are characterized relative to cognitive domains but they derive their value by construing the content of these domains in specific fashion (Langacker, 1976); (2) syntax does not correspond to an autonomous level of representation; it may be reduced to the structuring and symbolization of conceptual contents; and (3) there isn’t any real separation between grammar and lexicon.

 The first point is curious. As semantic structures are cognitive by nature (the word *cognitive* is used in the title of Langacker’s 1987 and 2008 books), it is difficult to understand how relational semantics in itself could be entirely specific to each individual language. The semantic specificity question will be discussed later in this chapter together with the Worf-Sapir linguistic relativity hypothesis. The second point flies in front of the traditional Chomskyan argument of the autonomy of syntax. Langacker insists that there is no separation between syntactic and semantic constituents; an implication that does not seem to follow judging from the neurolinguistics data available nowadays. The third point is reminiscent of the lexicalist hypothesis developed in generative grammar following the suggestions of Bresnan (1978) and Bresnan and Kaplan (1985) as documented in Chapter 1.

 Schlesinger (1988, 1995) proposed a theory in which semantic structures supply an intermediate level between cognitive structures and morphosyntax. He suggests that some case notions sometimes presented as semantic primitives actually are psychological constructs corresponding to cognitive notions. Schlesinger (1995) retains the universality of the cognitive notions underlying the semantic-syntactic cases but not that of the cases themselves, which may vary from one language to another. The semantic-relational nomenclature utilized by Schlesinger (1987) is close to the ones of Fillmore and Chafe (see below for the latter).

 Chafe’s semantic-relational theory (1973)[[24]](#footnote-24) is one of the most elaborated proposals available, and, to my taste, the most elegant one.

 Chafe (1973) expresses a strong opposition to any formal syntactic model pretending to capture the essence of the language system. He writes: “When we use language we start with something that we have to say - with meanings. If we are going to communicate these meanings, they must be structured in a way that conforms to the semantic structure of our language. Language then proceeds to give us a way to convert this semantic structure into sound. It is evident that the syntacticist model has nothing directly to do with language use” (p. 66). Chafe is convinced that semantic structure is the major component of language. His orientation is psychological and cognitive. He specifies: “Language enables a speaker to take concepts which are present in his own mind and to call forth these same concepts in the mind of his hearer. The sounds which travel from him to his hearer do not normally engender new conceptual units in the mind of the latter. They activate concepts already there, concepts which both the speaker and his hearer have in common….If concepts like these were idiosyncratic, language would not be able to function” (ibidem, p. 76).

Chafe (1973) further posits that the semantic structure identified in English language has general value. He analyzes an Iroquois language, Onondaga, a language of the polysynthetic type (in which words tend to consist of several morphemes agglutinated or fused).The semantic structure of this language differs very little, and only in trivial ways, from that of English (for example, Onondaga distinguishes between dual and plural and between punctual and iterative processes). The important differences between English and Onondaga arise largely as the result of post- semantic processes (see Chapter 6).

 In Chafe’s model, every sentence is built around a verb organizing one or several nouns. He specifies: “There seems no need for some independent symbol *S* as the starting point for the generation of sentences; the verb is all the starting point we need. What we may call for convenience a sentence is either a verb alone, a verb accompanied by one or more nouns, or a configuration of this kind to which one or more coordinate or subordinate verbs have been added” (1973, p. 98)[[25]](#footnote-25).

 Chafe’s theory is divided in two parts: a semantic level and a post-semantic one. The latter will be dealt with in the following chapter. It explains how the semantic-relational notions are set in correspondence with the syntactic ones.

 Fig. 5.1 illustrates schematically the general organization of Chafe’s (1973) theory of semantic structure.

 As already indicated, semantic structures are not naturally sequenced. They simply associate relational meanings which may be complex. Semantic relations may be compared to a mobile sculpture in a semantic space (the suggestion is from Chafe, 1973, p. 5). On their way toward surface structure, they need to be linearized (better flexibly sequenced). The notion of linearization in Chafe’s theory is more akin to that of sequence (as defined in Chapter 4) than to the restrictive notion of a line. By symbolization, it is meant the conversion of “linearized” semantic units into an arbitrary (idiom specific) configuration of symbolic units, the phonetic structure.

Central aspects of Chafe’s model concern the relationships between verbs and nouns.

Table 5.2 summarizes the categories of verb and noun relations according to Chafe’s (1973) model (the examples are from Chafe).

 Verbs Nouns

 1. State (*The wood is dry*) Patient

 2. Action (*Harriet sang*) Agent

 3. Process (*The wood dried*) Patient

 4. Action-process (*Michael dried the wood*) Agent, patient

 5. Ambient (*It’s hot; It’s snowing*) All-encompassing states or actions

 6. Experiential (*Tom wanted a drink*) Experiencer, patient

 7. Benefactive (*Tom owns a car*) Beneficiary, patient

 8. Instrumental (*The door opens with a key)* Instrument, patient

 9. Completable (The *candy costs ten cents*) Complement, patient

 10. Location (*The knife is in the box*) Location, patient

 The technical notion of agent refers to the animate entity responsible for an action. Only some nouns are eligible to occur as agents of action verbs. This depends on the ability of an entity to be specified as having the power to do something. Patient designates an entity in a particular state or an entity affected by an action or a process expressed by the verb. Chafe distinguishes between animate and inanimate patients. Ambient verbs involve particular states or actions (but not processes) covering the global environment.

 The generation of a semantic structure proceeds in the following way. At the outset, the verb is assigned a selectional unit, like state, action, process, action-process, or ambient. Once assigned, the selectional characteristic determines the rest of the structure in limiting the choice of a verb root and specifying the number and relation of accompanying nouns.

 Derivate operations in lexicon allow transforming a state into a process, for example in adding the feature “inchoative” (e.g., *open*, *opening*). In a converse way, the derivational unit “resultative” has the effect of transforming a process into a state (e.g., *break*, *broken*). Other derivative operations involve “causative” (transforming a process into an action-process; e.g., *The dish broke*; *Linda broke the dish*; identifying someone as the cause of the process) and “deprocessive” features (transforming an action-process into an action).

 Particular derivational units cannot apply in immediate succession to the same verb root (verb root + inchoative + resultative would result in a lexical monster); other can (for example, inchoative + causative; *The dish is broken*; *Linda is broking the dish*).

 Some derivations can be applied to nouns roots as well as to verb roots, to produce noun roots of a different kind or even to produce other verb roots. For example, a root which is intrinsically animate but non-human may be converted into one that is derivatively human through the derivational unit “anthropomorphizer” (e.g., *elephant* + anthropomorphizer *= Jumbo*). “Verbalizer” can transform a noun root into a verb root (e.g., *skin*, *to*- *skin*).

Categories 6 to 10 in Table 2 list secondary verbs. Experiential verbs involve mental (e.g., *know*, *want*, *like*, *remembe*r) as well as sensorial (*see*, *hear*, *feel*) verbs. They are state or process verbs. Intrinsic action verbs do not appear to involve an experiencer (intrinsic here means a verb that is not transformed through a derivational unit). Benefactive verbs express possession, which may be transitory (*John has the key*), non-transitory (*James owns a house*), alienable (*John has a sportscar*), or inalienable (*James has got a headache*). Completable verbs require to be accompanied by a noun specifying the patient involved.

 The various verbal selectional units listed in Table 2 can co-occur in relatively complex ways, for example, action-process+benefactive (*Mary sent Tom the tickets*), action-process+instrumental (*Tom opened the door with a key*), and action-process+location (*Tom threw the knife into the box*).

To this basic scheme, it is necessary to add the pragmatic indications of time, aspect, and mode, as well as the grammatical inflections introduced in surface (see below).

Simple sentences may be expanded into more complex sentences in which two, or more verbs, each with their accompanying nouns, are present. Clause conjunction may be viewed from a semantic-relational point of view, as a relation between two or more verbs by which the meanings expressed are considered to be of equal importance. Disjunction is a particular case of conjunction in which an alternative is proposed to one or more sentences.

 Particular semantic relations hold between clauses (causality, condition, circumstances, time, etc.), and they determine specific surface organizations. Chafe (1973) suggests that relative clauses are based on the semantic relation of partition. For example, a sentence such as *Girls who are beautiful are popular* may be viewed as a complex of a main indication “girls are popular” restricted to “(girls) who are beautiful, “through the intervention of a semantic relation called partitive; meaning that actually only a sub-class of the girls are popular. However, the partitive relation does not apply if the noun in the main clause is unique (hence cannot be partitioned). Chafe claims that there are no semantic structures which would lead to surface structures like *Marcia who is beautiful is popular*.

A particular group of sentences, labelled direct or indirect questions, have in common that pragmatically they are information or confirmation requests. Questions may be disjunctive (so-called yes-no interrogatives) or non-disjunctive (wh-questions). The linguistic response requested in the former questions is of a general type: approval/denial). Wh-questions request a particular information specified in the lexical form of the interrogative pronoun (*who*, *what*, *when*, etc.). The pragmatic indexation “question” and its subtypes determine a particular organization of sentence surface and, in English in some cases, the introduction of an auxiliary verb specified lexically as *do*. This causes the transposition of the first noun in the sentence with the verb immediately following it, which in this case turns to be the auxiliary.

As indicated, Chafe’s departing axiom is that the verb is the control center of the sentence. It determines to a significant degree by its own internal specification which types of obligatory semantic categories the rest of the sentence will contain. Chafe (1973) justifies his position with the observation that no noun need be present in sentences formed around ambient verbs. The meaning of these sentences seems to involve nothing but a simple predication in which there is no “thing” of which the predication is made. Moreover, supposing that one would start with nouns and let them determine the content of the verb. When there is a patient noun, one would say that the verb is “patient attached.” But that would not be sufficient. It would still be necessary to specify whether the verb is a state or a process and this specification can only be made within the verb itself. If, on the contrary, we begin with the verb and make this specification therein, there is no need for an additional selectional unit like “patient attached.” It is enough to say that a state or a process verb is accompanied by a patient noun (Table 2). This kind of situation, pleads Chafe “…adds support to our general hypothesis that the verb and not the noun is the controlling factor” (1973, p. 166).

Van Valin (1999, 2003) has proposed an elaborate semantic theory entitled “role and reference grammar,” which he argues falls between extreme formalism at one end and radical functionalism at the other. What differentiates Van Valin’s model from the standard formalist view is the idea that grammatical structure can only be explained with reference to its semantic and communicative functions. Syntax is not autonomous. Rather, it is viewed as motivated by semantic and pragmatic factors. Van Valin’s (1999, 2003) technical terminology is partially different from that of Chafe but the guiding principles of the two analyses have much in common.

 Lexical representation of verbs in Van Valin (2003) is more detailed than the corresponding scheme in Chafe (1973) but the former categories can be regrouped in a more economical way into the latter categories. Semantic roles are extremely detailed in Van Valin’s proposal regarding the state and the action verbs (activity verbs in Van Valin’s terminology). He distinguishes a long series of thematic relations in terms of argument positions. Although cognitively accessible, it is not sure that language users can actually make use of such a complicated semantic scheme. Van Valin addresses also the issue of interclausal relations. He identifies a series of interclausal semantic relations (causative, purposive, reason, circumstances, conditional, concessive, etc.) serving to express facets of an action, state, or process, that are then organized in complex syntactic structures.

Before leaving this section, it may be justified to examine a possible objection regarding Chafe’s and corresponding semantic theories. It concerns the fact that the semantic distinction between verb and noun may be partially arbitrary. Many nouns actually can refer to states, actions, or processes and various semantic notions may be equally expressed using a verb or a noun. Chafe’s (1973) partially deals with this question with his notion of derivational unit.

The problem may be more apparent than real. One has to keep in mind that in semantic theories, the words referring to syntactic classes are only utilized for convenience. We are so accustomed to their use when speaking of language that employing a corresponding semantic-relational vocabulary would probably appear cumbersome to a reader. In Tables 1 and 2, above, as well as in the accompanying commentaries, and actually in the whole chapter, the terms relative to formal linguistic classes could be replaced by corresponding semantic expressions, which would eliminate the objection formulated. For example, instead of saying that the sentence *The wood is dry* (Table 2) involves an article, a noun, a copular verb, and an adjective (attribute), one could say that the sequence in question expresses a relation of the type “definite patient in a (particular) state”; and correspondingly for the other examples and semantic considerations.

 My objective in this chapter was not to compile the material for a complete and definitive semantic-relational theory. None of the authors mentioned expressed this kind of pretension. They all are keen to indicate that their contribution is tentative and that such a complicated matter still needs a lot of work.

 **5.2. Generality of semantic relations**

 A crucial question regarding the semantic-relational component of language, particularly in the view that it may be directly related to morphosyntactic patterning, concerns its generality. Are the semantic-relational categories universal?

 Brown (1986) suggests that lexical semantics at a deeper level is common to all languages. They all organize the plant and animal species, and also artifacts, into vertical hierarchies with several levels in terms of class inclusion (hyponyms, hypernyms, and subcategories). Among the levels, one has a particular psychological relevance, the “basic object level,” which is the level with the highest degree of cue validity (Rosch, 1978). The categories at the basic object level have maximal similarity among instances within the category and maximal distinctiveness between categories. For example, members of a basic level category like *dog* have a number of features in common that distinguish them from other categories like *bird* or *fish*. The three categories, *dog*, *bird*, ad *fish*, are subordinate members of the superordinate (hypernymic) category *animal*.

Bickerton (2009) argues that the theta-grid of verbs (i.e., the list of their theta-roles or their argument structure) is universal. It is widely agreed that the semantic relations, roles, etc., are based on general cognitive notions, which does not mean that all concepts are necessarily turned into semantic notions (already James, 1892). Major concepts for describing the world and communicating correspond across the various idioms and dialects; otherwise no translation from one language to another would succeed.

 There are indications of limited variation in the interface between cognition and language between cultures, however. For example, in a number of languages the cognitive notion of plurality is subdivided into duality (referring to two entities) and plurality (more than two entities). Arabic as well as modern English, and other languages, has only one form of personal address in dyadic exchanges. That is, no distinction is entertained according to degree of acquaintance between interlocutors (even if Arabic may use the interjection *yâ* placed before the name of the addressee to express respect). Other languages like French have two pronominal (and corresponding verbal) forms distinguishing formal (second person plural) and informal (second person singular) addresses. Spanish uses the second person singular or plural for informal addresses and the third person singular (*uste*) or plural (*ustedes*) for formal addresses. German has a system close to Spanish (Sie, the third person plural pronoun is used in formal addresses). Standard Italian offers three possibilities: second person singular or plural pronoun for informal addresses, second person plural pronoun for formal addresses to singular persons (*voi*), and third person feminine singular pronoun (*lei, ella*) for even more formal addresses[[26]](#footnote-26).

Many languages express gender (natural or conventional) in a dichotomic way (an entity is considered to be either masculine or feminine). Many other languages have a third category, neutral, regrouping entities that cannot be considered to belong either to the masculine or the feminine gender. Some languages like Indian Hopi differentiate grammatically iterative from punctual actions or processes, statements of facts vs. hypotheses, direct from reported testimonials and remembrances. The Onondaga (Iroquois) language (Chafe, 1970, 1973) inflects verbs as descriptive, iterative, punctual (i.e., an event occurring only once), or purposive (only if the verbs refers to an action). But, and this is a key point, other languages can express the same pragmatic-semantic indications using paraphrases.

As indicated in Chapter 1, Everett (2008) argues for a cultural feature in the Pirahan Indians called “the immediacy of experience principle,” that may explain why these people do not produce recursive structures in their language while clearly demonstrating in their thinking that they are perfectly able to deal with conceptual relationships inserted into one another.

Other examples could be given. Regarding color names (Welch, for example, has only one term - *glas* - for *blue* and *green*) or kinship. Hungarian crosses two semantic features - sex and rank in the family- to refer to brothers and sisters. This yields two lexical terms (*younger*-*older-brother*; *younger*-*older-sister*) plus a generic term for designating *brother* or *sister* (Peterfalvi, 1970), whereas languages like English or French have only two corresponding lexical terms. Again in Welch and Hungarian, it is always possible to express corresponding semantic notions using paraphrases as in the above translation of the kinship terms from Hungarian into English, or in Welch to characterize a *glas* color as being from a particular composed tint.

 The above examples have in common that they concern peripheral aspects of language and are mostly related to the lexicon. In contrast, it would appear that there is no clear data showing that basic semantic relations are not shared by all languages.

The anthropologist Benjamin Worf (1956; inspired by his studies of American Indian languages) and (before) the American linguist Edward Sapir (1949)[[27]](#footnote-27) have proposed a linguistic relativity hypothesis (an idea already found in von Humboldt’s writings; Trabant, 1995). In its strong form, it suggests that each language represents and divides reality in particular ways through specific lexical and grammatical structures. Varying representations, it is argued, impinge upon people’s perception and thought. This is a version of the long-rejected belief in psychology that thought is language. To put it simply, the correct view is that language is based on concepts (thought) and not the other way around. More specifically, the linguistic relativity hypothesis also carries the idea that basic semantic aspects may vary substantially between languages, which very likely is not the case as indicated.

 Over the decennia, no empirical support has been registered for the Worf-Sapir hypothesis and contrary findings have accumulated. Worf (1956) claimed that Hopi language contained no words, grammatical forms or expressions referring directly to time (past, future, present, enduring). However, Malotki (1983) has carefully documented the production of tenses, words for units of time (days, parts of days, tomorrow, yesterday, etc.) in Hopi language. Their culture keeps records of time and events, using various devices (e.g., notted calendar strings). Worf (1956) also reported that the Eskimos have at least seven words (and probably more) for different varieties of snow. In some further reports, the count inflated to dozens of specific terms (Martin, 1986). As Pinker (1994) notes, seven or eight words for snow and snowy conditions in Eskimo languages would not be different from Standard English vocabulary.

In a weaker form, the Worf-Sapir hypothesis has been restated in saying that some limited effect of lexical “codability” may influence perception to some extent. Brown and Lenneberg (1954) had observed that colors corresponding to clear lexical categories in a language (hence having a better codability in that language) are recognized better and more rapidly than less codable colors. For example, a composed blue-green tint corresponding to the single word *glas* in Welch is recognized better by Welch-speaking than English-speaking subjects for whom this tint corresponds to two distinct lexical categories (*blue* and *green*). This study has been criticized as based on a wrong assumption, i.e., that the way languages cut up the color domain is arbitrary. As is known, the physiology of vision is standard across mankind (Berlin & Kay, 1969; see Lucy, 1997, and Saunders, 2000, however, for partially contrasted positions). It would be surprising indeed if language could reorganize the whole visual system in markedly different ways.

Carrol and Casagrande (1958) had reported that the ways in which English-, Navajo-, and Hopi-speaking subjects match and classify images, is influenced by their respective languages This experiment has also been criticized for conceptual and methodological reasons (Rosch, 1974).

Brown (1986) and Pinker (1994) have reviewed the relevant experimental literature mostly testing the weak version of the linguistic relativity hypothesis.

 They have both concluded that clear evidence for the truth of this hypothesis does not exist.

 More recently, Lucy (1992a, b, 1997), Robertson et al. (2004, 2005), and Robertson, Pak, and Hanley (2008) have presented new data and argued in favor a neo-Worfian hypothesis (a reformulation close to the weak form; see also Evans, 2014). They claim that divergent lexical systems (regarding, for example, color categories, grammatical gender, or time and space dimensions) influence speakers’ perception and representation of significant aspects of the physical and social environment. Pinker (2007), however, rejects this reformulation of the old problem arguing that the neo-Worfians have not demonstrated that the variety of language spoken can structure cognition in different ways. The fact that some languages may express concepts in ways that vary from other languages does not prove that the basic system of perception and thought necessarily differs in the respective speakers.

 CHAPTER SIX

 TWO MORPHOSYNTACTIC SYSTEMS

A number of largely convergent models of language production have been proposed, which have many important points in common (Fromkin, 1971, 1993; Garrett, 1982, 1988, 1995; Levelt, 1989, 1999; Bock & Levelt, 1994; Caramazza, 1997; Hagoort & Levelt, 2009; Peterson & Savoy, 1998; Pickering & Garrod, 2013; to name but a few; see Harley, 2014, for a synthesis).

Earlier models are based on recordings of speech errors (phonemic feature exchange, anticipation or perseveration, phoneme deletion, affix deletion, morpheme exchange, word blend exchange or substitution, phrase blend, and word order errors), spoonerisms, transcripts of spontaneous speech, data from experimental studies of sentence production (mostly with response latency as dependent variable), hesitations and other pausal phenomena, neuropsychological data from aphasia, and even Freud’s (1901) reported slips of the tongue (Ellis, 1980). More recent models incorporate data from the electrophysiological and neuroimaging literature (see below).

 Language production may be divided into three phases: conceptualization, formulation, and encoding. Speakers first conceive an intention and select relevant information from the environment or from memory. This stage yields a preverbal message in the sense that lexical representation is not activated yet. Formulation involves two major steps: the selection of individual words (lexicalization) and their organization in sentence form (syntactization). Phonological encoding, then, turns selected lexemes into sounds that are realized phonetically.

 Lexicalization is analyzed as a two-step process. In a first phase, speaker moves from semantic level to an intermediate level of word representation called lemma. According to Levelt (1989), lemmas are defined semantically and also syntactically (the grammatical roles that the future words can play in the future sentences). It could be, however, that the syntactic characteristics of lemmas are only activated in a syntagmatic context (and not in simple denotative tasks, for example; Pechmann, Zerbst, & Garrett, 2000; Vigliocco, Vinson, & Siri, 2005). In a second phase, the phonological form of the lemmas is retrieved from semantic memory and phonological encoding takes place. Although the issue is still discussed, it would seem that lexical selection and phonological encoding overlap to some extent.

In Garrett’s model, formulation proceeds through a series of discrete levels. At a first level, word order is not explicitly represented but the semantic content of the words selected is set in correspondence with syntactic roles (e.g., subject, objet, determiner).Then a full syntactic frame for the future sentence is generated. The seriality of Garrett’s model has been criticized. There is evidence that the levels of processing actually interact; for example, the statistical availability of a word affects syntactic planning (Bock, 1982).

In Levelt’s (1999) model, one moves from grammatical encoding during the lemmas phase onto surface structure where a morphological-phonological encoding takes place before phonetical encoding and articulation. There a proviso in the model for self-monitoring at various levels of the processing (particularly at the levels of the phonological word and the outgoing sound wave); the speaker comparing what is intended with its ongoing realization.

 Brain studies confirm the sequential processing of lexical, grammatical, and phonological information, as suggested in language production models. For example, Sahin et al. (2009) used intra-cranial electrophysiology (ICE) from unaffected brain tissue to record local field potentials from populations of neurons in Broca’s area in three English-speaking adult patients treated for epilepsy (aged from 38 to 51 years). They had to read silently words (nouns and verbs) verbatim or grammatically inflect them (singular/plural; present/past tense). ICE is known for its good temporal and spatial resolution. Neighboring probes within Broca’s area revealed the following activation sequence: lexical identification (after approximately 200 milliseconds - ms), grammatical (320ms), and phonological (450ms) aspects of the words. The temporal pattern of activation was identical for nouns and verbs suggesting that words from different lexical classes are inflected according to a common brain process. A control with the same patients and the same task using fMRI showed activation in the same brain area and in the same way as for 18 healthy adults. Principe et al. (2017) using electroencephalography with focal epileptic patients in a naming task, observed that the posterior-temporal region of the left-cerebral hemisphere is involved very early in the verbal responses (as soon as 100ms) after exposure to the pictures.

The available language production models tell us a great deal about the stages of syntactic planning, but they say little concrete regarding the syntactic processes themselves. I have suggested (Rondal, 2014) a production model based on two separate systems (semantic relational and formal morphosyntactic) upon which I would like to elaborate in the present opus.

 Fig. 6.1 sketches the major components of this model. Feedback loops between components have not been represented in the figure. There is still considerable uncertainty in the literature in their respect.

 As is obvious, language is a communication system. Communication is the transfer of information by signals or signs purporting to have an effect on the current or future behavior or state of mind of a conspecific. Pragmatics derives from the Greek *pragma*, meaning action, and in particular a course of action involved in human affairs. In classical Latin, *pragmaticus* means a business man.

 An ideation in the speaker under the form of a communicative intention is present at the onset of language production. Humans have an ability, probably innate (which does not exclude an epigenetic component) to understand that other humans have minds with mental states and intentions of the same kinds as those experienced by the self. Any communicative intention presupposes a “theory of mind” of the kind. Concretely, in a few hundreds of milliseconds, the communicative intention activates a pragmatic indexation and a particular semantic matrix composed of semantic relations and lemmas.

Pragmatic indexation specifies the objective of the message and the contrast between old and new information. The former refers to the information that the speaker can reasonably hold to be known by the addressee or that can be recovered from linguistic or extralinguistic context. This old-new contrast controls the use of the ellipsis and the emphasis (prosodic or syntactic, e.g., the use of the passive voice instead of the active). It also plays a role at the post-semantic level, i.e., the conversion of the semantic-relational deep structure into surface structure (see below).

Pragmatic indexation also determines the illocutionary type (declaration or request) of the utterance, person or participant deixis (distinction between first, second, or third person, or speaker, recipient, and “bystander”, respectively), optionally social deixis (e.g., polite forms), time and place deixis (e.g., proximal vs. distal reference), polarity (positive vs. negative), aspect (e.g., completion, duration, frequency of a given action, state or process), and mood (expressing probability, usualness, obligation, presumption, plausibility, degree, intensity, conditionality, obligation, permission, prohibition, exemption). Pragmatic indexation interacts with semantic relations as will be seen below.

 **6.1. Semantic-relational system**

As indicated in Chapter 5, semantic relations are not naturally ordered linearly. It is reasonable to suppose that following hours of exposure to a given input, language learners come to order their semantic relations according to the sequential regularities of the particular language they have been exposed to. In this sense, the semantic relations described in the preceding chapter end up organized in semantic-relational sequences (SRSs). Correspondingly, the lemmas are endowed with syntactic properties. I postulate a direct translation of these ordered semantic-relational relations in surface morphosyntactic patterns (SMSPs) through appropriate lexicalization. Sentence meaning, however, may still be modified in surface structure.

 The mechanism is documented, first with respect to syntactic structures, and, then regarding grammatical morphology.

 **6.1.1. Semantic-syntactic patterning**

Starting with a SRS definite reference+qualitative attribution, in languages like English, French, or Italian, one produces sentences of the type *Coffie is ready*; *Le café est prêt*; *Il caffè è pronto*. More frequent, in Italian is the corresponding utterance *E’ pronto il caffè* (literal gloss: *Is ready the coffee*). This exemplifies the double fact that semantic relations are ordered sequentially with respect to the dominant sequential patterns in the particular language and that they may correspond to two (or more) syntagmatic sequences with identical meaning. This is not always the case, however. Different word order in related expressions may correspond to different semantic relations. In French, the placement of the adjective with respect to the qualified entity may signal a difference in meaning. For example, the two phrases *un grand homme* (*a big man*) and *un homme grand* (*a man big*) do not have the same relational meaning. The former refers to a man of high reputation and achievement; the latter to a man of high physical stature. Usually, the anteposition of the adjective with respect to the noun suggests a more abstract attribution.

In an active declarative affirmative sentence of the type *The neighbor has borrowed my car to go to the maternity clinic with his wife*, the complex SRS is as follows: agent (definite reference) + action-process +patient (possessive reference) + goal (action-direction with definite reference) +accompaniment (possessive reference). Additional phrase or sentence regulations (including grammatical morphology; see below) are enforced during sentence expression.

 The pragmatic contrast between active and passive sentences is reveaing. Consider, for example, the two following sentences: (1) *My neighbor has borrowed my car*; (2) *My car has been borrowed by my neighbor*. The underlying relational meaning is the same in the two sentences. It corresponds to the semantic structure agent + action-process + patient. Most of the time, for pragmatic reason (emphasis on the patient), this structure is realized in surface as patient + action-process + agent (the agent can be omitted if it is easily retrievable from context or from previous information, which yields a so-called truncated passive). A declarative sentence beginning with the patient determines the production of the morphosyntactic marks of the passive, i.e., in English, Aux-be + past-participle on the main verb + agentive preposition.

 In a complex sentence with a subordinate completive, for example, *John said (that) the matter was not urgent*, the semantic-relational structure in the main clause is as follows: agent + action-process + completive.

The completive clause is of the semantic-relational type entity-attribution (disregarding temporal marking) with a pragmatic indication of negativity. In surface, the relationship between main and completive clauses is expressed (optionally) by the conjunctive element *that*.

The case of prepositions and articles is interesting to evoke in the present theoretical context. They may be specified through the grammatical relation that they encode between other lexemes. But, foremost, they express particular semantic relations. For example, preposition *from* refers to origin, starting-point, source, motive, etc. *To* indicates direction, comparison, recipient or possessor, person or thing affected by an action, etc.; and so on for the other members of this lexical class.

Articles in English specify definite or indefinite reference of the noun they accompany. In French, they also express a contrast between masculine and feminine reference (singular reference, *un*, *une*, *le*, *la*). Dutch distinguishes between definite masculine or feminine (*de*), and neutral (*het*, only for singular references). In Arabic, there is only one article: *al*. The definite-indefinite contrast is realized through the expression-nonexpression of the article. For instance, *bint Jamila* means *a beautiful girl* and *al bint al Jamila*, *the beautiful girl* (i.e., repeating the article in front of the adjective).

Speakers may rely equally (or more) on the semantic relation encoded by prepositions or articles when ordering their utterances than on grammatical definitions. The same can probably be said of coordinate and subordinate conjunctions in dealing with the sequential organization of clauses in compound and complex sentences. Coordinate conjunctions encode semantic relations of equality and union (*and*), opposition (*or*), negation (*nor*), consequence (*then*), etc. Subordinate conjunctions introduce hierarchic relationships between clauses (regarding time, causality, comparison, condition, etc.).

 Pronouns exist in all languages. They allow avoiding the repetition of a coreferring noun at no cost in communication providing that a link is maintained between pronoun and noun. The link may be anaphorical in discourse or retrieved from extralinguistc context. Each lexical sub-class of pronouns (personal, possessive, demonstrative, relatives, and interrogatives) has its own semantic and syntactic characteristics.

The example of personal pronouns may be utilized to illustrate the important role of relational semantics in pronominal syntax. Pronouns referring to an agent, a state, or the inanimate entity responsible of an action-process (e.g., in English, *I*, *you*, *she* or *he*, *we*, *they*) are placed before or after the verbal element depending on the pragmatic status of the sentence (declarative, interrogative, or imperative).

 Patient pronouns (e.g., *me*, *you*, *it*, *us*, *them*) are placed after the verbal element and may be introduced with a preposition. Other pronouns yet are placed either before or following the verbal element depending on their semantic relation with this element (complement, benefit, instrument, accompaniment, etc.).

There is a second major reason for postulating an important functional role of semantic relations in sentence treatment. It relates to the sentence discontinuous constituents. Let us consider central embedding in clause recursion, for example. The tree-diagram in Figure 6.2 represents the traditional structural syntactic analysis of the complex sentence *The man who was going to the station carried a heavy luggage* (disregarding tense and aspect marking on the verbs).

Fig. 6.2. Tree-diagram of the sentence *The man who was going to the station carried a heavy luggage*.

The embedded clause *who was going to the station* belongs to the first noun phrase of the main clause and for that reason is located lower in the hierarchical diagram than the main clause (*The man carried a heavy luggage*). S stands for sentence (or clause), NP symbolizes noun phrase, VP verb phrase, N noun, V verb, A article, Pro pronoun, Adj adjective, Prep preposition, PP prepositional phrase, Aux auxiliary, and Prt participle.

A relevant question is whether one needs to have in mind a deep hierarchical diagram of the kind illustrated in Figure 6.2 to produce and understand sentences with centrally embedded clauses.

There is evidence that learning nonadjacent dependencies (spatially or temporally discontinuous stimuli) is possible under specific conditions, even if such a learning is more difficult than learning adjacent dependencies (Perruchet & Pacton, 2006; Perruchet, Poulin-Charronnat, & Pacton, 2012).

 The identification of a relationship between elements A and C in a sequence of the AXC type is facilitated: (1) when there is a high level of similarity between A and C; (2) when the two elements also appear as isolated stimuli in the material presented; (3) when element X varies, which incites the subjects to pay particular attention to the constant elements in the display; (4) when A and C belong to the same category of items, differing from the category of X; and (5) when there are short pauses within the AXC sequence. It is also recommended to introduce complexity progressively in the learning episode (Lai & Poletiek, 2011).

 In opposition, double central embeddings, for example, sequences of the type ACXDB (where the relations are between A and B, and C and D), and triple embeddings (ACEXFDB) are markedly more difficult or even impossible to learn in a context of unintentional learning.

 Double central embeddings, although grammatical, are virtually inexistent in natural language use (for example, *The man that the police inspector that his superior admired has arrested carried a luggage suspect*). The obvious reason is the important additional constraint on working memory and the concomitant availability of procedurally simpler ways for expressing the same contents (through juxtaposition or coordination of clauses).

 As indicated in Chapter 3, Fitch and Martins (2014), have proposed a formal (hierarchical) syntactic buffer as a contributing solution to the problem of momentarily storing morphosyntactic information while performing some other operation in complex sentences such as those containing centrally embedded clauses.

 Semantic relations may be instrumental in this respect. They may play a facilitating role in treating complex sentences, in particular sentences with embedded clauses, and probably with linguistic recursion in general. The sentence in Figure 6.2 above involves two semantic relations: agent + action-process + patient (*The man carried a heavy luggage*); agent + action + complement (direction in this case) (*The man was going to the station*). In surface, the second one is integrated in the first one in what is called central embedding. The first clause is interrupted and the semantic relation involved is momentarily stored in working memory, acting as a semantic buffer and conceptual “splint,” the time needed to embed the second clause, before resuming the course of the first one. Caplan (1992, 1995) has documented a sequential span of at least a few words in regular monitoring of utterances.

 A second operation is involved in the production of relative clauses. It concerns the selection of the relative pronoun in order to avoid repeating the same lexeme in the main and the relative clauses (*The man the man was going to the station carried a heavy luggage\**). The following examples from several languages illustrate the ubiquity of the semantic roles in this operation.

In French, the selection of the relative pronoun differs with respect to its semantic role. This applies equally to centrally embedded and relative clauses derived to the right (hand side) of the main clause. Agent-pronoun is *qui* in active relatives (e.g., *L’homme qui achetait un billet de train se rendait à Strasbourg;* gloss: *The man who was buying a train ticket was going to Strasburg*; *J’ai rencontré un homme qui allait à Strasbourg*; *I met a man who was going to Strasburg*). The same is true for pronominal coreference in the context of a process or a state verb (e.g., *Le drap de lit qui sêche est agité par le vent*; literal gloss: *The sheet that is drying is shaken by the wind*; *L’homme qui est sur le quai regarde sa montre*; *The man who is on the platform is looking at his watch*).

Patient-pronoun is *que* at the active voice and *qui* at the passive (e.g., *L’homme que j’ai rencontré portait un parapluie*; literal gloss: *The man whom I met carried an umbrella*; *L’homme qui a été arrêté transportait un colis suspect*; *The man who was arrested carried a suspicious parcel*). There is no passive relative in *que* as one may only passivize a sentence beginning with an agent. Other relative pronouns exist (e.g., *où*, *à qui*, *de qui*, *avec lequel*, *dont*, etc.; gloss: *where*, *to whom*, *of whom*, *with whom*, *whose, etc.*) involving other semantic dimensions (location, benefit, possession, accompaniment, instrument, determination, etc.).

The system of relative pronouns in English is comparable to the French one with a lexical simplification: the relative pronoun *that* instead of *who* (agent and patient with human reference), or *which* (idem as *who* but with nonhuman reference) is grammatical for all semantic correspondences except those expressing possession (*whose*), place (*where*), time (*when*), or causality (*why*). Italian uses a single relative pronoun (*che*) in correspondence with the roles of agent and patient. *Chi* is a particular form, invariable, used only for singular reference, limited to animate entities, and employed with a demonstrative function. *Chi* may be used without explicit nominal coreference (e.g., *Chi dice questo sbaglia*; gloss: *Who says that is wrong*)*.* Other pronoun substitutions exist in relation with other semantic roles (*a cui*, *il quale,* *di cui*, etc.; literal gloss: *to whom*, *he who*, *from who*, *in that*).

The Spanish system is similar to the Italian one: *que* for the agent and patient roles; and other pronominal substitutions for other semantic roles*.* In standard Arabic,a relative pronoun is used only if its coreferent is definite. The Arabic language utilizes the same pronominal form in correspondence with the semantic roles of agent and patient (the particular form is inflected for number and gender: *alladhî*, singular masculine; *alladhîna*, plural masculine; *allatî*, singular feminine; and *allâtî*, plural feminine). A particular turn of clause exists involving an affixed personal pronoun, said of recall, also inflected for number and gender (*hu*, in the following example), located at the end of the relative clause and referring to the noun at the end of the main clause when the following relative pronoun is not agentive (for example, *zurtu Tûnis maa sadîqî Murâd alladhî kallamtuka aanhu*;literal gloss: *I visited Tunis with my friend Mourad that I spoke to you of him*).

 Normative grammars deal with relative pronominalization in terms of linguistic-functional categories. In French, for example, it is stated that the relative pronoun *qui* is used as grammatical subject of the verb, and *que* as grammatical object. Such recommendations imply that the speakers have ready access to the formal hierarchies founding these categories. This raises a problem, as said, even if such a recourse cannot be excluded in mature speakers (see next section). Incidentally, the relative pronoun *qui*, in French, is not used unequivocally as grammatical subject of the verb. It is employed in a variety of prepositional phrases headed by: *à qui*, *pour qui*, *avec qui*, *contre qui*, etc.; literal gloss: *to whom*, *for whom*, *with whom*, *against whom,* etc.) endowed with various grammatical functions. The same is true for *que*, which may function as direct object of the verb but also as attribute of the subject (e.g., *Chanceux que tu es*!; literal gloss: *Lucky that you are*), time complement (*L’été qu’il fit si chaud*; *The summer that it “made” so hot*), or grammatical subject (*Fais ce que bon te semble*; *Do what seems good to you*). These indications render quite complicated the use by the speakers of the linguistic formal and functional categories in the pronominalization operation in relative clauses.

 **6.1.2. Grammatical morphology as a self-editing process**

 The inflection of formally variable lexical elements (verbs, nouns, pronouns, adjectives, and articles, in French; verbs, nouns, and pronouns, in English) and agreement (concord) between particular items in sentences, is semantically motivated (notions or gender, plurality, tense, aspect, mode, etc.; Chafe, 1973).

 In traditional linguistics (for example, Hockett, 1942), grammatical morphological marking is treated as an operation taking the stem of the word as point of departure (for example, a verb like *bark*, or *think*) and producing the intended marking (for example, *barked* or *thought*, for the past tense). According to this view, both regular are irregular forms are generated by the same grammatical subsystem, except that the generation of the irregular forms necessitates an additional procedure. Generative grammar has much followed this tradition with the idea that some rules (said to be regular) apply to an open-ended set of items (like the add-*ed*, or add-*d*,to form regular past tenses), whereas other rules (irregular) are restricted to a finite list of exceptions (see Chapter 7, however, for additional indications on this matter).

Inflectional processing may be based on grammatical or on semantic roles and relations.

It may be hypothesized that particular semantic properties of the words first produced in a sentence determine the grammatical-morphological marking of the following words either in immediate proximity or at a distance in sentence space. For example, the plurality of an agent automatically determines plural marking on the verb stem. The semantic information in the agent word is momentarily stored in a working memory buffer until production of the verb. The general mechanism may be labelled sequential self-editing. There are numerous research indications showing that morphological agreement takes place in sentence surface (e.g., Bock & Cutting, 1992, for number agreement; see Harley, 2014, for more recent data). Auxiliary fronting in English questions or other surface constructions reversing the SVO order need not be a problem here because semantic relations, as indicated, are not linearly ordered. Hence, they may be supportive of morphological self-editing in whichever way is necessary.

Grammatical morphology varies substantially between languages. In a given language, some markings are said to be regular, meaning that they apply in the same way to all concerned word stems. Others differ and have to be learned case by case.

Regular marking is most often considered to be rule-governed (for example, Pinker, 1994, 1999). Yang (2016) contends that irregular morphological marking is also governed by rules. Given the questionable logical and empirical status of the notion of rules unknown to language users, it might be safer to speak of analogies implicitly detected by the learners with or without a clear awareness of the existence of corresponding normative rules. Ramscar (2002), for example, has documented the importance of semantic and morphophonological cues, and distributional frequencies in analogical generalizations related to grammatical inflections.

 **6.2. Formal morphosyntactic system**

 As documented in Chapter 3, there exist brain representations of linguistic formal and functional categories, and perhaps of some morphosyntactic operations. These representations may play a role at the post-semantic level of sentence production.

 In every language, post-semantic processes operate to convert a semantic-relational structure into a surface one. The semantic structures need to be converted into sequential patterns as speech is inherently linear. Another post-semantic process that will not be dealt with here because it has nothing to do with the semantic dimension of language is phonetic change. It concerns the fact that lexical symbolization does not lead directly to phonetic realization. It must transit through an underlying phonological structure in which various processes operate for producing a surface phonetic structure.

 Chafe (1973) argues that the contrast between old and new information (involved at the pragmatic indexation stage) has a determining role also post-semantically. The old-new contrast has been conceptualized by other authors under the names of presupposition-assertive focus, topic-comment, background-figure, primary and secondary topicalization, or theme-rheme ; for example, Bates, 1976; Levinson, 1984; Halliday, 1985). Old information is most often presented first in a sentence as an anchor point for the new information that follows. It can also remain unexpressed when obvious or recoverable from context or from a preceding exchange of information. Every language has its own way of representing old and new information in surface structure. Word order and intonation (differences in pitch level and sound intensity) are the devices most frequently used across languages (Levinson, 1984). Some languages, however, like Japanese employ specific particles as *wa* for indexing old information and *ga* for a new one.

In predominantly SVO (subject-verb-object) languages, like English or French, the old-new contrast is reflected in the grammatical organization of the sentence. Typically, in active sentences, a surface structure subject carries the old information. The formal status involves a particular sequential placement with respect to the verb and grammatical concord with it. Chafe (1973) indicates that the noun which conveys the old information will always be the agent, the beneficiary if there is no agent, and a patient if there is neither an agent or a beneficiary. A location noun necessarily occurs with a patient noun (e.g., *The box is under the table*). In such cases, if there are only two nouns, it is the patient noun that conveys the old information and the location noun the new one. The hierarchy for old information therefore is agent, beneficiary, patient, and location, and conversely for new information.

The interesting point is that the English surface features of grammatical subject and object do not directly represent semantic relations such as agent, patient, or the like, but are connected to these notions indirectly, through the fact that these relations play a role in the sentence distribution of old and new information (Chafe, 1973, p. 232-233).

 In passive sentences, the distribution of old and new information is different from the active ones. The function of the passive voice is to change the order of priorities for the distribution of information. The patient noun of a passive verb normally conveys the old information whereas the agent noun conveys the new one.

Truncated passives (e.g., *The bow was emptied; David was seen*), as opposed to “full” passives (e.g., *The box was emptied by John; David was seen by someone)*) allow an action-process or a process-experiential verb to be expressed without an accompanying agent or experiencer. This represents the “truncated passive” exception to the principle illustrated before stating that an agent noun must be assigned to an action verb, and an experiencer to an experiential verb.

 In so-called (syntactic) ergative languages (e.g., Basque, Georgian, Kurdish, Iroquois) as opposed to nonergative ones (English, French, and most Indo-European languages), a particular form (case) is assigned to the noun subjects of transitive verbs whereas nouns that are either the subject of intransitive verbs or the objects of transitive verbs, are marked otherwise. In ergative languages, the correspondence between semantic agent and surface subject as well as between semantic patient and surface object is direct, which is not the English situation, as said. In English, a sentence characterized as intransitive is composed semantically of a patient as the only noun (e.g., *The wood dried*) or a binomial whereby the other noun is a location (e.g., *The box is under the table*). Because it is old information, the patient noun in these sentences corresponds to the surface subject.

 It follows that in ergative languages, basic aspects of sentence surface can be directly related to semantic roles and relations. There are published indications regarding two ergative languages of the Austronesian family: Acehenese, a Malayo-Polynesian language spoken by native people of Aceh (Sumatra) in Indonesia (Duric, 1985), and Tagalog, another Malayo-Polynesian language spoken in the Philippines (Schachter, 1976; Maratsos (1989), suggesting that there is no need for the notions of grammatical subject and object of the verb in accounting for the morphosyntax of these languages.

 In English and, likely in other nonergative languages, one has to take into account the distribution of old and new information for connecting the grammatical roles of subject and object with corresponding semantic roles.

It would appear that there are two systems endowed with the theoretical power of regulating morphosyntactic functioning: a semantic-relational one and a formal-grammatical one. The two systems are separate including at the level of brain substrate (even if they have proved difficult, as said, to set apart in neurological investigation). However, it is possible to pass from one system to the other as Chafe’s (1973) post-semantic analysis suggests. The two systems correspond, if only given the self-evident axiom expressed in Chapter 3: anything in natural morphosyntax that can be described using the formal and functional categories of linguistics can also be accounted for relying on a network of semantic-relational categories.

As remarked in the preceding chapter, but worth repeating, there is no guarantee that formal morphosyntactic knowledge is complete even in adult language users compared with the proviso of formal grammars, given the poverty of stimulus argument and the absence of innate morphosyntactic representations.

 It could be asked why nature does not seem to have provided human beings with the genetic infrastructure necessary to develop natural morphosyntax in the deductive way theorized in generative linguistics and psycholinguistics? Kurzwiel (2006) suggests that it is in the nature of complex biological systems that their functioning is less than optimal. It is only as good as it needs to be to function correctly. Optimality implies additional biological costs that may not be functionally necessary. A large number of fine-grained distinctions proposed in formal linguistic theories are probably not relevant to the neuropsychological approach to language treatment, which operates according to broader conceptual considerations.

Pushing the reasoning one step further, one could ask whether the formal morphosyntactic system is really instrumental in regular language users or whether it simply represents an epiphenomenon of language exposure. This is a question that would certainly sound anathema to the proponents of the generative school in linguistics and psycholinguistics. Yet, brain representation alone does not guarantee people’s actual use of corresponding linguistic categories in language treatment. Formal morphosyntactic knowledge is needed, by definition, for metalinguistic tasks. It may prove useful when there is little relevant information available from extralinguistic context and perhaps for particularly complex sentences where semantic information is not sufficiently clear. But, beyond that, the question remains: linguistic categories to what extent and for what regular functional purpose?

 CHAPTER SEVEN

 IMPLICIT MORPHOSYNTACTIC LEARNING

 A number of explanatory theories of natural morphosyntactic development have been proposed over the last 50 years or so (without going back to previous suggestions). They proceed from different theoretical backgrounds and may vary substantially.

 Several theoretical account are in direct line with generative grammar (and its successive versions) and the hypothesis of representational innateness: for example, McNeil’s (1966, 1971), Pinker (1984, 1985, 1994), and more recently Yang (2016).

 Other theoretical accounts are more on the learning side: either in line with more traditional learning theories, for example, Staats (1971, 1974), or proposing a more open learning approach integrating semantic and syntactic features and stressing the active role of the child as a “procedure discoverer” (Braine, 1971, 1976; Stemmer, 1973; Crystal, 1987), or even an elaborated parental- teaching/child learning paradigm (Moerk, 1983, 1992, 2000).

 Other contributions yet stress the determining role of cognitive structures in morphosyntactic development: for example, Sinclair (1971, 1973), Brown (1973), Slobin (1973), Bruner (1975a, 1975b), Bates (1976), Karmiloff-Smith (1979), Bates and MacWhinney (1982), Schlesinger (1988), and Tomasello (2003).

 For partial syntheses of this extensive literature, one may want to see Rondal (2006), Linden (2007), Gleason and Ratner (2009), Saxton (2010), Ambridge and Lieven (2011), and Harley (2014).

 Current theories while projecting many interesting insights on morphosyntactic development, fail to account for the implicit character of this development; quite obvious as it would seem given that language users have no clear conscious knowledge of the morphosyntactic mechanisms that they activate when producing sentences. The implicit character of morphosyntactic knowledge is mentioned passingly, when it is mentioned at all. No theoretical treatment or precise model of it has been suggested.

 In my view, and as will be explained in what follows, both the semantic-relational and the formal-syntactic bases of morphosyntactic functioning are learned implicitly. One needs to account theoretically for this learning. There are enough data in the huge language developmental literature to ground a precise modelling exercise.

 It likely takes the child additional developmental time for learning the syntactic information that one finds represented in adults’brain. This implies, as suggested, that the first developmental system underlying combinative language cannot be formal but has instead to be based on relational semantics. And if, as suggested also, the former one is never complete in regular language users, this means that the semantic-relational system may (must) remain the dominant one all life on.

 **7.1. Conceptualizing implicit learning**

A number of research works in the field of implicit learning have demonstrated the reality of unintentional learning; Implicit learning may be defined as: “A mode of adaptation in which the behavior of a subject shows sensitivity to the structure of a situation when this adaptation cannot be attributed to an intentional exploitation of an explicit knowledge of this structure (Perruchet & Nicolas, 1998, p. 15; my translation). This kind of learning, at variance with explicit learning, is often described as autonomous with respect to conscious inspection and associated with incidental learning conditions (Meulemans, 1998).

Several of the definitional characteristics of implicit learning are still the object of discussion, for example, the question to decide whether there is a need to keep implicit learning completely separated from explicit learning or whether both types of learning are present in variable proportion in all learning tasks. What is meant by the expression “implicit learning” may vary between authors. For Reber (1993), for example, implicit is synonymous with unconscious. Other authors conceptualize the notion of implicitness as referring overall to unintentional processes yielding a knowledge that cannot be verbalized (for example, Perruchet & Vinter, 1998). Paradis (2004) defines implicit learning as determining a competence acquired incidentally, stored implicitly, and used automatically. He drastically separates implicit from explicit knowledge. They are qualitatively distinct entities relying upon separate neural substrates (see below). No matter how systematic or repeated, explicit knowledge never becomes implicit knowledge. It is restricted to the outputs of behavioral adaptation. The learner may be aware of the external aspects of his behavior not of the processes underlying it. Conversely, implicit competence relates to underlying organizational processes of behavior and it never becomes explicit knowledge.

Another issue concerns the relative robustness of implicit learning and memory according to age (children, adults, and aging persons) and intellectual ability. Reber (1993) argues in favor of an invariance in the ability for implicit learning across ages and intellectual levels. Fletcher, Mayberry, and Bennet (2000), and Witt and Vinter (2013a) are less assertive. Implicit learning could correspond to a learning mechanism evolutively more ancient and more robust than explicit learning (quite sensitive to age and intellectual level, as is known). This is the position defended by Reber among others. In opposition, there are works the results of which seem to support the hypothesis of noteworthy differences in people’s implicit learning ability according to age; possibly tied to the maturation of brain structures underlying implicit learning (Thomas et al., 2004) and to the concomitant development of conceptual knowledge in children when this knowledge is relevant to the situation, context, and/or the material proposed for learning (Murphy, McKone, & Slee, 2003).

Since the pioneering work of Reber (1967), one of the experimental paradigms often used in implicit learning research is learning a finite-state artificial grammar. An automaton produces linear sequences of items (for example, quadruplets of consonants presented visually) following particular transitional rules. Participants are not informed of the existence of such rules. In the learning phase, they are requested to memorize the sequences presented. In the test phase, they are informed of the existence of transitional rules but not of the exact nature of these rules. They are then invited to sort out new sequences of consonants according to whether they are grammatical (i.e., conforming to the transitions they have been exposed to in the learning phase) or not. An interview of the participants is often carried out at the end of the experimental session in order to verify whether they have developed a conscious knowledge of the rules and are able to verbalize them.

Results generally show that most subjects can differentiate the grammatical sequences from the ungrammatical ones much beyond chance level, as if they had discovered the transitional rules involved. However, they are unable to account for these rules verbally.

Reber (1967, 1969, 1989, 1993) explains this implicit form of learning as demonstrating a unconscious process of rule abstraction. He sees a proof for it in the fact that participants are able to judge the grammaticality of sequences of items to which they have not been exposed. According to Reber, this observation excludes the possibility of mere learning of superficial characteristics of the stimuli during the learning phase of the experiment.

 Reber’s explanation has been contested and seems mostly abandoned nowadays. It can be, however, that participants could discover the transitional rules in the most simple cases but be restricted to an implicit analysis of the most salient stimulus properties (perceptual, positional, and/or statistical) in more complicated cases (Witt & Vinter, 2012).

 Other interpretations have been proposed. According to Brooks (1978) ad Vokey and Brooks (1992), there is no process of rule abstraction in implicit learning of artificial grammars, only memorization of entire items during the learning phase. In the test phase, subjects compare the new items with those already memorized. The grammaticality judgements are based on the similarity/non-similarity of the sequences proposed in the learning and test phases.

 This later interpretation has been criticized in terms of insufficient parsimony and problematic attentional resources. Perruchet (1994) has shown experimentally that subjects do not actually memorize entire items, only fragments of items; for example, bigrams or trigrams of consonants located at the beginning or the end of frequently repeated sequences. There is a tendency for the isolated fragments (labelled “chunks”) to increase in size with exposure to the material displayed until providing a correct syntax of this material.

 A complementary explanation has been proposed by Perruchet and Vinter (2002; see also Perruchet, Vinter, & Pacton, 2007). They propose that chunks isolated from the material and determining mental representations isomorphous with this material may access consciousness and are a valid substitute for rule knowledge. The mental representations are the product of elementary associative operations. The repeated simultaneous exposure to two initially independent events leads naturally to the formation of a chunk composed of the two units. There is no need of an intention to learn or something of the kind to account for this phenomenon; joining the two events within the same attentional focus is enough to guarantee chunking.

 As alluded to above, Paradis (2004) is not convinced that implicit knowledge may ever access consciousness. Learners may become aware of surface aspects of their behavioral adaptations but not of the underlying processes, mechanisms, and representations, which stored in procedural memory (see below), cannot be accessed consciously.

Implicit learning is learning without intention to learn but not without attention, and associative learning can be specified as an automatic process linking together the task components present on the scene and captured in the attentional focus of the learner (Perruchet & Poulin-Charronnat, 2015, who recall the “principle of belonging,” posited by Thorndike, 1932, to account for the creation of an associative link between two stimuli).

Several experiments demonstrate the necessity of attention in implicit learning. For example, Stadler (1995) has shown that adding a cognitive task sufficiently compelling to the implicit learning of another task reduces the quantity and the quality of the latter.

 Most of the literature on associative learning has been concerned with adjacent dependencies. As introduced in Chapter 6, a sophisticated cognitive function as language involves also dealing with nonadjacent dependencies (for instance, a meaning indication of plurality introduced at the beginning of a sentence in the noun acting as grammatical subject is passed over to the verb which may be placed at a distance in the sequence). Pacton and Perruchet (2008) have observed that when participants are confronted with tasks involving relations between both adjacent and nonadjacent elements, they exclusively learn the type of relation that they are requested to process irrespective of the distance between the two kinds of element in the stimulus material. The joint attentional processing of two events whether adjacent or nonadjacent is a sufficient condition for learning the relation between them. Pacton and Perruchet (2008) add that the well-documented contiguity effect in associative learning could be a by-product of the fact that attention generally focuses on contiguous events. However, even if the attentional focus at a given moment in a learning task has a higher probability to capture events that that are close on spatial or temporal dimensions, they do not need to be strictly contiguous (also Gomez, 2002). Attentional focus can also encompass nonadjacent events even if it is true that the tendency to associate nonadjacent items in a sequence decreases as a function of the distance between them (Cleeremans, 1993).

 Pacton, Sobaco, and Perruchet (2015) have conducted another experiment in which participants were simply requested to read aloud sequences of digits or syllables (i.e., without intentional processing of specific elements of the stimuli). Only adjacent dependencies were learned. It was observed that the experimental subjects were highly sensitive to violations involving the spacing between paired elements. Importantly, Pacton et al. (2015) propose: (1) that the presence or absence of an intermediate event between paired elements not only plays a role in the mental representation of an association but is an intrinsic component (not just “an obstacle”) in the association between nonadjacent elements; and (2) that nonadjacent dependency learning incorporates the possible presence of variable intervening events either as a proper component of attentional coding or as an automatic by-product of the attentional processing of neighboring elements (p. 198).

 It is also important to keep in mind that implicit learning can handle abstract events. Associative processes operate also on abstract entities and the initial chunks in implicit learning tasks serve as a basis for further abstraction (Perruchet & Poulin-Charronnat, 2015). This means that complex representations can enter into associative links without rules-based operations. Williams and Kuribara (2008; see also Williams & Rebuschat, 2012) exposed native English speakers to mixed artificial language composed of English words organized in sentences according to Japanese syntax (unknown to the participants). Results showed that they were able to learn some aspects of the canonical word order of Japanese. During the test phase, participants accepted as correctly formed sequences containing different words from the training phase; an indication that some abstract learning had taken place. However, they could not generalize their beginning syntactic knowledge to more complex sentences suggesting that no rule induction was involved.

 **7.2. Implicit learning of natural morphosyntax**

Besides general context, there are several differences between implicit learning of an artificial grammar in experimental conditions and morphosyntactic learning of natural language in natural conditions.

They are: (1) the interpersonal context of natural language learning involving the lexical, content, and morphosyntactic adaptations characteristic of verbal interactions between adults and language-learning children (see below) as opposed to the experimental context of artificial grammar learning; (2) the particular meaning and communicative value of these interactions; (3) the absence of specific instruction regarding natural language learning and specific intention to learn something in children ; and (4) the time dimension: a couple of hours for learning an artificial grammar against thousands of hours of verbal interaction in the course of several years for natural language acquisition.

There are also interesting similarities between the context of artificial grammar learning and natural morphosyntactic development: (1) both relate to procedural learning (i.e., learning to grasp a particular surface organization of the stimulus material); (2) child in language learning like participant in artificial grammar learning knows nothing of the syntactic disposal (s)he is supposed to learn; (3) child also ignores that (s)he is placed in a situation of morphosyntactic learning; (4) a degree of selective attention is necessary in both learning contexts without which, theoreticians are keen to indicate, there can be no learning as implicit learning cannot be assimilated to a passive impregnation from particular properties of the environment ; and (5) in both learning contexts, the procedures provide an important number of repetitions of particular aspects of the material to learn, determining noticeable differences in the relative familiarity with certain elements of the input in comparison with other non- or less-relevant aspects.

 In what follows, an implicit learning model of natural morphosyntax is proposed and justified empirically.

 **7.2.1. Syntactic learning**

 Perruchet and Nicolas (1998) have suggested that early syntactic acquisition in natural language is achieved through implicit learning of sequences from the input.

 A noteworthy capacity for sequential implicit learning and memorization is indeed present in babies. For example, Saffran, Aslin, and Newport (1996) have exposed eight-month-old infants to a continued flow of artificial speech composed of “words” of three syllables without meaning, repeated in random order during two minutes (e.g., *padotibidakutupiropadoti*). The babies proved able to differentiate new words (not presented before) from more familiar ones relying on syllabic order (controlled behavioral evidence: gaze direction and head orientation toward the source of new sounds). In a second phase, the experimenters exposed the infants to another task consisting in differentiating words already learned from “nonwords”, i.e., sequences composed of three syllables but extracted from the end of one word and the beginning of a following one (for example, *tibida*). The babies had been exposed to these syllables but diversely distributed. They again proved able to differentiate nonwords from familiar ones.

The robustness of the sequential capacity is attested in a research work conducted by Cashon et al. (2016) with children with Williams syndrome, aged between 8 and 20 months, and using the same methodology.

 Marcus et al. (1999) have presented 7-month-old babies with triplets of syllables without meaning sequenced ABB (e.g., *gatiti* or *linana*). The infants were then exposed either to corresponding triplets or triplets sequenced ABA (e.g., *wofewo*). They demonstrated a statistically significant tendency to pay more attention to the new sequences. Marcus et al. (1999) interpreted their data as attesting rule learning in young babies, which is not warranted. Instead, these data appear to document a remarkable sequential sensitivity in babies as young as 7 months.

Gomez and Gerken (1999, 2000) have presented nonsensical one-syllable words for a few dozens of seconds to children aged 12 months (e.g., *vot*, *pel*, *pic*, *rud*, *tam*). The sequences began and ended with the same words but word order in the sequences varied systematically. At the test phase, new sequences were presented. The infants exhibited a clear preference for the unfamiliar sequences demonstrating sensitivity to relatively complex sequential patterns.

 Kalouguina and Shi (2013) have exposed French-speaking children aged 11 and 14 months to utterances in Russian composed of three words whose syllables were sequenced ABC immediately transformed into BAC. During the test phase, the same children were presented with new sequences either corresponding to those learned before or to new sequences characterized by a different transformation (i.e., from ABC to ACB). Only the positional information of the syllables allowed differentiating the sequences. The 14-year-old but not the 11-year-old children could identify the similarities and differences in the stimuli beyond chance level.

 These abilities appear to generalize across voice differences. For example, Graf Estes (2012) examined word segmentation in 11-month-olds and found that infants can recognize segmented words across a change in the speaker’s voice (although less easily than recognizing them in a consistent voice).

 Perruchet and Poulin-Charronnat (2015) argue that children, equipped with a basic sequential ability, extract from the language input and memorize unanalyzed sequences of adjacent words; small ones at the beginning, then longer ones in recombining sequences already mastered. The sequences extracted from parental input are meaningful in the communicative context, pragmatically relevant, and fairly accessible to the cognitive level of the children. They are clearly set in evidence at the beginning or the end of the utterances, frequently repeated (Bannard & Mattews, 2008), and correspond to perceptually salient features of the input. Parental utterances to younger children are also clearly intonated. The final syllables are elongated, which slows down the rate of speech and helps the child to segment the discourse into manageable units. Pitch is elevated and contributes to keeping attention focused on the language exchange.

 Learning is explained by the intervention of general mechanisms operating unconsciously (MacWhinney, 2010) and establishing associations between contiguous occurrences of behavioral events (Estes, 1970). The learners segment the input in sequences of primitive elements on the basis of perceptions automatically activated by the presence of these elements within the same attentional focus. The first sequences extracted from the language input facilitate the receptive treatment of longer extracts, participate to the construction of corresponding mental representations, and serve as a basis for utterance production. The associative treatment is considered to proceed unconsciously, automatically triggered by the attentional processes, while the perceptions and mental representations may access consciousness. The constructed representations are isomorphous with the sequential structures exemplified in the input. Perruchet and Poulin-Charronnat (2015) argue that this learning model allows accounting for complex behavioral adaptations without having to evoke rule knowledge.

 It is worth noting that attention and consciousness are separate systems. Attention does not automatically determine conscious awareness. Gazzaniga (2008) indicates: ”Experiments have shown that in order for a stimulus to reach consciousness, it needs a minimal amount of time to be present, and it needs to have a certain degree of clarity” (p. 285). However, “…although attention may be present, it may not be enough for a stimulus to make it to consciousness” (ibidem, p. 286).

 Longitudinal analyses of mother-child verbal interaction (cf. the Childes Language Data Bank, MacWhinney, 2000) illustrate the dynamics of an extraction model.

 For instance, Moerk (1983) has provided an extensive analysis of the language exchanged between Eve (one of the “Harvard children”) and her mother. It contains numerous episodes where the mother of Eve proposes a given formulation not yet productive in Eve’s repertoire; for example, the construction *in the waste basket* (stressing the preposition and the article) at several times interspersed in the verbal turns. In following interactive episodes, Eve’s mother checks her child’s ability to use correctly this prepositional phrase and then generalizes it to other phrases constructed according to the same pattern.

 In mothers’ speech to language-learning children, Waterfall et al. (2010) have documented dative alternation with either a double object construction (e.g., *Give Mommy the ball*) or a prepositional object construction (*Give the ball to Mommy*) and variation in object complement verbs (e.g., *Who do you think started the whole thing?* vs. *Who do you think started it?).* A related work by Onnis, Waterfall, and Edelman (2008) has investigated the role of variation sets in language acquisition by manipulating their distribution in two artificial language experiments with adults. They concluded that variation sets may aid the acquisition of multiword phrases. This may be considered a bona fide indirect argument in favor of a particular role of variation sets in syntactic development. Variations in utterance positions, coupled with contiguous presentations, may facilitate children’s identification of the repetitive part of a variation set. However, no empirical demonstration was supplied by Waterfall et al. (2010) of the plausible effect of variation sets on actual syntactic development.

 Moerk (1983) interpreted his data as evidence of conscious parental language teaching vis-à-vis the children (this is the title of his book). This is probably an overstatement. It has not been attested that most parents act in the same didactic way with their children as Eve’s mother, even if they do simplify and adapt their input to their their language-learning children (see below). In this respect, I agree with Chomsky (1965; quoting Von Humboldt, 1836) when he suggests: “… one cannot really teach language but can only present the conditions under which it will develop spontaneously in the mind in its own way (p. 51). I am basically in agreement with this statement (not with Chomsky’s related Platonistic view, also shared with Humboldt, that “learning is a matter of drawing out what is innate in the mind” (ibidem).

 I have contributed a longitudinal interactive corpus in French figuring one mother and her male child, named Stéphane, in twice half-an-hour weekly sessions of natural interaction at home for 30 months. Stéphane was 27-month-old at the beginning of the recording period and 57 month-old at the end (Rondal, 1994b; the entire corpus is deposited within the Childes Language Data Bank (it can be accessed at: http://childes.talkbank.org/data/Romance/French/rondal.zip. ). Stéphane’s MLU (mean length of utterance) was 2.25 at the beginning of the study and 4.50 at the end (MLU computed in number of words per utterance; cf. Rondal et al., 1987).

 The analysis of this corpus (Rondal, 2015) also illustrates the dynamic of the extraction model.

 At the two-word stage, the child typically extracts and reproduces with the same sequential order two consecutive words (referring to action-process on people and physical entities: toys and common home objects) from an immediately preceding maternal utterance (e.g., Mother: *Va chercher l’auto*; gloss: *Go get the car*; Child: *Chercher auto*; *Get car*). As soon as a sequential pattern is used productively, it is generalized to sequences incorporating other (newly acquired) lexemes (e.g., *Get book; Get key; Get candy*; etc.).

 The emergence of simple declarative active sentences is easily retraceable at the three-word stage: nominal phrase (grammatical subject) + transitive or intransitive verb phrase (verb and direct object or verb alone) coding for a semantic relationship agent + action-process + patient (e.g., *Fafa conduit auto; gloss: Fafa drives car*) or agent-action (e.g., *C’est Fafa conduit; It’s Fafa driving*). Child: (1) extracts the corresponding sequential pattern from preceding maternal utterances maintaining the same word order; (2) expresses the learned sequential pattern with other vocabulary items (newly learned or retaken from an immediately preceding maternal utterance); (3) expresses an original production without relation to immediately preceding maternal input; and (4) when the expressive pattern is stabilized, paradigmatic substitution begins.

 State and experiential verbs are gradually inserted into the sequential patterns expressed by the child. Sentences slowly increase in length through the identification of new semantic relationships between verbs and nouns; for example, location, possession, benefit, instrument, complement. The child’s productions are regularly acknowledged by the mother often expanding them with missing morphosyntactic elements. However, no direct corrective feedback is formulated.

 Child gradually makes his production syntactically more complex extracting more advanced patterns from maternal input, reproducing them rather closely at the beginning (except for relatively longer instances that are still beyond his short-term memory capacity). Active declarative negatives with the negative element inserted within the sentence are produced as well as requests in action or information formulated either as declaratives, imperatives, or interrogative sentences.

 Compound sentences are produced then complex ones, involving the production of subordinate completive, circumstantial, causal, temporal, relative, and conditional subordinates. Without exception, subordinate clauses are derived to the right of the main clause, which is also the case in maternal speech.

 As regards syntactic morphology, and already noted by Brown (1973) and Moerk (1983), no syntactic morpheme is learned and stabilized immediately. There is always a period of production/no production followed by a gradual increase in production in obligatory contexts. The singular forms of the present tense of the indicative (the person is morphologically differentiated in French) appear first borrowed from immediately preceding maternal utterances and then gradually used freely. They allow opposing the indicative to the infinitive or imperative mode. The past tense is next produced referring to a completely terminated action, state or experience at the moment of production. The imperfect appears slightly later (duration in the past), and then the future (immediate or at short time interval).

 Verbs are morphologically marked in a piecemeal manner that is, one by one, extracted from maternal input. Singular personal pronouns are the first to be expressed. The developmental order is: first-person (*moi, je- me*; *I),* then third and second person pronouns. Plural personal pronouns appear later. The same evolution is observed for the possessive adjectives and pronouns. As to the articles, the singular definite forms are used first (*le, la; the*) followed by the indefinite forms (*un, une; a*).

 The plausible empirical validity of an extraction model of the kind suggested by Perruchet and Poulin-Charronnat (2015) depends upon the well-formedness of the input to the language-learning child. If it were not the case, it would be impossible to explain from where grammatical knowledge could come in language users (given that a genetic predisposition in that respect does not appear to exist).

 There have been authoritative declarations to the contrary in the linguistic and psycholinguistic literature. For example, Chomsky (1965) writes: “It seems clear that many children acquire first or second languages quite successfully even though no special care is taken to teach them and no special attention is given to their progress. It also seems apparent that much of the actual speech observed consists of fragments and deviant expressions of a variety of sorts. Thus it seems that a child must have the ability to ‘invent’ a generative grammar that defines well-formedness and assigns interpretations to sentences even though the primary linguistic data that he uses as a basis for this act of theory construction may, from the point of view of the theory he constructs, be deficient in various respects” (pp. 200-201).

 Fodor (1966) suggests: “One point about the corpus (i.e., the corpus of language to which the child is exposed; my addition) should, however, be noticed. If it is anything like a randomly selected corpus of adult utterances, it must contain a very substantial number of false starts, slips, grammatical mistakes, and so forth” (p. 108).

Chomsky’s and Fodor’s statements are not based on actual data. A number of empirical studies from the early 1970s have established beyond reasonable doubt that the language addressed to language-learning children by the parents and other adults is grammatically well-formed (Broen, 1972; Slobin, 1975; Newport, Gleitman, & Gleitman, 1977; Marcus, 1993).

More recently, Perfors et al. (2011), in the course of the research summarized in Chapter 3), have screened the adult language addressed to Adam, another one of the “Harvard children”, for grammatical well-formedness. The original corpus contained 25,755 individual sentence tokens, of which 443, i.e., only 1.72%, were ill-formed grammatically according to standard normative English grammar. The authors also report that the final corpus analyzed, i.e., after removing topicalized sentences[[28]](#footnote-28), sentences containing subordinate phrases, sentential complements, conjunctions, serial verb constructions ( for reasons of computational tractability as well as the difficulty involved in designing grammars for them), and ungrammatical sentences, contained 21,671 sentence tokens, of which 7371 were sentence fragments.

 I have assessed the grammatical well-formedness of the mother’s speech to her child Stéphane in the longitudinal corpus of language mentioned above (Rondal, 1994b). The French normative grammar of reference used was the *Le bon usage* of Grevisse and Goosse (2011). The whole corpus contains 58,734 conversational turns (approximatively half by the mother).Two thousands utterance tokens covering the whole corpus were randomly selected for the well-formedness analysis (in ten blocks of 200 consecutive maternal utterances addressed to the child interspersed with the child’s productions). They were analyzed in three categories: incomplete sentences, grammatically well-formed, and grammatically ill-formed sentences.

Results show very few ill-formed maternal sentences (no more than four or five per block; i.e., 2.5% or less). About a dozen incomplete sentences were identified per block (i.e., around 6%). They were classified as incomplete without taking into account the conversational and the extralinguistic contexts. The sentences were not pragmatically incomplete in the sense that their meaning was obvious when taking these sources into account (for example, many elliptic statements were responses to a preceding child question or comment; e.g., child: *Joue avec ça*; gloss: *Play with that*; mother: *avec l’auto*? Gloss: *with the car?).* Examples of grammatically ill-formed sentences are: *maman du cirage oui sur le soulier* gloss: *mommy polish yes on the shoe*; *tomber sur ta tête ah oui*; *fall on your head of yes*; *la clé pour le tableau aussi*; *the key to the board also*. As these examples suggest, the ungrammatical utterances were perfectly understandable when situated in the conversational context with the child.

Another important characteristic of parental speech to language-learning children must be taken into account. As a number of studies published from the early nineteen sixties (Rondal, 1985, 2006, and Moek, 2000, for reviews) have demonstrated, this type of language is adapted in continuity to and modified gradually in correspondence with the developmental level of the child.

The adaptations concern virtually all linguistic aspects: relatively larger acoustic distance between vowels (i.e., separation of formants or distinctive frequencies of the acoustic signal) rendering them easier to process categorically, prosody and intonation, rate of speech, articulatory precision, selection of lexical terms, semantic relations and semantic content, utterance length, morphosyntactic complexity, and discursive organization. Integrating various studies (Rondal, 1985), it is possible to show that the modifications in parental speech (compared to the language exchanged between adults) start in the first weeks of life of the baby and are observable until early adolescence although gradually to a lesser extent..

There is no structural difference in parental speech addressed to boys or girls. Mothers and fathers (as well as caregivers and other familiar adults) adapt their level of language to the children in the same way, even if some differences have been reported in style of interaction (for example, fathers appearing to produce more requests in clarification than mothers).

 Mothers’ adaptation to the language level of the children is fine-tuned. At every moment in development between MLU 1,00 and 5,00 in children, maternal MLU is superior to child’s MLU by a factor of 2.50 with younger children and 1.00 with older children (Rondal, 1985). The double temporal evolution (mother-child) is positively linearly correlated: Bravais-Pearson index .69, statistically significant at p < .001 level, between 2 and 5 years in Moerk’s data (1985); .55, significant at p < .001, in Rondal’s data (1978; entire corpora deposited in the Childes Language Data Bank; MacWhinney, 2000).

Pinker (1994) has counter argued that “motherese”[[29]](#footnote-29), as he calls it, is found only in Occidental cultures (see Lieven, 1994, however, for a synthesis of several intercultural researches demonstrating the contrary). As a case in point, Polynesian parents, it was reported (Slobin, 1981), speak little to their children until they have reached a certain level of cognitive development. This could suggest that parental linguistic adaptations actually do not play a decisive role in language development. Additional observations have shown, however, that in these societies the grand-parents and the older siblings are in charge of interacting with and speaking to younger children with, it would seem, the same kind of language simplifications and adaptations as those documented in Western families.

 This latter indication may be set in relationship with a series of observations showing that older siblings in Western families also adapt the language they address to their younger sisters or brothers in corresponding ways to those of the parents (for example, Shatz & Gelman, 1973; Sachs & Devin, 1976; Beaudichon, Sigurdsson, & Trelles, 1978).

 It is likely that the input adaptations from the interlocutors when interacting with children are favorable to language learning and in particular to morphosyntactic learning (Hills, 2013). However, Newport et al. (1977) and Pinker (1994) have questioned the simplicity of child-directed speech. They argue that the high proportion of questions in adult-child interaction (between 33 and 53% in the studies reviewed by Rondal, 1985) may render the speech exchanged actually more complicated to deal with. Pinker (1994) adds: “Motherese is riddled with questions containing *who*, *what*, and *where*, which are among the most complicated constructions in English” (p. 279). The important proportion of fragments of sentences, as documented above, in parental speech addressed to children could also be problematic in a learning perspective.

Despite these possible caveats, there are numerous studies demonstrating significant effects of parental speech on various aspects of language development in children. Regarding syntax in particular, Barnes et al. (1983) observed a positive correlation between the quantity of language addressed by parents and progressive MLU increase in their children; Newport et al. (1977) and Furrow, Nelson, and Benedict (1979) have reported positive correlations between proportions of yes/no questions with auxiliary fronting and individual rates of acquisition in children. Rowland and Pine (2000) and Rowland et al. (2003) have shown that the relative frequencies of appearance of wh-terms and auxiliary fronting in parental language to the children predict the order of acquisition of corresponding syntactic structures. Huttenlocher et al. (2003) have documented a strong relationship between the individual rates of acquisition of compound and complex sentences in children and parental use of these structures in dyadic interaction sessions.

 Ambridge et al. (2015) have reviewed several more recent studies demonstrating frequency effects of parental speech on the acquisition rates of a series of syntactic structures. For example, several works regarding various languages, some non-Indo-European, show a significant correlation between order of acquisition of interrogative lexemes (pronouns, adjectives, and adverbs) and the relative frequencies of appearance of these terms in parental input (proportion of variance explained: around 36%).

 Full passive sentences have a low level of occurrence in languages like English or French. In Sesotho (a Bantou language), however, passive sentences are approximately ten times more frequent than in English. This has to do with the fact that in Sesotho the lexemes positioned at the beginning of a sentence must obligatorily refer to old information. This constrains speakers to formulate a full passive in response to any question referring to an old entity. For instance, to a question like *Where did you get that*? The answer must be something like *I was given it by X* rather than *X gave it to me*. This pragmatic convention determines a high proportion of passive sentences in Sesotho language, which favors an earlier chronology of development of these sentences (Kline & Demuth, 2010).

 The above indications should not be taken to suggest that syntax-learning children typically perform complex computations regarding statistical aspects of the input they are exposed to. It means that children are able to distinguish frequent forms from less frequent ones in the language they receive and, it would seem, concentrate on the former. Denison and Xu (2014) have documented a strategy of the kind in children as young as 12 months.

The relationship between implicit and statistical learning has been discussed in the literature. Ding et al. (2016) indicate that the difference between rule and statistics concerns the level of abstraction. For example, after being exposed to a number of noun phrases, one may induce that a given word can be placed before or after another word. One might also induce a rule saying that a class of words can be used to modify another class of words. Seidenberg, MacDonald, and Saffran (2002) suggest that such abstraction could be an implicit process (with no conscious awareness) during language acquisition or an explicit process when learning grammar in school. As said, Perruchet and colleagues disregard the notion of rule in explaining implicit learning, admitting, however, that it cannot be completely ruled out (cf. Witt & Vinter, 2012).

 Ellis (2015) indicates that language knowledge involves statistical knowledge, adding that there must be some cognitive mechanism that tallies the frequencies of occurrence of various language sequences. He adds that this knowledge does not correspond to formal rules but rather to a huge collection of particular memories. Language learning is essentially learning sequential probabilities of the units of language. Perruchet (2005; see also, Perruchet & Pacton, 2006) argues, in opposition, that implicit learning does not imply mentally performing statistical computations. It may go the other way around, i.e., sequences of associated units repeatedly identified in the input may access consciousness and it is these properties of conscious experience that make the learner sensitive to particular statistical characteristics of the input.

 What matters most for syntactic development, it would seem, is the language input itself. There has been a long debate in the literature regarding positive and negative evidence in morphosyntactic development. As indicated in Chapter 3, learnability theorists have advocated the position that without negative evidence (as to what is grammatical in a language and what is not), the grammar of a natural language cannot be learned. Wexler, Culicover, and Hamburger (1975) ruled that negative evidence is more powerful in development than positive evidence (simple input). There is little doubt that parental feedback contingent upon children’s language productions exist in large numbers. Most usually, however, they do not supply negative evidence to the child as to her/his preceding language production (see Rondal, 1985, for a synthesis of a series of studies). The issue does not relate to the frequency of parental feedbacks but rather to their nature (also Pinker, 1984). Leaving aside a low percentage of punctual corrections (more often related to lexical misuse or pronunciation - with the older child in the latter case), the huge majority of remaining parental feedbacks in natural interaction with children is more concerned with referential adequacy, truth value, and general semantic considerations, than with grammar and morphosyntax.

 There are four cases documented in the interaction studies: (1) parental approval of semantically and grammatically correct child’s utterance; (2) disapproval of semantically and grammatically incorrect child’s utterance; (3) approval of semantically correct but grammatically incorrect child’s utterance; and (4) disapproval of semantically incorrect but grammatically correct child’s utterance. Cases 1 and 2 above do not allow knowing whether parental feedback is motivated by the grammatical or the semantic aspect of the child’s utterance. Cases 3 and 4 involve the risk of inducing the child in error as to either the semantics or the grammaticality of her/his preceding utterance. It follows that on the basis of a single feedback it is virtually impossible for the child to conclude about the well-formedness of her/his utterance. (S)he would need to compare a series of feedbacks related to the same type of utterance in order to make a judgement. This seems to be a very complicated endeavor.

 Marcus (1993) has calculated that a given child should repeat the same type of utterance a large number of times (from one hundred to several hundreds) and memorize the relative valence of the adult feedbacks, in order to be able at the end to decide whether or not this type of utterance is well-formed according to the standards of the linguistic community. It is known that children tend to repeat themselves little. Pinker (1989) has looked for self-repetitions in the corpora of the Harvard children (80,000 utterances). He has identified, besides a few more productive routines, no self-repetition superior to three times, except one grammatical error committed eleven times on the whole corpus by one of the children.

 Parents could use, voluntarily or not, implicit means to convey a grammatical feedback to the children. For example, it has been documented (Hirsh-Pasek, Treiman, & Schneiderman, 1984) that mothers most of the time do not repeat (verbatim or partially) those children’s utterances that are grammatically well formed. Demetras, Post, and Snow (1986) have also observed that following a well-formed utterance produced by the child, mothers’ most likely behavior is to go on with her speech. Following an ill-formed utterance, mothers tend to produce a clarification request. Chouinard and Clark (2003) have reported that parents tend to reformulate children’s ill-formed utterances more often than repeating well-formed utterances.

 It could be then that implicit grammatical feedbacks are actually produced, at least to some extent, in the interactions between parents and children. However, it has not been demonstrated that children understand them as such or that they actually make use of them. There may be a correspondence problem involved. How would a child know to what precise part or aspect of her/his utterance the information conveyed in the implicit feedback relates?

 The important point for an extraction model is that the input to the language-learning child is mostly grammatical. Given that this prerequisite is met, as the observations suggest, the model may function from positive evidence only. Children’s possible errors, overgeneralizations, undergeneralizations, and so forth, may be gradually filtered out through an implicit process of comparison with the input.

 **7.2.1.1. Semantic-relational framework**

 Child uses several types of information (prosodic, phonotactic, meaning and references of the entities and events in the environment) in order to segment the continuous flow of the speech heard. Parental speech to the young child contains a high percentage of isolated words, repeated and proposed at regular intervals (Brent & Siskind, 2001). Once identified and memorized, they serve as anchor points for apprehending short sequences of contiguous words extracted from the input.

Perruchet and Poulin-Charronnat (2015) suggest that the first sequences extracted from input are composed of two or three items given the limitations of the young child’s attentional capacity. They increase gradually in length. Strongly associated items in learned sequences become combinable units for longer sequences. An important augmentation in the linguistic material interiorized by the child takes place in the space of a few months. This increase, it would seem, necessitates the elaboration of a reduction system also able to serve as a basis for the generation of a large quantity of patterned utterances. The question is to know what happens to the thousands of sequences extracted over the hundreds of hours of exposure to language input. A reduction system means an abstraction system of some kind. But which one?

Traditional psycholinguistic suggests that children must access the distributional equivalents of the linguistic formal and functional categories for organizing economically the extracts memorized, making additional progress in morphosyntactic ability, and being able to generate an infinity of grammatical sentences. As said, implicit learning involves also developing abstract knowledge regarding, for example, the order and position of items including nonadjacent ones. Complex representations may enter in the associative links that are at the basis of learning. Although adding that these representations are not necessarily akin to the grammatical categories and formal hierarchies of structural linguistics, Perruchet and Poulin-Charronnat (2015) do not exclude this possibility in a kind of second developmental phase (but they leave the question open, p.159).

Tomasello (2003), while advocating a functionally-based distributional analysis of syntactic learning, does not exclude the same possibility either. At the same time, however, and somewhat contrarily, he questions the psychological reality of linguistic grammars (p. 303). In a later work, Abbot-Smith and Tomasello (2006) argue for “exemplar learning.” Children are supposed to store in memory exemplars of syntactic constructions; for example, many exemplars of transitive constructions of the SVO (subject-verb-object) type in English, from which they generalize and may develop abstractions corresponding to the traditional syntactic categories.

 Ullman (2004) has proposed a double mechanism for language learning, labelled declarative-procedural. In this model, the declarative mechanism underlain by declarative memory is in charge of semantic learning and the learning of irregular grammatical morphology (see Section 7.2.2., below). The procedural mechanism depending upon procedural memory is responsible for phonological and syntactic learning as well as the learning of regular grammatical morphology. Procedural learning according to Ullman (2004) is implicit. Ullman, however, is convinced that formal rules and linguistic categories are operational in morphosyntactic learning. Semanticists, as Chafe (1973), Van Valin (1999, 2003), and others, have also made it clear that, for them, at the post-semantic level of language production, the abstract notions involved are of a formal nature. This may be the case (see next section for a discussion).

There are other possibilities, however.

Ibbotson and Tomasello (2016; already Tomasello, 2003) suggest that children acquire language (and morphosyntatic regulations) via their cognitive abilities and the realizing of other people’s intentions. These ideas are not new. As indicated in Chapter 3, there have been suggestions as to the role of cognitive development in morphosyntactic development (e.g., Bever, 1970; Sinclair, 1971, 1973). The problem, as said, is that cognitive development in itself does not carry specific information regarding the content of morphosyntactic development and regulations. It is, of course, a general prerequisite at least to some extent (remember the cases of the “hyperlinguistic” people with severe cognitive handicap illustrated in Chapter 3). The same can be said for the capacity to read other people’s intention advocated by Tomasello (2003) and Ibbotson and Tomasello (2016), involving joint attention, interest to what other people are looking at and doing, and a capacity to imitate other people’s action. Theory of mind (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, 1991) can be added as a necessary component of the communication and language ability, as suggested by its deficit in the autistic spectrum. Although these abilities are prerequisite for language development, they do not carry in themselves specific information as to how to organize sentences morphosyntactically.

 An interesting proposal is that of Ambridge and Lieven (2015) defining a constructivist framework for early morphosyntactic development. Children are considered to start out with holophrases (i.e., a single word expressing what would typically take a sentence in adult language), which develop into lexical schemata via a process of early abstraction, and then into adult-like abstract constructions. As a prototypical example, their account of the English determiner system is as follows. Children hear and store strings of the following types: *a ball*, *the ball*; *a book*, *the book*; *a doggie*, *the doggie*; *a man*, *the car*; etc. They then treat these exemplars to form two “slot-and frame” schemata: *a* (X) and *the* (Y). The two repertoires are orthogonal at the beginning. The use of X and Y above are meant to emphasize the claim that children have not formed formal categories such as the noun category (Ambridge & Lieven, 2015, add that they may never do; also Ambridge, in press, see Section 7.2.1.2. below). With more language exposure, the slots in the particular schemata are enriched with semantic (e.g., for the articles, the constrast between discrete and nondiscrete entity) and pragmatic properties (e.g., the contrast between specific and nonspecific or discourse-new vs. discourse-old reference).

The acquisition of basic word order follows the same path. Across a number of stored utterances extracted from the input, children form schemata of the types: agent-action [e.g., *Daddy eat*(*s*)], and action-patient (*eat* X). The overlaps between many such strings are considered sufficient for children to integrate these two schemata into the more complex schema agent-action-patient. The question then is whether such semantically-based constructions develop into functional linguistic categories (e.g., subject-verb-object). Ambridge and Lieven (2015) remark that for transitive constructions (only as a case in point) the developmental endpoint could differ from a “simple” SVO structure and correspond rater to a series of frames involving particular semantic relations [e.g., contact (*John hit bill),* causative (*John broke the plate*), experiencer-theme (*John heard bill*), theme-experiencer (*John scared Bill*), “weigh” constructions (*John weighs 180 pounds*), “contain-construction (*The tent sleeps four people*)].

 Movement operations (as theorized in the generative approach) are not posited in Ambridge and Lieven’s (2015) account. Rather passive and question constructions are considered to be abstracted from input utterances relying on the same processes as outlined above and this is also the case for relative- and complement-clause constructions.

 In accordance with Ambridge and Lieven’s (2015) constructivist theory, I suggest that semantic relations actually provide a first (and possibly continuing) potent mechanism for reducing the repertoire of sequences extracted from the input by the language learner to a set of accessible abstractions without the need to rely on formal rules and hierarchies.

As indicated, one interesting characteristic of semantic relations is their flexibility and combinative capacity. They may be set in correspondence with varied formulations in sentence surface. There is another important aspect of semantic relations (already alluded to and illustrated in Chapter 6). It is the fact that underlying surface expression, they may act as a short-term memory buffer for momentarily keeping relevant information and as a bridge, or a conceptual splint, for relating sentence nonadjacent parts. In Section 7.1., mention was made of the likely importance of the intermediate element(s) in learning nonadjacent dependencies in sequences of items. Semantic relations may be relevant in this respect as, by their very nature, they maintain an underlying intermediate link between related distant items whenever a surface expression is momentarily interrupted for inserting supplementary material.

Concretely, regarding development, once the child has acquired a repertoire of simple utterances, it may be hypothesized that the corresponding cognitive representations trigger the mental activation of corresponding semantic relations. The infants’ initial semantic representations may be imprecise. They continue to sharpen in the second year of life and beyond (Bergelson & Aslin, 2017).

 The major meaning combinations formulated in early child language (not necessarily acquired at the same time by all the children) were already identified by Braine (1963, 1976), Braine & Wells (1978), Schlesinger (1971), and Brown (1973). Formulated according to Chafe’s (1973) terminology, they may be listed as follows: agent + action (e.g., *horsie running*) ; action-process + patient (*drink milk*) ; agent + action-process (*baby eat*) ; action + location (*run garden*) ; action + instrument (*cut knife*); action + beneficiary (*give daddy*); action + accompaniment (*go mamma*); entity + location (*car garage*); entity + -possessor (*car mamma*); entity + attribution (*baby little*); notice + entity (*hi ducky);* and experiential + patient (*see daddy*).

At the three- and four-word stage, one finds: agent + action-process + patient (*baby brush doggie*); agent + action + location (*baby run garden*); agent + -process + patient + location (*baby wash doggie bath*).

The child’s repertoire of semantic relations is gradually enriched and complexified until reaching the adult level as characterized in the semantic theories analyzed in Chapter 5. These relations not being naturally sequentialized , as indicated, have the capacity to underlie various syntagmatic formulae in correspondence with conventional syntactic patterns in the community language. A precise correspondence emerges gradually as a direct function of exposure to this language.

Evans (2014) advocates a semantically-based account of early syntax. She acknowledges that relational meanings are abstractions in the sense that they are not directly tied to the words appearing in sentence surface. This abstract character allows the speaker to convey a host of lexically different surface realizations based on similar semantic relations. When children hear expressions that embody a given (cognitively accessible) semantic relation, they memorize the corresponding surface pattern. “By around the age of four, infants have a complex mental corpus, made up of words, fixed-patterns of words (e.g., idioms), and more abstract sentence-level constructions, that have been extracted from the language they have been exposed to” (Evans, 2014, p. 247).

A semantically-based theory of early syntactic development was already suggested by earlier researchers in the field (for example, Ervin-Tripp, 1966), also confirming the generality of semantic relations in children’s combinative constructions across languages (Ervin-Tripp, 1971). Braine (1963, 1971, 1976) distinguished three kinds of formal patterning in early corpora of word combinations (in several languages: Finnish, Samoan, Swedish, Hebrew, and Quechuan), which he called positional, positional associative, and “groping.” Braine considered that they reflected children’s early search for ways to express basic semantic relations and generalize them to new expressive contexts. He further remarked: “Thus, no special acquisition theory is required to account for hierarchical organization. The idea that children are acquiring positional formulae that map meaning into form predicts that hierarchical organization may appear whenever the semantic basis of the formulae permit one formula to be embedded in another, and there is a result where the context would make sense” (1976, p.91).

Bowerman (1973) concluded from her study of early syntactic development in four Finnish-speaking children, that the kinds of notions children use in producing their two- and three-word constructions may be primarily semantic. She added: “As the child matures linguistically, he begins to recognize regularities in the way different semantic concepts are dealt with and to gradually reorganize his knowledge into syntactic concepts, which are more abstract. At any point in time, some of the concepts which are functional in the child’s competence may be primarily semantic and others primarily syntactic. If this true, the optimal grammar for child language must be capable of operating with both semantic and syntactic concepts. It must also be flexible enough to represent shifts over time to new levels of abstraction, so that, for example, a sentence constituent which at one time might be represented as an ‘agent’ would at a later time be represented as ‘sentence-subject” (p. 227).

Bowerman’s conclusive suggestion is close to the spirit of the present essay as seen in preceding sections and as will be seen again in what follows.

 **7.2.1.2. Formal syntactic learning**

 Along with more exposure to language input children may gradually develop an implicit knowledge of the distributional equivalents of a least some of the linguistic formal and functional categories. It is a difficult enterprise and it needs some environmental help.

 A mechanism conceptualized by Pinker (1984, 1987), and already suggested by Grimshaw (1981) and Macnamara (1982), under the name of semantic bootstrapping (in a sense, the reverse of syntactic bootstrapping - Chapter 3), can explain, perhaps in interaction with other strategies (see below), the transition from a semantically based approach to morphosyntax to a more formal one. There seems to be enough empirical data corroborating such a hypothesis.

 The idea is that children exposed to the expression of semantic relations and word semantic roles in the input will use these semantic categories as cues (“flags”) to related syntactic categories, which, in turn, would help them to discover other correlated properties of those grammatical notions. As children associate more words with syntactic categories, they can begin noticing other properties that can help them identify these syntactic categories in the absence of semantic evidence. Furthermore, from conceptual relations they may identify functional syntactic relations. For example, phrases structure organization could be learned with the help of lexical entries plus semantic-syntactic correlations pertaining to phrase structures (for example, grammatical subject precedes verb phrase). This would lead children to ultimately bootstrap their way to an understanding of the language’s grammar and formal expression. If correct, the semantic bootstrapping theory would offer a good example of the interaction between semantic-relational and formal-syntactic systems in learning language syntax.

Pinker (1987) rejects two alternative views to the semantic bootstrapping hypothesis. Prosodic bootstrapping (Morgan & Newport, 1981) posits that the child records the intonation contours, stress patterns, and pause distribution in input sentences and, on this basis, infers the phrase structure trees of the sentences. Correlational bootstrapping (Maratsos & Chalkley, 1981; Braine, 1987) assumes that the child analyzes distributional properties of the input, such as serial positions, inflections, and so on, and in so doing construct his grammatical categories. Pinker states that these theories do not provide any departure basis for the children’s grammatical constructions and that many linguistic relevant properties are not perceptually marked in the input.

 It is correct to mention, however, that in further contributions (for example, Pinker, 1989), Pinker has partially modified his previous assumptions into a slightly different type of acquisition mechanism labelled “constraint acquisition model,” accommodating the possibility of multidirectional and more opportunistic (probabilistic) learning. The semantic information continues to be considered as the most important source to get the formal learning model started, but other sources of evidence (e.g., those advocated by the correlational bootstrapping account) are no longer rejected once a certain amount of learning has already taken place.

The validity of a semantic bootstrapping theory depends upon two assumptions. First, it requires that children are able to perceive and identify a set of semantic relations and roles in the language they are exposed to. We already know that this type of knowledge is available to the language-learning child from early on in development.

 Second, children must identify the links between semantic and formal syntactic categories. As said (Chapter 3), Pinker postulates the necessity for at least the formal notions of noun and verb to be prefigured genetically. Otherwise, the link between semantic and formal categories would be difficult or impossible to establish (the child not knowing beforehand which particular language aspects to look for and connect to). Pinker (1984) adds that the bootstrapping hypothesis is meant to explain how a correlation between semantic and syntactic constructions can help the learner to fix parameters in innate rule schemata (p.679). Accordingly, even if the child is innately equipped with innate constraints dictating that there exist noun phrases, etc., (s)he must still find tokens in the input so that instances of language behavior can be used to fix parameters or to apply universal syntactic principles.

 It is questionable, however, whether such a priori information is actually necessary for the semantic bootstrapping mechanism to function. In my advocating semantic bootstrapping as a possible way from relational semantic to formal syntax, I dissociate myself from Pinker’s innateness representational postulate. It is worth restating that children may only need to have at disposal a sufficient amount of semantic-relational/formal- syntactic pairings in order to recognize that there may be a systematic link between particular members of the two sets of categories; They do not have to and do not actually develop a technical theory of formal and functional categories. At best an implicit knowledge of the distributional equivalents of these categories is all they need.

Do we have empirical indications that the language input to children is compatible with a semantic bootstrapping theory freed from representational innateness? It would seem that this is the case.

 Hochberg and Pinker (1985) have examined the semantic-syntactic correspondence in parental speech to the Harvard children. They did it at two points in time: the first, when the children were at Brown’s stage 1 (MLU around 1.75), the second at stage 5 (MLU around 4.00). In stage 1, results showed that 90% of the nouns referred to physical entities or people; virtually 100% of the things and people were labelled using nouns; 88% of all actions were referred to using verbs; 73% of the verbs consisted of physical actions or states; 35% of the adjectives referred to physical properties; 100% of the physical properties were referred to using adjectives; about 64% of the prepositions referred to spatial relations; virtually 100% of the spatial relations were expressed using prepositions or adverbs; 100% of the agents expressed corresponded to grammatical subjects; and 85% of the patients were expressed as grammatical objects.

 No full passive sentence was identified in the corpora. There were only a few nominalizations (i.e., the use of a word that is not a noun as a noun with or without morphological transformation; e.g., the word *change* in the sentence: *Politic needs a change*).

 Rondal and Cession (1990) have also documented a good correspondence between semantic and formal syntactic categories in mothers’ language to 18 English-speaking children aged 1 year 8 months to 2 years 8 months, from the Minneapolis-St Paul area in the Unites States. The children were classified in three MLU subgroups (n = 6): MLU 1.05-1.58; 1.69-2.09; and 2.27-2.88. The entire corpora are consultable in the Childes Language Data Bank (MacWhinnney, 2000).

 A comparison of Hochberg and Pinker’s (1985) stage 1 data with those of Rondal and Cession (1990) reveals a large measure of agreement between the two studies. A vast majority (at times 100%) of the relevant semantic notions (action, agent, physical entity, etc.) in mothers’ speech were found to correspond (with very low standard deviations on the means of the dependent variables) with the syntactic elements canonically paired with them in the grammar of the language. One-way analyses of variance (on MLU levels) were carried out on the measures exhibiting some degree of developmental variation in order to test the significance of this variability. None yielded any significant effect at the conventional p < .05 level except for the category oblique object within the semantic category source-goal-location and the instrument one (p < .037). A follow-up analysis showed that the children in the first MLU group had significantly fewer source-goal-location and instrument categories expressed as oblique objects than the children in the other two MLU groups.

 It is of interest to note that in Hochberg and Pinker’s (1985) data the differences between stage 1 and stage 5 in percentage of the semantic categories expressed by canonical syntactic categories are almost non-existent for names, changes of state, and spatial relations. They are inconsistent between the three parents for the attributions. Summing up the major data in the two above studies, it may be concluded that the MLU interval studied (i.e., from about MLU 1.05 to 4.00) covers a period of time when the language input to the child is maximally coherent in terms of semantic-syntactic correspondence.

 As documented in the study of Rondal and Cession (1990), the input to the young child contains a substantial proportion of mental verbs (16% of all verbs on the whole) and state verbs (22%) against 62% for action verbs. From there the child could go two ways. (S)he could assume either that actions, states , and mental events are all expressed by one single global grammatical category, verb, or that actions belong to a category, say verb1, states to verb2, and mental events to verb 3. These subcategories would then be fused into a single adult-like verb category later in development. Another possibility yet, perhaps more plausible, is that the child temporarily filters states and mental events out of her/his verb category.

 Also, the use of semantic-relational information to infer the distributional presence of syntactic entities may work only (or work better) with active sentences and sentences in which no emphatic modification of the canonical word order has taken place. In English or French, for example, agents in passive sentences (e.g., *John was bitten by a wasp*) are oblique objects (i.e., preposition phrases) and patient are grammatical subjects, while in active sentences agents are grammatical subjects. This apparent contradiction would create a difficulty for a young child looking for a way to relate semantic roles and syntactic functions. The child might filter dubious instances out of the exploitable input using linguistic or nonlinguistic signals. Or the adults may do it for the benefit of the language-learning child. The well-documented fact in languages like English or French that full passive sentences addressed to children are extremely rare (e.g., Maratsos et al., 1985; Thibaut, Rondal, & Kaens, 1995), can perhaps be considered as an indication that parents unconsciously filter out such possibly misleading constructions when addressing young children. Once a basic semantic-syntactic correspondence is in order, the child may learn how to categorize semantically more “opaque” constructions (i.e., with respect to their syntactic correspondence) by observing their distribution within the known structures.

 Additional support for the semantic bootstrapping theory is illustrated in an experimental work reported by Gropen et al. (1991). They used locative verbs to verify whether children and adults could predict the syntactic environment of the verb from its relational meaning. Participants were taught novel verbs for actions involving the transfer of an entity to a container. They were then tested to establish whether they could produce the moved object as opposed to the container argument as the direct object of the verb. When the semantic relation was one of a container changing state by being filled with a moving entity (e.g., glass filled with water), children as well as adults were more likely to express the moving entity as complement of the verb. In opposition, there was no tendency in the participants to express the moving entity as complement of the verb when faced with a locative (directional) relation (e.g., coffee poured into a cup). The authors offered the interpretation that the verb’s syntactic framework could be predicted accurately from its semantic argument structure.

 If one looks at children’s early language productions, as documented in the preceding section, their nouns and verbs shows different characteristics of occurrence and they tend to correspond to physical entities and actions, states or action-processes. In the same way, what may be linguistically defined as grammatical subjects and objects correspond most of the time to agents or patients of actions, states, or action-processes (Maratsos & Chalkley, 1981; Macnamara, 1982).

 There have been several criticisms of the semantic bootstrapping hypothesis; some rejecting the theory altogether claiming (for example, Gleitman, 1990) that verbs are learned according to their syntactic properties and not their semantic ones (so-called syntactic bootstrapping, which is bound to meet conceptual difficulties in explaining syntactic acquisition as argued before in this book). Gleiman (1990) further argues that the semantic bootstrapping hypothesis underestimates the difficulty of word learning. She mentions the so-called *gavagai* or inscrutability of reference problem (as raised by Quine, 1960). Briefly said: how would a child hearing a word like *gavagai* orany other new word in a given language, know to which particular aspect of the extralinguistic context this word refers (passing entity, person, clothes, eyes, color, etc.)? The answer has been provided with the so-called fast mapping strategy (based on observations and experiments; as described by Mervis & Bertrand, 1995, and Waxman & Booth, 2001, for example). If not helped directly by the parents or caregivers, children guess at the possible meaning of the word heard given the context of the utterance. In succeeding instances, they may be led to modifying their first guess until reaching conventional meaning.

Ambridge, Pine, and Lieven (2014) question the semantic bootstrapping hypothesis as presented by Pinker (1984, 1987, 1989), considering in general (as said in Chapter 1), that a posited universal-grammar type of knowledge does not simplify the task facing the language learner. Regarding formal syntactic learning, in particular, they argue that children have a capacity for distributional learning, i.e., probabilistically grouping together words appearing in similar linguistic contexts. Children can detect trends in sentences where, for example, determiners go with nouns. This would be sufficient for learning syntax. Ambridge et al. (2014) suggest that children have the capacity for exploiting even imperfect correlations between semantic roles and morphosyntactic categories (realized by word order or grammatical morphology). Initial semantically-based categories can be expanded into corresponding syntactic categories via distributional similarity. This helps solving the conundrum often opposed to the semantic bootstrapping hypothesis regarding particular sentences in which, for example, the grammatical subject is not an agent (as in *The situation justified the measures*; example provided by Ambridge et al., 2014). Such sentences can still be analyzed by comparison with agent-action-process-patient ones. The same kind of similarity analysis can probably be achieved by children exposed to syntactically ergative languages (where, for example, the patient role may not become the grammatical object, like in English and many Indo-European languages, but most often the grammatical subject of the sentence). It can also be remembered that, as analyzed in Chapter 5, there is less need for the notions of grammatical subject and object of the verb in accounting for the syntax of ergative languages given that in these languages basic aspects of sentence surface can be directly related to semantic roles and relations.

 Ambridge et al. (2014) conclude that the learning strategies mentioned above obviate the need for innate linking rules as posited by Pinker.

 However, even if a bootstrapping theory expurgated from its innate representational component and possibly associated with syntactic similarity analyses, could do the job of launching the learning of both formal and functional syntactic learning in children, there is no guarantee, as already suggested, that this learning would ever be complete in comparison with the linguistic provisos. Important individual differences might exist between people in this respect.

 Theoretically children may use distributional learning to identify clusters corresponding to formal syntactic categories (and sub-categories as, for example, masculine and feminine grammatical gender; Reeder, Newport, & Aslin, 2017). The question is whether they actually do it and whether having identified such correspondence, they use it in morphosyntactic treatment, or whether they rather rely directly on distributionally defined clusters (also Freudenthal, Pine, & Gobet, 2005), or even whether they can use both information in conjunction already very early in development. The later possibility is suggested in an experimental work conducted by Xiaoxue He and Lidz (2017) showing that English-learning infants around 18 months are able to learn novel verbs relying on morphosyntactic cues for verb categorization and using the link between verb and event to specify verb meaning.

 In a paper in press, Ambridge explains why he has come to the conclusion that syntactic categories may be psychologically illusory. He confronts the illusory claim with two alternative conceptions: the traditional assumption that adult speakers are in possession of the categories used in linguistics to describe the syntax of their language, which means that children somehow must develop these categories; and the assumption that syntactic categories are induced from the input via some kind of distributional procedure.

 In the first conception, syntactic categories are innate and development may consist of using words’ semantics to assign them to relevant syntactic categories; this is the semantic bootstrapping hypothesis (at least according to its first proponent, Pinker) discussed above. An alternative along the same line of reasoning is the prosodic bootstrapping, in which children are assumed to use prosodic information to split sentences into phrases and then distributional indications (like the presence of articles or their absence) to assign words to syntactic categories, for example, noun and verb (Morgan & Newport, 1981; Christophe et al., 2008). The major problem with prosodic bootstrapping is that it is not clear how children would know which words are indicative of which category and how they would identify these words in the first place, except if they are innately specified. Ambridge (in press) mentions another nativist account of early syntactic development, that of Mintz (2003) under the name of “frequent frames” hypothesis.” It also assumes that syntactic categories are innately specified by universal grammar and posits the existence of a mechanism guiding the linking between these categories (needing to be present before linking can begin) and distributional frames. For example, distributional clusters containing a reference to concrete objects receive the label “noun”, and the clusters that take nouns as arguments the label “verb.”

While rejecting the innateness claim, most constructivist authors write as if they endorsed the psychological reality of the formal and functional linguistic categories. It is not clear whether this is just an expressive facility or whether they really believe that these categories necessarily constitute the end point of morphosyntactic development. As reported in previous chapters, it has been demonstrated through computational studies, experimental artificial grammar studies, and child-directed speech statistical analyses (also Ambridge, in press, for a summary of additional recent works) that formal distributional analyses applied to corpora of language work relatively well in classifying words according to syntactic category. But the question is not to know whether corpora of language contain distributional cues which operated by sophisticated algorithms can identify syntactic categories. They can. The interesting question is whether children exploit these cues and, if yes, how they do given that they are far from enjoying processing and memory capabilities comparable to those of the robots implementing the algorithms used in the works mentioned.

Ambridge (in press) argue that there is good evidence for multi-word exemplar representation and continuous update in real time in children. He suggests that children do not need to erase these representations or impose over them a catalogue of syntactic categories. Analogically comparing and generalizing across sets of exemplars stored in memory may be sufficient for any practical purpose. This is in this sense that it may be illusory to believe that syntactic categories have psychological reality.

 It remains, however, that exemplar-based models require people to keep in memory huge quantities of utterances. As suggested in the preceding section (7.2.1.1.), semantic relations, as “first-level” abstractions emerging naturally from general cognition and closer to sentence surface than formal syntactic categories, are reasonable candidates for providing the tools necessary to reduce the huge exemplar load in memory to more manageable proportions.

 Metalinguistic and metagrammatical teaching at school or later in life may assist formal learning in favoring people’s awareness of the relationships between semantic relations and linguistic formal and functional categories. However, judging from the lack of conscious grammatical knowledge in children and adults, as analyzed in Chapter 3, this kind of teaching appears inefficient. Additionally, the large number of illiterate adults on the planet, from what is known, not displaying particular morphosyntactic limitations in oral language when they are immune of disorder, does not plead in favor of the necessity of explicit grammatical teaching for regular language functioning.

 More generally, it is not known what language users actually do in language functioning with formal knowledge. Does formal information have a distinct role in language processing or is it an epiphenomenon? According to Paradis (2004), it is an epiphenomenon almost by definition as it may be considered that declarative knowledge never translates into procedural one and *vice versa*.

A last indication worth considering in this debate is provided in a series of experimental studies regarding written language by Pacton et al. (2001), Pacton, Perruchet, and Fayol (2005) (see also Deacon, Pacton, & Conrad, 2008).

 These works concern the implicit learning and transfer of some orthographic lexical regularity in French language. Results show that no orthographic rule is abstracted by the participants (adults and primary school children) in nonword judgements and completion tasks, following a massive exposure to artificial words. During the control phase, they rely on statistical regularities in the sequences of letters (for example, those letters that may be reduplicated, non-reduplication of vowels and morphemes, their positioning in the words (for example, the sound *è* followed by *t* is always transcribed as -*ette* in the role of diminutive affix; e.g., *voiturette*; gloss: *small car*). Even when there are morphological regularities (derivational or inflectional) or semantic ones that could motivate an abstract rule, they are not exploited by the participants. The subjects mostly utilize associative regularities between orthographic features related to co-occurrence of letters, sounds, and meaning elements, in the words, between the words in sentences, as well as orthographic analogies between words.

Extrapolating from written to oral language, it would seem that even where there is an opportunity for rule learning, it is not the strategy preferred by standard learners.

 In a conclusive nutshell, although semantic-relational and formal-morphosyntactic learning may interact, possibly for optimizing language functioning, it is quite possible, perhaps more likely, that the former system retains priority over the latter one in language users even beyond developmental times.

 **7.2.2. Learning grammatical morphology**

 As indicated in Chapter 6, a spread out conception in psycholinguistics is that the acquisition of regular grammatical morphology proceeds through the recognition of specific rules linking word stems and various affixes in order to express particular semantic indications. In contrast, according to the so-called dual-route theory (Pinker, 1999; Taatgen & Anderson, 2002), the exceptions to these rules (irregular cases of grammatical morphology) must be treated as paired associations between the stem and the affix, and memorized as such.

The matter may be more complex, however. For instance, just taking the example of the past tense in English, there are several sorts of irregular past tense formation: (1) vowel change (e.g., *break*-*broke*); (2) vowel change + *d* or *t* (*tell*-*told*); (3) final consonant becoming *d* (*have*-*had*); (4) total change (*go*-*went*); and (5) no change (*put*-*put*). Yang (2016) favors a rule-based approach where regular forms are computed by a productive rule and the irregular ones by a set of unproductive rules.

For the sake of brevity, the following discussion mostly relates on the past-tense marking in English, which is probably the single most studied topic in grammatical morphology (Pinker, 1999). For recent empirical indications regarding other grammatical markings and other idioms than English, one may see Yang’s extensive coverage (2016).

A central question is to know which particular mechanisms are involved in morphological marking and how they are learned. Rejecting rule induction from input[[30]](#footnote-30) (or rule deduction from an innate representational basis) as void of explanatory power, as argued before in this book, other possible mechanisms involve input statistical and frequency effects, analogical generalizations from semantic and morphophonological similarities, and perhaps a cognitive strategy defined by Yang (2016) under the name of tolerance principle.

 MacWhinney (1975) has shown that the development of morphology starts with rote memorization before children discover some of the processes of word formation in their native language. Pinker (1984) expresses the same idea in suggesting that acquisition goes from word-specific to more general paradigms. For example, the verbal affixes (e.g., -*s*, -*ed*, or -*ing*) are linked at the beginning to a restricted number of verbs and then gradually extended to a larger series of items. Concerning in particular the past tense, as indicated in Chapter 3, a typical U-shaped learning curve has been documented (Pinker, 1999). In a first stage, irregular verbs are inflected correctly just as easily as regular ones (e.g., *eat*-*ate*; *walk-walked*), before being over-regularized (*eat-eated*) at a later stage (Brown, 1973; Kuczaj, 1977). In a third stage, children come to distinguishing regular from irregular verbs and inflect them both correctly.

 Slobin (1978) has reported on the grammatical development of his daughter Heida, which provides an interesting case study in point regarding the acquisition of the English past-tense inflection (between 4 years 2 months and 4 years 9 months)

 At 4 years and 2 months, Heida would not accept (vehemently refused!) the adult irregular past-tense forms, insisting on her own forms, for example, *comed* (instead of *came*), *goed* (*went*), *knowed* (*knew*). At 4 and half years, she began accepting the adult irregular forms while keeping with her over-regularizations, and pretending that both types of forms were correct. She often hesitated, however. For example, asked about the past tense of *know*, she suggested *knowed*; however, when proposed *knew*, she accepted it as correct. In the same way, she suggested *wined* for the past tense of *win*; rejected *wan*, and accepted *won.* A couple of months later (with no corrective feedback from her environment), Heida started showing indications that she had begun giving preference to the adult forms. However, if following her use of a correct irregular form, one would (somewhat perversely) suggest that the incorrect irregular form could actually be the correct one, she would demonstrate signs of embarrassment. At 4 years 9 months, Heida gave up and recognized that the adult forms are the only correct ones and that her previous over-regularizations were erroneous.

 It would seem that a statistically-based model could account for this typical U-shaped learning. In a conception of the kind, beyond the first stage whereby only a verbatim repetition and rote memorization of a limited number of fixed forms takes place, the child becomes more sensitive to the probability distributions of the inflected forms in the input. As regular forms are more frequent than irregular ones (by definition) in the language of the community, the child could be led to model her/his grammatical morphological marking after the more frequent forms with the consequence that a number of irregular forms that are grammatical will be over-regularized. Then, with longer language exposure, the child realizes that (s)he is the only one producing odd forms and proceeds to clean her/his production repertoire in accordance with the standards in the matter, accepting the existence of exceptions to the regular patterns.

 Regular forms may not invariably be the most frequent ones in the language input to the children, however. For example, regarding the past tense in English, in 49 hours of adult speech to Eve (one of the Harvard children, aged between 18 and 26 months at that time), there were 292 irregular and 99 regular verbs (Slobin, 1971). Yang (2016) reports that 56% of the most frequent verbs in the language addressed to the Harvard children were irregular. These observations seem to go against the preceding suggestion. However, regular forms exhibit more formal coherence (by definition). For the past-tense inflection, for example, there are two regular markings (stem + *ed* or stem + *d*) vs. five variants for the irregular forms (as illustrated above). Even in fewer numbers, the case being, regular forms could be perceptually more salient.

Rule-oriented psycholinguists are keen to assimilate the intermediate stage in the U-shape evolution to a stage signaling the onset of the productive rule add-*ed* (or add-*d*). Pinker and Prince (1988) theorized that the learners internalize morphological rules covering only regular processes; the irregular forms being attributed to a process of analogy. But Albright and Hayes (2003), using a version of Berko’s (1958) “wug test” (whereby the participants are invited to inflect artificial words corresponding to fancy entities represented in pictures; e.g., “*Here we have one wug, and here there are two of them, there are two …* “), observed that the ratings of novel pasts actually depended on the phonological shape of the word stem in the same way for irregular as for regular pasts.

Numerous frequency effects of parental speech upon children’s learning of grammatical morphology have been reported (already suggested by Berko, 1958).

 Lahey et al. (1992), as well as Wilson (2003), have documented an important variability between children, in the timing of grammatical morpheme mastery. One of the differential keys is the frequency of corresponding morphemes in parents’child-directed speech. Moerk (1980) has shown that among the English morphemes acoustically well distinguishable (i.e., the articles *the* and *a*, the present progressive *-ing*, prepositions *in* and *on*, and the regular past form *-ed*), those most often used by the parents of the Harvard children, are the first to reach 90% of correct production in obligatory contexts. Linear correlation coefficients between grammatical morpheme frequency in parental speech and order of acquisition are between .56 and .67 for the three children (proportion of developmental variance explained around 30%). Brown (1973) had reported contradictory results, arguing that there were no correlation between relative frequencies of given grammatical morphemes (14 analyzed) in parental speech and children’s order of acquisition of these morphemes. However, he had mixed in his analysis both the acoustically well-distinct grammatical morphemes (here above) and others less acoustically distinct, such as the copula (contracted or not), the auxiliaries (contracted or not), regular plural on nouns, and third person of present tense of the indicative mode for verbs. For the latter morphemes, there is indeed no relationship between parental speech and early development in children, probably because they are not noticed by the children at those early times. Brown (1973) used a rank-order correlation analysis (Spearman rank correlation coefficient). The two contradictory series of data annihilated each other in the computation yielding an indication of statistical non-significance.

Ambridge et al. (2015) have reviewed a number of more recent research works, both naturalistic and experimental, on the acquisition of inflected forms by children in several languages. They confirm that there is a close relationship between frequency of use of grammatical morphemes in parental speech and rate of corresponding development in children. For example, Dabrowska and Szczerbinski (2006) found a significant correlation between the input frequency of genitive, dative, and accusative Polish noun case-marking inflections, and the children’s correct corresponding inflections of novel words. Regarding irregular verb and noun forms, several studies show that in English the high-frequency irregulars are less likely to be over-regularized than the lower-frequency ones; for example, *blew* and *feet*, for the former, as opposed to *drank* (*drinked*) and *shelves* (*shelfs*), for the latter (Maslen et al., 2004). In the same way, but concerning relative frequency, Matthews and Theakston (2006) found, for example, that children produce more often *two mouse* rather than the correct *two mice* because the latter is less frequent in the input; as opposed to *two feet* rather than the incorrect *two foot*, the former being preserved by its higher frequency in the input.

The learning mechanism needed to account for the pervasive frequency effects documented are not compatible with a conception that children would learn inflectional morphology by filling predefined squares in a kind of innate abstract chessboard structure (Wilson, 2003); unless some *ad hoc* parameter considerations (more gently, ancillary hypotheses with no independent theoretical motivation) like, for example for verb inflections, the existence of an innately given tense parameter with either a positive valence (i.e., requiring tense marking) or a negative one, set in relationship with differential input frequencies in different idioms (e.g., Legate & Yang, 2007).

 A more realistic learning mechanism, it would seem, involves storing and reusing strings of word from the input. Reusing strings of words implies an implicit analogical strategy whereby the child having identified an inflectional pattern generalizes it (rightly or wrongly) to semantically and/or morphophonologically corresponding items. The global statistical distribution of these items in the input will then influence the child into confirming or correcting her/his previous inflectional usages.

 Tomasello (2003) has also suggested that the morphosyntactic representations of the young child are based on examples and analogies before gradually evolving towards more formal types of representation (p. 315-318).

Assuming that the major mechanism involved in grammatical-morphological learning is analogy-based, on which language support does it act? Is the recognition of the syntactic categories and grammatical functions of the words in the sentence necessary for effectuating the proper grammatical markings? A clear knowledge of the linguistic formal and functional categories would certainly be of great help. It is not the only possible basis for inflectional marking, however. Morphophonological cues may play a role in inducing a kind of partial regularity in irregular forms. For example, in English, it is the case that a series of verbs ending in -*ing*, or -*ink* have irregular past tenses (*sing*-*sang*; *ring*-*rang*; *drink*-*drank*; *stink*-*stank*; *spring*- *sprang* or *sprung*; *fling*-*flung*; *slink*-*slunk*; *think*-*thought*; *etc.*). Semantic indications also play an important role (see also Ramscar, 2002). Proximal and distal association between words in sentences corresponding to particular semantic relations (e.g., agent-action, patient-state, patient-process) may underlie grammatical morphological marking to a certain extent (e.g., grammatical concord between subject and verb; gender and number correspondences between article and noun, grammatical subject and attribute, in some languages). Distributional information (e.g., a statistics of co-occurring words) may also play a role (Redington, Chater, & Finch, 1998), and of course an accurate (even if largely implicit) knowledge of exceptions (McClelland & Patterson, 2002).

Exceptions to the rules are at the center of Yang’s (2016) theory of language development. He argues that they have the consequence of increasing the processing complexity of language. This may lead, it is claimed, to a cost-benefit adaptation in the learner under the form of a threshold function for production applicable in the same principled way to phonology, morphology, and syntax. The title of Yang’s book (i.e., *The price of linguistic productivity*) is intended to analogize this problem in language learning with the price of goods as determined by the balance between supply and demand. The learner postulates a productive rule only if it results in a more efficient organization of the language. It helps maintaining the number of exceptions below a critical threshold.

The challenge is to find out how children do to identify productive and exceptional patterns in language and acquire them accordingly. Yang (2016) suggests that language learning corresponds to a gradual productive generalization of rules where productivity is evaluated according to a tolerance principle motivated by the indication that “…children are remarkably adept at keeping exceptions from spilling over to the core of the grammar” (p.15). Yang is categorical in stating that “children occasionally extend productive rules, as appropriate, but almost never generalize lexical forms through analogy” (ibidem). However, he does not supply convincing data for confirming his negative claim.

The execution of the tolerance principle goes as follows. One starts with a rule, count the number of items in a distribution that meet the structural description of the rule and those who do not, and compare the number of exceptions to a critical threshold. If the number of exceptions is lower than the threshold, the productivity of the rule is confirmed (possibly only temporarily; new data could oblige to restart the operation). Generalization takes place after a sufficiently high amount of positive evidence has been accumulated. If not confirmed, the rule is dropped and the learner looks for another one in the rule space provided by universal grammar. An (irregular) inflectional rule may be, for example, “Change rime to *ought*.” It takes care of the past tense of such verbs as *think*, *buy*, *bring*, or *seek*. Another one may be: ”Shorten the vowel before suffixation,” accounting for the past-tense formation of such verbs as *lose* (*lost*), *sleep* (*slept*), *keep* (*kept*), or *feel* (*felt*).

Taking into account the additional processing time determined by exceptions and particularly those regarding frequent items in language use (according to Zipf’s law; see Chapter 1),Yang (2016) defines the tolerance principle as the ratio of N, the number of items (types not just tokens) to which the rule is applicable, to In(N), its natural logarithm[[31]](#footnote-31). This measure is also applicable to phonological and syntactic learning, guiding the learners toward the morphosyntax of their language; although Yang’s essay mostly document grammatical-morphological learning. Yang admits: “At the present time, it is not clear how the tolerance principle is executed as a cognitive mechanism of learning: surely children doesn’t (sic) use calculators. Conceivably, the calibration of productivity is implemented via the real-time competition of rules and lexical listing…” (2016, p.64).

 In his (2016) book, Yang repetitively dwells upon what he calls the sparsity of data available to language learners (even writing of “highly impoverished input”, p.216), a modern version of the poverty of stimulus argument in the generative tradition. He supplies several analyses from language corpora addressed to language-learning children borrowed from the childes world library, which appear to confirm his position[[32]](#footnote-32). The most extensive of these corpora cover only a few hundred hours of verbal interaction between adults and children as opposed to thousands of hours during several years in common situations of children’s language exposure. There is no guarantee then that these few hundred hours of documented interaction are truly representative of the composition of the total input to children.

 In a section entitled “Smaller is better” (2016, pp. 66-67), Yang explains that for mathematical reason the tolerance principle works better with smaller samples; that is, the acquisition of a productive rule is easier when N is lower, which follows from the way the tolerance threshold is computed. He goes on suggesting that a young child with a limited vocabulary has a better chance of learning the rules of language than an adult who generally knows more words. This could be dubbed Yang’s paradox: i.e., the idea that: “…the sparsity of language is much more of a blessing than a curse” (p. 226). It turns upside down the logic of the poverty of stimulus argument, promoting it to a developmental advantage.

 Yang adds that this indication could explain the (alledged) fact that adults are poor rule learners in comparison to children. This is a curious reasoning. First, there is no proof that this is really a fact rather than just a popular idea. The learning backgrounds and contexts of adults and children vary in many respects that cannot be adequately controlled in experimental settings. Second, regarding language in particular, adults have been children in the first place. They may be supposed to already know “the rules” of their first language. In natural second-language learning, the situation is not different, it would seem, from that of children in what regards the quantity of vocabulary items adults are exposed to at a given learning time.

 Yang’s (2016) suggestions may have the merit of originality but are unrealistic. Children are supposed to compute complex data over limited periods of exposure to language input, navigate in a space of abstract universal[[33]](#footnote-33) rules (which Yang states explicitly are “inevitable” as the need for “an overarching linguistic theory” cannot be put into question) with no conscious knowledge of these rules, compute proportions of items meeting each rule or not, decide whether keeping the rule or dropping it, and, if the latter, restart the whole process again until reaching rule productivity (at least for the core aspects of language). As ultimate rampart, Yang recalls to mind the Chomskyan belief according to which the uniformity of language acquisition in all languages supports the idea that there exists a specific general biological capacity for language (Chomsky, 1965), with no indication as to whether it should be considered representational or procedural.

 Analogical generalization probably remains a more realistic candidate for an explanation of the acquisition of grammatical morphology. It is known that analogies play an important role in nominalization morphology for the forging of new words in all languages. Why would this general process suddenly disappear from language learning at the level of grammatical morphology?

Analogical treatment functions whether the input is satisfactorily saturated in the structures to be learned or not. The learner opportunely uses whatever (s)he has at hand, exploiting full regularities, partial regularities, and common patterns identifiable in sets of irregular forms; developing incorrect (default or so-called unproductive) generalizations or omitting them altogether. Correspondingly, learning is complete or incomplete (from a linguistic point of view) and may remain so for a given time or even definitively in first- or more often in second-language acquisition, depending on the available input and its characteristic distributional statistics. Last but not least, an analogical account allows dispensing with the gratuitous need for a pre-existing rule-space in which learners would dig following the computation of a tolerance principle based on critical thresholds of which they have no knowledge.

 **7.3. Devoted memory systems**

 Tulving (1972, 1983) proposed to distinguish episodic and semantic memories. Episodic memory is the memory for life events and episodes situated in a particular temporal and spatial framework. Actually this memory corresponds to a cognitive reconstruction of these episodes on the basis of partial indications that are memorized but can be modified in time and under the effect of more recent life episodes. Semantic memory stores general knowledge on the world: factual, verbal, visual and spatial. This kind of memory is less depending than episodic memory on the context of origin and is the object of a generalization. Squire and Zola (1998) argue that episodic and semantic memories constitute two parallel sub-systems of a more general form of memory labelled declarative memory (referring to the capacity to remember newly learned information). They may interact. For example, episodic memory may provide an entrance door to semantic memory.

 The anatomic substrates of the two forms of memory differ. Episodic memory relies on a vast cerebral network involving internal temporal, diencephalic, thalamic, and hypothalamic regions in addition to various prefrontal regions. The neural substrate of semantic memory involves the medial-temporal lobes (and particularly the hippocampal region) and diencephalic structures (Nyberg & Cabeza, 2000).

 There is also the need to distinguish an explicit (intentional) and an implicit treatment and storing of the information (Graf & Schacter, 1985). Declarative memory (i.e., episodic + semantic memory) is considered to be an explicit system. An implicit non-declarative system regroups a series of registers: procedural skills and habits, priming[[34]](#footnote-34) and perceptual learning, emotional and skeletal responses, and non-associative learning, i.e. simple reflex pathways (Ardila, 2012).

Double dissociations between the two systems have been documented in developmental and adult-onset language disorders (e.g., brain lesions, Alzheimer disease, Parkinson disease, Korsakoff syndrome, acute depressive states). Implicit memory involves the frontal (including Broca’s and several other premotor areas), parietal and superior-temporal cortices of the left brain, the left-basal ganglia (most importantly the striatum with the caudate nucleus and the putamen), and the right-neocerebellar structures. Explicit memory depends on the integrity of the prefrontal cortex and relies on a series of medial-temporal lobe structures on both sides of the brain. These include the hippocampal and parahippocampal regions as well as the anterior cingulated cortex. Research in brain hemodynamics, electrophysiology, and magneto-encephalography has revealed the existence of different mechanisms involved in the two kinds of memory at the neuronal and molecular levels (Kandel, 2005).

The above indication concern long-term retention or long-term memory. A system of short-term retention or short-term memory was identified long ago (Chapter 2). It has been reconceptualized under the name of working memory from the seminal work of Alan Baddeley (1990). This system is sub-divided into an auditory-vocal and a visual-spatial component. The organic substrate involves a series of neural structures located bilaterally within the prefrontal cortex.

 As alluded to in Chapter 7, Ullman (2004) has proposed a dual neuropsychological model of memory and language learning labelled declarative/procedural model (already introduced in Ulman et al., 1997). In this model, lexical knowledge as well as irregular grammatical morphology is considered to be stored in declarative memory. In opposition, syntactic learning, regular grammatical morphology, and phonology depend on procedural memory. Procedural learning plays a major role in identifying context-dependent relations between elements or events in real-time sequences. Learning occurs on an ongoing basis during multiple presentations. The knowledge acquired applies automatically to corresponding or analogically-related material. At the neural level, automation of behavior corresponds to interactive and recurrent loops between cortical, cerebellar, and subcortical structures. The right cerebellum in connection with the left-cerebral hemisphere is acknowledged as playing an important role in a broad language network (Fabbro, 1999; Vias & Dick, 2017).

Ullman (2004) suggests that his declarative memory system is closely related to the “ventral stream” system proposed by Goodale and Milner (1992; also Milner & Goodale, 2008). This system was initially conceptualized for visual objects (recognition, identification and storage in long-term memory). It has been extended to speech and language representations (Hickok, 2012; Hickok & Poepel, 2004).The ventral stream projects ventro-laterally toward the inferior-posterior-middle-temporal cortex which serves as an interface between sound-based representations of speech in the superior-temporal gyrus and widely distributed conceptual representations.

 Correspondingly, Ullman (2004) indicate that his procedural system is closely related to the “dorsal stream” identified by Goodale and Milner (1992), rooted in posterior-parietal cortex and connected to frontal-premotor regions. The dorsal pathways in this conception might be particularly relevant for phonological and syntactic processing, as it projects dorso-posteriorly involving a region in the posterior-Sylvian fissure at the parietal–temporal boundary, and ultimately projecting to frontal regions (Hickok, 2012). A close interpretation has been proposed recently by Arbib (2017).

Ullman (2004) also makes his, Squire and Zola’s (1996) indication that the rules in the procedural system are rigid, inflexible, and not influenced by other mental systems. This system, unlike declarative memory, appears to be “informationally encapsulated” (p.237); an idea (modified and adapted from Fodor, 1983)..

Ullman (2004) has reviewed neuroimaging and electrophysiological studies of normal language processing as well as from developmental and adult-onset language- (e.g., aphasia) and non-language disorders (e.g., Alzheimer’s disease). He argues that this evidence supports his declarative/procedural model. In particular, numerous double dissociations have been documented suggesting that the two systems are largely independent of each other although interacting in important respects and forming a dynamic network (Eichenbaum & Cohen, 2001).

In opposition, Paradis (2004, see also Paradis, 2009) insists that the two systems do not actually interact. He indicates: “Declarative knowledge does not admit of implicit proceduralization” (p.41). But Paradis acknowledges that the two systems are related and may substitute for each other: “A computational procedure automatically generates an output that may become the object of declarative knowledge” (ibidem).

Ullman’s basic interactive position may appear to be contrary to the notion of informational encapsulation proposed as an intrinsic characteristic of the procedural system. It need not be perhaps. The two systems may have their proper characteristics, including informational encapsulation for procedural memory, and yet be able to interact “on neutral ground”, complement and substitute each other to some extent in acquiring analogous knowledge, including knowledge of sequences (Ullman, 2004, p. 243), as independent self-contained systems.

 Interestingly, Ullman (2004) suggests that the same or similar language knowledge can be acquired by both systems: “… superior aspects of the temporal lobe may play a role in the storage of knowledge about procedural memory related relations of structures representations” (p. 247). He adds: “The rapid lexical/declarative storage of sequences of lexical forms should provide a database from which grammatical rules can gradually and implicitly be abstracted by the procedural memory system” (ibidem).

Ullman reviews neuroimaging experiments in healthy adults demonstrating interactions between the two memory systems at brain level. In these experiments, procedural learning is shown to activate the caudate nucleus while deactivating the medial temporal lobe substrate of declarative memory. Intersubject comparisons showed that the degree of activity in the caudate nucleus correlates negatively with the degree of activation of the medial-temporal lobe. This suggests that individuals may vary with their respective dependencies on the two systems.

The theory presented in this book is globally in agreement with Ullman’s (2004) suggestions, with two reservations, however. First, Ullman equates semantics with lexicon (not an unusual stance in neuropsychological writing on the topic). As illustrated in this book, lexical semantics is only a part of the semantic component. Another, most important, aspect is relational semantics whose role in morphosyntactic functioning has been stressed in the present essay. It has to be conceptualized as belonging to the procedural system. Semantic relations correspond to abstract recipes regulating morphosyntatic patterns. They are expressed in sentence surface through the medium of words which depends on declarative memory and combinatory sequences that may be consciously scrutinized; however, intrinsically, semantic relations are procedural matters. Second, Ullman has adopted the traditional view that mental grammar is governed by rules unknown to language users; a conception rejected in the present essay.

 A last consideration is in order regarding working memory. As illustrated, short-term memory plays an important role in language functioning. Baddeley’s (1986, 1990, 2012) model of working memory[[35]](#footnote-35) is an expanded version of the classical short-term memory device. Usual accounts of working memory are based on the assumption that it operates on information represented consciously. The tasks used for assessing it are clearly of the explicit processing type. Current theory suggests that the encoding of information in working memory, maintenance, retrieval, and use in decision making, all proceed from conscious contents.

Auditory-vocal (sometimes labelled phonological-) working memory is often viewed as a necessary ingredient of language acquisition, including regarding its morphosyntactic aspects (Gathercole & Baddeley, 1993). However, if, as argued, morphosyntactic learning is a form of implicit learning, it should rely more upon implicit rather than explicit memory. Relating auditory-vocal working memory and morphosyntactic development in nondisabled children, one may observe a discrepancy in the respective time scales. The working memory span represents “the number of items of whatever length that can be uttered in about two seconds” (Baddeley, 1990, p.74). This span is slow to increase in children, in part for neurological maturational reasons. The process of refreshing the information in order to keep it alive in the system (through internal recoding or private speech) as well as the existence of more efficient encoding strategies, is not operational before five or six years (Gathercole et al., 2004). By that time and even before, morphosyntactic development is already well under way.

Implicit memory is better off in this respect because it matures neurologically before explicit memory (Thomas et al., 2004) and continues to develop with age as a function of augmented knowledge base (Murphy, Mckone, & Slee, 2003). There has been some suggestion and empirical indications that working memory may also be engaged by incidental exposure to visual or auditory items. For instance, Hassin et al. (2009) argue that working memory can operate outside of conscious awareness and without conscious intention for active maintenance of information for relatively short periods of time and context relevant updating of information. In experimental tasks involving perceptual judgement of visual patterns according to the usual methodology in implicit learning research, Hassin et al. (2009) have provided findings suggestive of the existence of nonconscious implicit memorizing. Corresponding data have been reported for implicit memory of verbal material (words and nonwords) by McKone (1995). Evidence seems to accumulate in favor of the existence of an implicit working memory challenging the view that working memory processes are contingent only on conscious awareness (Soto, Mäntylä, & Silvanto, 2011). It appears that unconscious contents may be maintained within working memory and that complex cognitive computations may be performed online (Velichkowsky, 2017). The matter, however, remains insufficiently investigated.

 **7.4. Developmental language disorders**

Several pathological conditions may provide natural experiments in morphosyntactic implicit learning. Natural experiment is intended here in Bronfenbrenner’s (1979) sense of a particular disorder caused by an unfortunate conjunction of natural causes that besides remediation supplies an interesting situation for testing hypotheses about, for example, the course of language development. These hypotheses would seldom be amenable to experimental manipulation by design because they would be unethical under normally planned conditions. This is the case for specific language impairment in otherwise normally developing children and several genetic syndromes determining an important degree of cognitive handicap.

 **7.4.1. Specific language impairment**

 Specific language impairment (SLI) is a developmental language disorder whose prevalence is estimated to be around 4% in 5-year-old children (Bishop & Edmunson, 1987; Tomblin et al., 1997). It is diagnosed in children not developing language normally but without cognitive handicap, autism spectrum disorder, apraxia, acquired brain damage, hearing loss, or emotional and social problems. As documented in Chapter 3, research has showed that the impairment has a genetic etiology running in families. It can result from a single-gene mutation or the combined influence of several gene variants (Plomin & Dale, 2000; Bishop, 2002).

The most common form of SLI is the receptive/expressive phonological-syntactic variety (or grammatical-SLI; SLI in what follows). Major characteristics involve short and syntactically-simplified sentences, misuse of pronouns, and omission of grammatical-morphological features. Word comprehension is usually spared (Weismer & Hesketh, 1996), but retrieval of lexical knowledge (e.g., word finding) may be problematic in many individuals (Weckerly, Wulfeck, & Reilly, 2001). Early pragmatic and relational semantic functioning appears to be preserved, however, in spite of sometimes severe morphosyntactic deficits (Arosio et al., 2017).

 Recent research work throws a light on the likely underpinning of this deficiency as a weakness regarding morphosyntactic implicit learning.

 Desmottes, Meulemans, and Maillart (2016a) utilized an experimental paradigm called serial search task, with 24 typically developing (TD) and 24 children with SLI. The TD and SLI children were matched for chronological age (7 to 12 years) and nonverbal intellectual quotient. Concrete bisyllabic words were used illustrated by pictures. The pictures were presented in a 2x2 design on a computer screen and the arrangement of the four pictures varied for each trial. Children were instructed to locate the picture depicting each auditory- presented word with a touchscreen. Response time and accuracy were the dependent variables. After completing the last sequence in the learning task, an interview was made for probing a possible declarative knowledge of the sequences in the participants. Children then were informed of the existence of regular sequences and requested to reproduce the sequences to which they had been exposed in a series of other trials.

 None of the children either TD or with SLI demonstrated explicit knowledge of the sequence regularities. Data from the reaction times indicated that TD children implicitly learned both the spoken words as well as the motor sequences. In contrast, a majority of children with SLI gave no indication of implicit learning of the sequence regularities in the motor as well as in the verbal task. These data support the hypothesis of an amodal limitation in implicit sequence learning in children with SLI. However, in a comparable experiment, Gabriel et al. (2014) reported similar sequence learning in a visual and auditory serial research task in TD and children with SLI. The important interindividual variability in children with SLI registered in the two studies may explain, at least partially, the discrepancy in the results obtained.

 Desmottes et al. (2016a) analyzed the product-moment correlation between the performance of their subjects on the serial search task and language tests assessing, lexical, grammatical, and phonological abilities. Higher scores on the serial search task (verbal as well as motor learning) were significantly associated with better morphosyntactic abilities in TD children but only with lexical abilities in children with SLI.

 In another study, Desmottes, Meulemans, and Maillart (2016b) used a slightly different research paradigm called visual-motor serial reaction time. In this technique, a stimulus appears successively in one of several possible locations on a computer screen. Participants are requested to press a button upon detection of the stimulus and reaction times are measured. Unknown to the participants, there is a sequence involved in the succession of the stimuli. A decrease in reaction times is considered to be an indication that they have detected a pattern in the sequential display even if transitional rules cannot be verbalized. Desmottes et al. (2016b) observed that most children with SLI showed a beginning of sequence learning during the early stages of the experiment. As training blocks progressed, however, only TD children improved their sequence knowledge. Moreover, children with SLI failed to exhibit consolidation gains at 24- hour and one-week time intervals following the end of the learning session, as did TD children.

 These data are interpreted by the authors as globally supportive of the theoretical suggestion of Ullman and Pierpont (2005). This suggestion is based on Ullman’s (2004) neuropsychological model (summarized in the preceding section). Ullman and Pierpont (2005) argue that SLI corresponds in major part to a fundamental deficit in procedural memory and that it is not specific to language, impacting also the motor domain. Hsu and Bishop (2014) also found that SLI affects the treatment and maintenance of nonverbal as well as verbal information. This suggest that persons with SLI have a dysfunction affecting the basal ganglia (in particular the striatum) leading to a difficulty in the implicit treatment of sequences of items, linguistic as well as nonlinguistic, and in detecting and learning input regularities.

 In contrast, according to Ullman and Pierpont’s (2005) model, declarative memory is expected to remain intact in children with SLI. This has been confirmed in a study of Lum et al. (2011) in which measures of (explicit) working memory, declarative and procedural memory in 51 children with SLI and 51 TD children matched for chronological age (10 years). Working memory was assessed with the Working Memory Test Battery for Children, declarative memory with the Children's Memory Scale, and procedural memory with a visual-spatial serial reaction time task. Results showed that compared to TD children, children with SLI were impaired in procedural memory, even when holding working memory constant. In contrast, as predicted by Ullman and Pierpont’s model, they were spared in declarative memory for visual information and in verbal declarative memory after controlling for working memory and level of language development.

 The assessment of visual-spatial working memory suggested that this function was intact in children with SLI, whereas verbal working memory was impaired, even when language deficits were held constant.

 Lum et al. (2011) computed a series of product-moment correlations on their data. They revealed that neither form of explicit working memory was associated with either lexical or morphosyntactic abilities in children with SLI or in TD children. Declarative memory measures correlated significantly with lexical abilities in both groups of children. As expected, morphosyntactic abiity was associated with procedural memory in TD children but not in children with SLI. Only declarative memory correlated with morphosyntactic ability in children with SLI.

 Ullman (2004) stressed the relevance of his dual model for the clinical rehabilitation of individuals with SLI. The possible interaction between declarative and procedural systems indeed suggests that people with disorders of the grammatical/procedural system could benefit from a rehabilitation strategy based on memorization of complex morphosyntactic forms in declarative memory. Such strategy could presumably be supplemented with other therapeutic approaches including neuropharmacology (Ullman & Pierpont, 2005).

 **7.4.2. Cognitive handicap**

Morphosyntactic difficulties are common place in genetic syndromes of cognitive handicap (Chapman, 2003; Rondal, 2009; for reviews and analyses). Although there appears to be noticeable differences between particular syndromes (cf. Rondal & Perera, 2006), let us recall the major morphosyntactic limitations in the language of persons with Down syndrome (DS). They are the best known to date and may serve as a general guide line in the matter.

As to syntax, one generally observes in the language of persons with DS (outside of exceptional cases as reported in Chapter 2): (1) a lesser comprehension and reduced and unstable use of articles, prepositions, auxiliaries, copulas, pronouns, and conjunctions even when they are matched with TD peers on mental age; (2) no syntactic comprehension of reversible passives; (3) no comprehension of temporal sentences with clause order not matching the order of ongoing events; (4) fewer reversals of order of grammatical subject and copula or auxiliary verb in interrogative sentences; and (5) a reduced use and limited understanding of longer sentences, compound, and overall complex sentences with subordinate clauses.

Regarding grammatical morphology, the observations indicate typically: (1) a reduced use of verbal inflections; frequent omissions include contracted *will*, contracted *am*, contracted *is*, third-person singular, regular past-tense, and progressive form in -*ing*; (2) a reduced and unstable marking of number on nouns; (3) omission of number agreement between grammatical subjects and verbs.

However, in their vast majority, children with DS develop language to a large extent including regarding basic morphosyntactic aspects. Word order in sentences, for example, is usually correct). Importantly, when they combine words in sentences, persons with DS express the same range of relational meanings as reported in the language development of TD children. They also understand correctly the same set of basic of semantic relations when they are realized in the speech of other people (Rondal & Edwards, 1997). However, as indicated above, morphosyntactic development remains incomplete most of the time. A remaining question is why.

The cognitive handicap of these children cannot, in itself, be a satisfactory explanation for their morphosyntactic difficulties (although it may contributes to the problem), given that, as illustrated in Chapter 2, cases of advanced morphosyntactic abilities in several adolescents and adults with Down syndrome and regular cognitive levels for the condition, have been documented.

Neither language input to children with Down syndrome can be at stake. It has long been established that this input is appropriate in quality and quantity for language acquisition. Parental language input to children with DS is equivalent in all respects to the language addressed by parents to TD children when these and the children with DS are matched for level of language development (Rondal, 1978; Rondal & Docquier, 2006).

 Rondal and Guazzo (2012; see also Rondal, 2017) have suggested that the morphosyntactic limitations in persons with DS stem in major part from a deficit in implicit learning and procedural memory. This hypothesis may perhaps be extended to related conditions of cognitive handicap with a genetic etiology (see below).

 Mean total brain volume in children and young adults with DS is smaller than in controls (by around 18%), mostly due to an underdevelopment (hypoplasia) of the frontal lobes, reduction in superior-temporal white matter volume, smaller planum temporale, flattening of the occipital lobes, reduction of the cerebellum to approximately 30% of its normal volume, and a 20% reduction in density of the hypothalamic neuron nuclei (Wisniewsky, Kida, & Brown, 1996; Wisniewsky et al., 2006). Nadel (2006) has summarized a series of high-resolution PET scan and MRI investigations confirming the preceding indications and reporting parietal and hippocampal volume reduction as well as evidence of dysfunction in basal ganglia in persons with DS.

 Perceptual apprehension of sequential units in the input is involved in implicit learning. Visual perceptual deficits have been reported in persons with DS (Wang et al., 2015). A mild to moderate permanent hearing loss, mostly of the transmission type tied to abnormal middle-ear status, affects an important proportion of children with DS (25%) according to a recent survey by Nightengale et al. (2017). These shortcomings could complicate the extraction of relevant material from the input in implicit learning.

In DS, sustained attention is better preserved for visual stimuli than for auditory ones (Faught, Conners, & Himmelberger, 2016). A weakness in reducing attentional focus to more limited sequences of units in the input may be involved in the restricted processing of verbal data demonstrated in these subjects. In the same way, the limitations documented in auditory short-term (working) and long-term memory in DS (Conners, 2003; Carlesimo, Marotta, & Vicari, 1997) may probably render less efficient the processing of encoding relevant incidental information and recollecting it for later use (Witt & Vinter, 2013b).

Several experimental studies with participants with cognitive handicap of genetic etiologies have been carried out to test the robustness of implicit learning and procedural memory across the spectrum of intellectual variability. The experimental protocols were adapted in various ways in order to conform to the reduced cognitive abilities of the participants (see below for an illustration). Results indicate that implicit learning is a robust paradigm in the sense that its general principles apply to persons with cognitive handicap as well as to TD subjects. However, a number of differences have also been documented suggesting that the former subjects present limitations in some aspects of implicit learning and procedural memory.

 A thorough investigation is that of Witt (2011). He compared a sample of 120 TD children divided in age groups between 5 and 8 years with a sample of children with cognitive handicap (of organic origin; etiologies not supplied in the report), intelligence quotients between 50 and 70 points; chronological ages: 9 years and 6 months to 10 years and 1 month; mental ages: 5 years and 8 months to 6 years and 4 months. The finite-state automaton in Witt’s series of experiments generated sequences of 3, 4 or 5 colored flags representing teams of small pandas in a computerized video game of cord drawing presented in series of 8 sequences. Three types of series of colors were used: in one series, the sequential grammar allowed the successive repetitions of adjacent colors (e.g., blue-yellow-yellow; blue-yellow-yellow-green; red-green-green-yellow-blue; and so on); in another series, the grammar allowed only the successive repetition of nonadjacent colors (e.g., blue-yellow-blue; red-green-yellow-red; blue-yellow-green-yellow-blue; and so on). A third series presented the colors in random serial order and served as a control condition. In the test phase, the participants were informed that in the second day of the tournament, the organizers had forgotten to place the colors on the flags which, as a consequence, had remained white. Participants were invited to set series of 3, 4, or 5 colors on corresponding cardboards, selecting the colors one by one from a random color display. The instruction was simply to produce beautiful flags in a plausible attempt not to focus participants’ attention onto simply reproducing flags seen during the learning phase. The test phase was followed by a short debriefing session in the form of a preformulated and standardized questionnaire in which the experimenter asked the children whether they knew why they had been invited to play that particular game. The objective was to verify whether the participants had developed a conscious knowledge of the rules of the game and, if yes, whether they could verbalize them correctly.

Results showed that children with a cognitive handicap demonstrate an ability to learn sequential relations implicitly. This indication attests to preserved capabilities of implicit learning in these children. It confirms previous reports with children, adolescents, young and aging adults with cognitive handicap with and without DS (Wyatt & Conners, 1998; Atwell, Conners, & Merrill, 2003; Vinter & Detable, 2003). It is also congruent with Reber’s postulate regarding the robustness of basic aspects of the implicit learning paradigm across ages and intellectual levels (as mentioned in Section 7.1).

 Witt’s data are in opposition to the indications of Fletcher, Maybery, and Bennett (2000) suggesting that implicit learning would be largely inefficient below an intellectual quotient of 60 or so and/or a mental age around 6 years. It should be added, though, that Fletcher et al.’s experiments have been criticized on the ground of having too much of an explicit learning component. It is known that persons with cognitive handicap practically always score lower than their TD peers in explicit learning tasks.

 However, in Witt’s (2011) experiment, children with cognitive handicap, contrary to their mental age-matched TD peers, appeared to be sensitive only to adjacent repetitions of pairs of elements. No positional information was coded mentally. They seemed to be restricted to identifying only the perceptually more prominent characteristics of the stimuli.

 Bussy et al. (2011) used the technique of visual-motor serial reaction time with adolescents with Fragile-X syndrome (FXS)[[36]](#footnote-36), DS, and TD children, all matched for mental age. It was interesting to involve participants with FXS in the experiment given that they present anatomical abnormalities of the striatum, one of the basal ganglia involved in implicit memory. A statistically significant decrease in reaction times was recorded in the FXS group between the random blocks and some blocks with repeated sequences (not all, however, due to an important within group variability). In the DS group, the reaction times decreased significantly between the blocks with repeated sequences and the random ones except when there was an interfering random block located between two blocks with repeated sequences.

 A few studies have investigated the implicit learning ability of subjects with Williams syndrome (WS). Schellenberg, et al. (2003) investigated implicit learning in 27 children and adults with WS matched for chronological age with TD subjects (ages between 9 and 50 years). They used a form of the standard artificial grammar learning paradigm. Results indicated that the WS subjects are able to implicitly learn the characteristics of a simple artificial grammar although at a lower level than TD individuals.

 Vicari (2001) compared implicit and explicit memory processes in 12 children with WS, 14 with DS, and 32 TD children, all matched for mental age (close to 6 years and 6 months). Tests of verbal and visual implicit and explicit memory were administered (e.g., serial reaction time; a stem completion task in which the subjects had to complete words to which they had been exposed inadvertently before; and free recall of a list of unrelated words). Children with DS and TD controls demonstrated comparable implicit memory abilities but expectedly TD participants obtained higher scores in the explicit learning tasks. Children with WS demonstrated lower explicit memory abilities than TD controls. They proved also less efficient in some implicit memory tasks particularly the serial reaction time.

 In conclusion to this still premature research literature, it would appear that persons with an important cognitive handicap of genetic origin do not differ markedly from TD people in some of the implicit learning tasks utilized in the experimental studies mentioned. They are capable of implicit learning but often at a lower level of performance than their TD peers. It may be indeed, as argued by some theorists in the domain, that general cognitive ability and intellectual level are not primary explanatory variables when it comes to the ability for implicit learning. Additional research work with more systematic and balanced research designs is needed for moving ahead in this field.

 An indication worth to be stressed is Witt’s (2011) demonstration that children with cognitive handicap can implicitly learn sequential dependencies between adjacent units in pairs of stimuli but not nonadjacent ones or even adjacent dependencies beyond pairs of stimuli. In other words, they are limited to detecting and memorizing only simple, immediate relationships, and relationships that are perceptually closer and located within shorter attention spans. They are also unable to encode positional information on the items in visual displays. Data reported by Bussy et al. (2011) are convergent with Witt’s findings. Adolescents with DS could implicitly learn sequential relations in a task of serial reaction time. However, when a block with random presentation of the stimuli was placed in between the blocks with repeated sequences, learning dropped to nil. This suggests a weakness in maintaining relevant memory traces in implicit learning and resisting intervening distractors.

 The above data, although clearly preliminary, appear to go in the theoretical direction suggested above. In particular, children with DS, and possibly other genetic condition conducive to an important degree of cognitive handicap, are capable of acquiring (albeit with marked delays) a basic morphosyntax allowing them to produce and understand relatively short and structurally simple sentences as well as a kind of elementary (and somewhat chaotic from the point of view of the system) grammatical morphology. Most of the time, however, they do not progress much beyond that level of development. It may be hypothesized that their limitations in implicit learning and procedural memory contribute to their difficulties in morphosyntactic learning and development. Future research on this question should help clarifying the matter and perhaps lead to the definition of more efficient rehabilitation strategies (Rondal & Guazzo, 2012).

 CONCLUSION

Recapitulating the major points of the present analysis, it may be reckoned that there exist two systems with the capacity of supporting natural morphosyntactic functioning: a formal and a semantic-relational one. They may substitute each other but also interact for optimizing morphosyntactic functioning.

Practically from its foundation as a discipline, theoretical psycholinguistics has leaned on the morphosyntactic conceptions developed in structural and generative linguistics without, almost never for a long time, seriously questioning these conceptions, which claimed to account not only for the way idioms may be described grammatically but also for how people proceed when producing and understanding sentences. This was a curious collective stand if one considers that it amounted to letting a discipline, psycholinguistics, whose object is language behavior and representations in bone and flesh human beings, be dictated its contents and objectives by another discipline, a methodologically distant albeit conceptually related one, theoretical linguistics.

 Several reasons for this course of events have been analyzed. One of them, maybe the major one, has to do with the elegant and powerful new formulations in theoretical linguistics that appeared in the nineteen fifties and sixties under the name of generative and transformational linguistics, a more psychologically oriented version of structural linguistics. These formulations seem to have fascinated leading cognitive psychologists of the time to such an extent as to motivating them to embrace uncritically the new linguistic orientation and decide that the psycholinguists’ major task was indeed to find ways of integrating the linguists’ intuitions regarding grammatical competence into some plausible psychological account. However, the experimental attempts to verify the theoretical predictions issuable from generative grammar have mostly failed as illustrated in Chapter 2, and numerous observational data gathered in psychological research including developmental psycholinguistics do not support the generative theories.

A particular aspect of generative linguistics inherited from traditional structural linguistics but forcefully maintained, was (and still is to a large extent) the *a priori* conception that semantics is a subsidiary component of the linguistic system restricted to interpreting the output of the morphosyntactic one. To the excuse, so to speak, of structural linguists, it may be recalled that semantics, and particularly relational) semantics, had remained an underdeveloped domain (often confused with lexical semantics) until the early nineteen sixties (but, of course the relative disinterest of structural linguistics for semantic matters also contributed to its lack of significant development). From that time on, interest arose considerably for the communicative and functional aspects of language, and in particular its pragmatic and semantic foundations.

Another major aspect of generative linguistics was its reliance on the notion of universal grammar and its ontological corollary the notion of linguistic representational innateness, still accepted by a number of present-day psycholinguists. This notion had the advantage, so to speak, to dispense the discipline with the task of searching for and empirically validating a genuine theory of grammatical development.

 It is probably correct to say that in recent years generative linguistics has lost a good deal of its attractive power among psycholinguists (and among linguists as well). The concept of universal grammar has been reworked several times. It is still debated and even rejected by some linguists. The hypothesis of an innate prefiguration of key grammatical notions appears to be deprived of molecular genetic support despite continuing theoretical affirmations to the contrary. What is coming to the front instead, is the realization that the genetic and related brain machineries underlie a whole set of mechanisms devoted to language treatment but not specific to it. As a consequence, the notion of grammatical specificity with respect to the rest of the cognitive apparatus, long a pilaster of generative theorizing, is now being abandoned even by its proponents.

So, “the generative king is naked,” but where do we stand and in which direction should we plan to go in the explanation of morphosyntactic functioning and development?

 It may reasonably be supposed that sentences are treated on line (not exclusively nor even mostly linearly, however) and self-edited. Natural morphosyntax is learned as no clear representational prefiguration has been documented. It is learned implicitly as regular speakers/hearers have no conscious knowledge of the grammatical dispositions prevailing in their community idiom. It is constructed incidentally by children from appropriate and adapted input through the setting of systematic correspondences between semantic relations and morphosyntactic patterns, and the discovery of productive analogies. There is no conscious intention to learn the grammatical aspects of the idiom in natural language acquisition, which is different from explicit and metalinguistic learning. It follows that a correct theory of morphosyntactic functioning and acquisition must be part of a general theory of implicit learning, which means also that the devoted memory systems have much to do with procedural memory.

I have proposed that semantic relations provide the first system for the generation and regulation of natural morphosyntax. *First*, here, has to be understood as qualifying as system that develops chronologically in the first place and may remain the most important one all life on in language users. We now have sufficiently detailed semantic structural theories to account for morphosyntactic regulation. Semantic relations, it is worth recalling, are also abstractions, just as the linguistic categories summarizing them. It may be considered, however, that they are more accessible to the learners as they stem from an interface with general cognition.

It seems that there exists genuine brain representations of the distributional equivalents of linguistic formal and functional categories (although their precise nature and many aspects still remain unclear pending more efficient neuroimaging and related techniques of investigation). There is no guarantee, however, that they are neither complete with regard to the linguistic provisos nor even that they are actually involved in the morphosyntactic treatment of regular language users. There may be important individual differences between people in this respect. These formal brain representations may provide a second system for the regulation of natural morphosyntax, independent from but possibly interacting with the semantic-relational one. This second system needs also to be learned for the same reason as recalled above. It must derive from exposure to language input. Experimental research in connectionist and probabilistic approaches to grammar suggest that a distributional learning of at least some linguistic formal and functional categories is possible mediating powerful extraction, computation, and memory devices. Transposing to the natural context, it may hypothesized that such a learning is possible only following a relatively long exposure to the distributional properties of the language input, and, in particular for children, in taking advantage of the systematic correspondence between semantic and morphosyntactic categories in parental input as documented in the context of the semantic bootstrapping hypothesis. These are the reasons why the formal system has to be conceptualized as a second morphosyntactic system and not, as traditionally considered the first or the only one.

 In this respect, psycholinguistics can be said to have constructed its morphosyntactic theories in the reverse way from what appears today; that is, defining the formal system as the first or even the only relevant one for organizing morphosyntax; curiously not realizing that as language being pretty much a matter of exchanging semantic information in pragmatically motivated frameworks, it would have been surprising indeed if this semantic priority would not be reflected also in the basic organization of the morphosyntactic component.

 I also argued that natural morphosyntatic functioning is fundamentally sequential in its treatment, allowing speakers to self-edit their production. However, as discussed, sequential processing need not be assimilated to a simple linear device. Sequences can integrate hierarchical loops and sophisticated distal dependencies between items. In that sense, the opposition traditionally entertained between sequences and hierarchies in many theoretical discussions, is probably exaggerated. There is no logical necessity to define the two concepts as being orthogonal to each other.

 BIBLIOGRAPHY

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, *23*, 275-290.

Abend, O., Kwiatkovski, T., Smith, N., Goldwater, S., & Steedman, M. (2017). Bootstrapping language acquisition. *Cognition*, *164*, 116-143.

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*, 119-161.

Ambridge, B. (in press). Syntactic categories in child language acquisition: Innate, induced or illusory. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science*. Amsterdam: Elsevier.

Ambridge, B., Kidd, E., Rowland, C., & Theakston, A. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*, 239-273.

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge, UK : Cambridge University Press.

Ambridge, B., & Lieven, E. (2015). A contructivist account of child language acquisition. In B. MacWhinney & W. O’Grady (Eds.), *Handbook of language emergence* (pp. 478-510). Hoboken, NY: Wiley Blackwell.

Ambridge, B., Pine, J., & Lieven, E. (2014). Child language acquisition: Why universal grammar doesn’t help. *Language*, *90*, e53-e90.

Ambridge, B., Pine, J., & Lieven, E. (2015). Explanatory adequacy is not enough: Response to commentators on “Child language acquisition: Why universal grammar does not help.” *Language*, *91*, e116-e126.

Anderson, J., & Dubois-Charlier, F. (Eds.). (1975). La grammaire des cas. *Langages*, *38*, entire number.

Arbib, M. (2002). Beyond the mirror system: Imitation and the evolution of language. In C. Nehamiv & K. Dantenkam (Eds.), *Imitation in animals and artefacts* (pp. 2229-280). Cambridge, MA: MIT Press.

Arbib, M. (2017). Dorsal and ventral streams in the evolution of the language-ready brain: Linking language to the world. *Journal of Neurolinguistics*, *43b*, 228-253.

Ardila, A. (2012). Interaction between lexical and grammatical language systems in the brain. *Physics of Life Review*, *9*, 198-214.

Arnaud, A., & Lancelot, C. (1660). *Grammaire générale et raisonnée contenant les fondements de l’art de parler, expliqués d’une manière claire et naturelle*. Paris : Le Petit.

Arosio,F., Foppolo, F., Pagliarini, E., Perugini, M., & Guasti, M. (2017). Semantic and pragmatic abilities can be spared in Italian children with SLI. *Language Learning and Development*, *13*, 418-429.

Atwell, J., Conners, F., & Merrill, E (2003). Implicit and explicit learning in young adults with mental retardation. *American Journal of Mental Retardation*, *108*, 55-68.

Bach, E., & Harms, R. (Eds.). (1968). *Universals in linguistic theory*. New York: Holt, Rinehart & Winston.

Baddeley, A. (1986). *Working memory*. Oxford, UK: Oxford Science Publications.

Baddeley, A. (1990). *Human memory*. Hillsdale, NJ: Erlbaum.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Baker, G., & Hacker, P. (1984). *Language, sense and nonsense*. Oxford, UK: Blackwell.

Bannard, C., & Mattews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children’s repetitions of four-word combinations. *Psychological Science, 19*, 241-248.

Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of adult speech which predict children’s language development.  *Journal of Child Language*, *10*, 65-84.

Baron-Cohen, S. (1991). The theory of mind deficit in autism: How specific is it? *British Journal of Developmental Psychology*, *9*, 301-314.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind ? *Cognition*, *21*, 37-46.

Barsky, R. (1997). *Noam Chomsky. A life of dissent*. Cambridge, MA : MIT Press.

Barsky, R. (2010). *Zeilig Harris : From American linguistics to socialist Zionism*. Cambridge, MA: MIT Press.

Bates, E. (1976). *Langage in context: The acquisition of pragmatics*. New York: Academic Press.

Bates, E., & MacWhinney, B. (1982). A functionalist approach to grammatical development. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173-218). Cambridge, UK: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157-193). Hillsdale, NJ: Erlbaum.

Beaudichon, J., Sigurdsson, T., & Trelles, C. (1978). Étude chez l’enfant de l’adaptation verbale à l’interlocuteur. *Psychologie Française*, *23*, 213-220.

Begg, I. (1971). Recognition memory for sentence meaning and wording. *Journal of Verbal Learning and Verbal Behavior*, *10*, 176-181.

Bellugi, U., Marks, S., Bihrle, A., & Sabo, H. (1988). Dissociation between language and cognitive functions in Williams syndrome. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 177-189). London: Churchill Livingstone.

Benveniste, E. (1950). La phrase nominale. *Bulletin de la Société de Linguistique de Paris, 132*, 151-167.

Bergelson, E., & Aslin, R. (2017). Semantic specificity in one-year-olds’ word comprehension. *Language Learning and Development*, *13*, 481-501.

Berko, J. (1958). The child’s learning of English morphology. *Word*, *14*, 150-157.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

Berlingeri, M., Crepaldi, D., Roberti, R., Scialfa, G., Luzzatti, C., & Paulesu, E. (2007, January). *Nouns and verbs in the brain: Grammatical class and task specificity effects as revealed by fMRI*. Paper presented at the 25th European Workshop on Cognitive Neuropsychology, Bressanone, Italy (unpublished).

Berndt, R., Mitchum, C., Haendiges, A., & Sandon, J. (1997). Verb retrieval in aphasia. I. Characterizing single word impairments. *Brain and Language*, *56*, 68-106.

Berwick, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Berwick, R., & Chomsky, N. (2016). *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Berwick, R., & Chomsky, N. (2017). Why only us: Recent questions and answers. *Journal of Neurolinguistics*, *43b*, 166-177.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York: Wiley.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, *7*, 173-188.

Bickerton, D. (2009). Recursion: Core of complexity or artifact of analysis. In T. Givon & M. Shibatani (Eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution* (pp. 531-544). Amsterdam: Benjamins.

Bierwisch, M., & Schreuder, R. (1993). From concepts to lexical items. In W. Levelt (Ed.), *Lexical access in speech production* (pp. 23-60). London: Blackwell.

Bloomfield, L. (1933). *Language*. Chicago, IL: University of Chicago Press.

Bishop, D. (2002). The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders*, *35*, 311-328.

Bishop, D., & Edmunson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, *52*, 156-173.

Blumenthal, A. (1967). Prompt recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, *6*, 203-206.

Bock, J. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*, 1-47.

Bock, J., & Cutting, J. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99-127.

Bock, J., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego, CA: Academic Press.

Bolinger, D. (1968). Judgements of grammaticality. *Lingua*, *21*, 34-40.

Bolinger, D. (1975). *Aspects of language*. New York: Harcourt, Brace & Jovanovich.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123-172). Boston, MA: Reidel.

Bouchard, D. (1988). French *voici*/*voilà* and the analysis of pro-drop. *Language*, *64*, 89-100.

Bowerman, M. (1973). *Early syntactic development*. Cambridge, UK: Cambridge University Press.

Braine, M. (1963). The ontogeny of English phrase structure: The first phrase. *Language*, *39*, 1-13.

Braine, M. (1971). On two types of model of the internalization of grammar. In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 153- 186). New York: Academic Press.

Braine, M. (1976). Children first word combinations. *Monographs of the Society for Research in Child Development*, Serial No. 164.

Braine, M. (1987). What is learned in acquiring word classes? A step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65-87). Hillsdale, NJ: Erlbaum.

Braine, M., & Wells, R. (1978). Case-like categories in children. The actor and some related categories. *Cognitive Psychology*, *10*, 100-122.

Brame, M. (1976). *Conjectures and refutations in syntax and semantics*. New York: North-Holland Publishers.

Bransford, J., Barclay, J., & Franks, J. (1972). Sentence memory: Constructive vs. interpretive approach. *Cognitive Psychology*, *3*, 193-209.

Bredart, S., & Rondal, J.A. (1982). *L’analyse du langage chez l’enfant. Les activités métalinguistiques*. Bruxelles: Mardaga.

Brenowitz, E., & Beecher, M. (2005). Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neuroscience*, *28*, 127-132.

Brent, M., & Siskind, J. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81*, B33–B44. doi:10.1016/S0010-0277(01)00122-6.

Bresnan, J. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 1-59). Cambridge, MA: MIT Press.

Broen, P. (1972). The verbal environment of the language-learning child. *American Speech and Hearing Association Monographs*, Serial No 17.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-215). Hillsdale, NJ: Erlbaum.

Brooks, P., Tomasello, M., Lewis, L., & Dodson, K. (1999). Children’overgeneralization of fixed transitivity verbs. *Child Development*, *70*, 1325-1337.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Brown, R. (1986). *Social psychology. The second edition*. New York: Free Press.

Brown, R., & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology*, *49*, 454-462.

Bruner, J. (1975a). The ontogenesis of speech acts. *Journal of Child Language*, *2*, 1-20.

Bruner, J. (1975b). From communication to language. A psychological perspective. *Cognition*, *3*, 255-287.

Bussy, G., Charrin, E., Brun, A., Curie, A., & Des Portes, V. (2011). Implicit procedural learning in Fragile X and Down syndrome. *Journal of Intellectual Disability Research*, *55*, 521-528.

Bybee, J. (2002). Sequentiality as the basis of constituent structure. In T. Givon & B. Malle (Eds.), *The evolution of language out of pre-language* (pp. 109-134). Amsterdam: Benjamins.

Caplan, D. (1992). *Language structure processing and disorders*. Cambridge, MA: MIT Press.

Caplan, D. (1995). The cognitive neuroscience of syntactic processing. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 871-879). Cambridge, MA: MIT Press.

Cappa, S., & Perani, D. (2003). The neural correlates of noun and verb processing. *Journal of Neurolinguistics*, *16*, 183-189.

Caramazza, A. (1991). *Issues in reading, writing and speaking. A neuropsychological perspective*. Boston, MA: Kluwer.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177-208.

Caramazza, A., & Hillis, A. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*, 788-790.

Carlesimo, G., Marotta, L., & Vicari, S. (1997). Long-term memory in mental retardation: Evidence for a specific impairment in subjects with Down’s syndrome. *Neuropsychologia*, *35*, 71-79.

Carroll, J.B., & Casagrande, J. (1958). The function of language classifications in behavior. In E. Maccoby, T. Newcomb, & E. Hartley (Eds.), *Readings in social psychology* (pp. 18-31). New York: Holt, Rinehart & Winston.

Cashon, C., Ha, O., Graf Estes, K., Saffran, J., & Mervis, C. (2016). Infants with Williams syndrome detect statistical regularities in continuous speech. *Cognition, 154*, 165-168.

Chafe, W. (1970). *A semantically based sketch of Onondaga*. Bloomington, IN: Indiana University Publications in Anthropology and Linguistics (Memoir N°25).

Chafe, W. (1973). *Meaning and the structure of language*. Chicago, IL: University of Chicago Press.

Chafe, W. (2002). Searching for meaning in language. A memoir. *Historiographica Linguistica*, *29*, 245-261.

Champagnol, R. (1993). *Signification du langage*. Paris: Presses Universitaires de France.

Chapman, R. (2003). Language and communication in individuals with Down syndrome. In L. Abbeduto (Ed.), *International review of research in mental retardation* (pp. 1-34). New York, NY: Academic Press.

Chaudhari, D., Damani, O., & Laxman, S. (2011). Lexical co-occurrence, statistical significance, and word association. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processsing* (pp. 1058-1068). Edinburgh, UK: Association for Computational Linguistics.

Chi, J., Doaling, E., & Gilles, F. (1977). Left-right asymmetries of the temporal speech areas in the human fetus. *Archives of Neurology*, *34*, 346-348.

Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10*. Cambridge, MA: MIT Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague, The Netherlands: Mouton.

Chomsky, N. (1959). Review of *Verbal behavior* by B.F. Skinner. *Language*, *35*, 26-58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1966). *Cartesian linguistics: A chapter in the history of rationalist thought*. New York: Harper & Row.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt, Brace & World.

Chomsky, N. (1969). *La nature formelle du langage*. Paris: Seuil.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184-221). Boston, MA: Ginn.

Chomsky, N. (1971). Deep structure, surface structure and semantic interpretation. In D. Steinberg & L. Jakobovitz (Eds.), *Semantics* (pp. 183-216). New York: Cambridge University Press.

Chomsky, N. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Chomsky, N. (1975). *The logical structure of linguistic theory*. Chicago, IL: The University of Chicago Press.

Chomsky, N. (1976). *Reflections on language*. London: Fontana.

Chomsky, N. (1979). *Essays on form and interpretation*. Amsterdam: North-Holland.

Chomsky, N. (1980). *Rules and representations*. Oxford, UK: Blackwell.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, N. (1984). *Modular approaches to the study of mind*. San Diego, CA: San Diego State University Press.

Chomsky, N. (1986). *Knowledge of language*. Its nature, origin and use. New York: Praeger.

Chomsky, N. (1988). *Language and problems of knowledge*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (1997). Preface to J.Y. Pollock, *Langage et cognition. Introduction au programme minimaliste de la grammaire générative* (p. XIII-XIX). Paris : Presses Universitaires de France.

Chomsky, N. (2001). *Beyond explanatory adequacy.* Cambridge, MA: MIT Press.

Chomsky, N. (2007). Approaching UG from below.In U. Sauerland & H. Gartner (Eds.), *Chomsky’s minimalism and the view from syntax-semantics* (pp. 1-29). New York: Mouton de Gruyter.

Chomsky, N. (2011). Language and other cognitive systems. What is special about language? *Language Learning and Development, 7*, 263-278.

Chouinard, M., & Clark, E. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, *30*, 637-669.

Christiansen, M;, Kelly, M., Shillcock, R., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition*, *116*, 382-393.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, *51*, 61-75.

Clark, H., & Clark, E. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.

Cleeremans, A. (1993). *Mechanisms of implicit learning: A connectionist model of sentence processing.* Cambridge, MA: MIT Press.

Clifton, P., & Odom, P. (1966). Similarity relations among certain English sentence constructions. *Psychological Monographs*, Serial N° 613.

Conners, F. (2003). Phonological working memory difficulty and related interventions. In J.A. Rondal & S. Buckley (Eds.), *Speech and language intervention in Down syndrome* (pp. 31-49). London: Whurr.

Cordemoy, G. de (1668, 1968). *Discours physique de la parole.* Paris : Lambert (republished in *Cahiers pour l’Analyse*, Bibliothèque du Graphe, Paris, Seuil).

Costermans, J. (1980). *Psychologie du langage*. Brussels : Mardaga.

Craik, F., & Massani, P. (1969). Age and intelligence differences in coding and retrieval of word lists. *British Journal of Psychology*, *63*, 315-319.

Crystal, D. (1987). *Child language, learning and linguistics*. London: Arnold.

Curtiss, S. (1982). Developmental dissociations of language and cognition. In L. Obler & L. Menn (Eds.), *Exceptional language and linguistics* (pp. 285-312). New York: Academic Press.

Curtiss, S. (1988). *The special talent of grammar acquisition*. In L. Obler & D. Fein (Eds.), The exceptional brain (pp. 364-386). New York: Guilford.

Dabrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*. Published online. doi: [10.3389/fpsyg.2015.00852](https://dx.doi.org/10.3389/fpsyg.2015.00852)

Dabrowska, E., & Szczerbinsky, M. (2006). Polish children’s productivity with case marking: The role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, *33*, 559-597.

Damasio, A., & Tranel, D. (1993). Nouns and verbs are retrieved with differentially distributed neural systems. *Proceedings of the Natural Academy of Sciences of the United States of America*, *90*, 4957-4960.

Darwin, C. (1871). *The descent of man*. London: Murray.

Davis, M. (1994). *Emil L. Post: His life and work*. Bâle, Switserland: Birkhaüser.

Deacon, H., Pacton, S., & Conrad, M. (2008). A statistical learning perspective on children’s learning about graphotactic and morphological regularities in spelling. *Canadian Psychology*, *49*, 118-124.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, *12*, 449-460.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*, 2013-2015.

Dell, G., Chang, F., & Griffin, Z. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*, 517-542.

Demetras, M., Post, K., & Snow, C. (1986). Feedback to first language learners: The role of repetitions and clarification questions. *Journal of Child Language*, *13*, 275-292.

Denison, S., & Xu, F. (2014). The origin of probabilistic inference in human infants. *Cognition, 130*, 335-347.

Dennett, D. (1994). Review of Ray Jackendoff, *Patterns in the Mind: Language and Human Nature*, Harvester-Wheatsheaf, 1993, and Steven Pinker, *The Language Instinct: How the Mind Creates Language*, Morrow, 1993. *The London Review of Books*, *16* (12), 10-11.

Desmottes, L., Meulemans, T., & Maillart, C. (2016a). Implicit spoken words and motor sequences learning are impaired in children with specific language impairment. *Journal of the International Neuropsychological Society*, *22*, 520-529.

Desmottes, L., Meulemans, T., & Maillart, C. (2016b). Later learning stages in procedural memory are impaired in children with specific language impairment. *Research in Developmental Disabilities*, *48*, 53-68.

De Villiers, P., & De Villiers, J. (1972). Early judgements of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research*, *1*, 299-310.

De Villiers, J., & De Villiers, P. (1978). *Language acquisition*. Cambridge, MA : Harvard University Press.

De Villiers, J., & De Villiers, P. (1974). Competence and performance in child language : Are children competent to judge? *Journal of Child Language*, *1*, 11-22.

Dik, S. (1991). Functional grammar. In F. Droste & J. Joseph (Eds.), *Linguistic theory and grammatical description* (pp. 247-274). Amsterdam: Benjamins.

Ding, N., Melloni, L., Tian, X., & Poeppel, D. (2016). Rule-based and word-level statistics-based processing of language: Insight from neuroscience. *Language, Cognition and Neuroscience*. Advance online publication. doi:10.1080/23273798.2016.1215477

Duric, M. (1985). *A grammar of Acehenese*. Dordrecht, The Netherlands: Foris.

East, M., & Hofer, H. (1991). Loud calling in a female-dominant mammalian society: Behavioral context and functions of whooping of spotted hyaenas. *Animal Behaviour*, *42*, 651-669.

Egger, E. (2017). *Apollonius Dyscole: Essai sur l’histoire des théories grammaticales dans l’antiquité*. Paris : Forgotten books.

Eichenbaum, H., & Cohen, N. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York: Oxford University Press;

Eldredge, N. (1985). *Time frames: The rethinking of Darwinian evolution and the theory of punctuated equilibria*. New York: Simon & Schuster.

Ellis, A. (1980). On the Freudian theory of speech errors. In V. Fromkin (Ed.), *Errors in linguistic performance* (pp. 123-132). New York: Academic Press.

Ellis, N. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.),  *Implicit and explicit learning of languages* (pp. 3-23). Amsterdam: Benjamins.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, *14*, 179-211.

Elman, J. (1991). Distributed representation, simple recurrent network, and grammatical structure. *Machine Learning*, *7*, 195-225.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smit, A., Parisi, D., & Plunkett, K. (1997). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Elman, J., & Zipser, D. (1988). Learning the hidden structure of speech. *Journal of the Acoustical Society of America*, *8*, 1615-1626.

Emonds, J. (1985). *A unified theory of syntactic categories*. Dordrecht, The Netherlands: Foris.

Ervin-Tripp, S. (1966). Language development. In M. Hoffman & L. Hoffman (Eds.), *Review of child development research* (Vol. 2, pp. 55-105). Ann Arbor, MI: University of Michigan Press.

Ervin-Tripp, S. (1971). An overview of theories of grammatical development. In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 189-212). New York: Academic Press.

Estes, W. (1970). *Learning theory and mental development.* New York: Academic Press.

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429-494.

Evans, V. (2014). *The language myth. Why language is not an instinct*. Cambridge, UK: Cambridge University Press.

Everett, D. (2005). Cultural constraints on grammar and cognition in Pirahan language. *Current Anthropology*, *46*, 621-646.

Everett, D. (2007, March). Cultural constraints on grammar in Pirahan: A reply to Nevins, Pesetsky, and Rodrigues. *LingBuzz*, from http://ling.auf.net/lingBuzz/000427

Everett, D. (2008). *Don’t sleep, there are snakes. Life and language in the Amazonian jungle*. New York: Pantheon.

Everett, D. (2009). A response to some criticisms. *Language*, *85*, 405-442.

Everett, D. (2012). *Language: The cultural tool*. New York: Vintage Press.

Everett, D. (2017). Grammar came later: Triality of patterning and the gradual evolution of language. *Journal of Neurolinguistics*, *43b*, 133-165.

Fabbro, F. (1999). *The neurolinguistics of bilingualism*. Hove, UK: Psychology Press.

Faught, G., Conners, F., & Himmelberger, Z. (2016). Auditory and visually sustained attention in Down syndrome. *Research in Developmental Disabilities*, *53*, 135-146.

Fetzer, A. (2004). *Recontextualizing context: Grammaticality meets appropriateness*. Amsterdam: Benjamins.

Fillmore, C. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 1-87). New York: Holt, Rinehart & Winston.

Fillmore, C. (1977). The case for case reopened. In P. Cole & J. Sadock (Eds.), *Syntax and semantics* (Vol.8, pp. 59-81). New York: Academic Press.

Fitch, W., Friederici, A., & Hagoort, P. (2012). Pattern perception and computational complexity. *Philosophical Transactions of the Royal Society of Britain*, *367*, 1925-1932.

Fitch, T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*, 179-210.

Fitch, W., & Martins, M. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *The Year in Cognitive Neuroscience*, *1316*, 87-104.

Fitz, H. (2010). Statistical learning of complex questions. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2692-2698). Austin, TX: Cognitive Science Society.

Fletcher, J., Maybery, M., & Bennett, S. (2000). Implicit learning differences: A question of developmental level? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 26*, 246-52.

Flöel, A., De Vries, M., Scholz, J., Breitenstein, C., & Johansen-Berg , H. (2009). White matter integrity in the vicinity of Broca’s area predicts grammar learning success. *NeuroImage*, *47*, 1974-1981.

Fodor, J. (1966). How to learn to talk: Some simple ways. In F. Smith & G. Miller (Eds.), *The genesis of language* (pp. 105-128). Cambridge, MA: MIT Press.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. (1985). Introduction. In Précis of the modularity of mind. *Behavioral and Brain Sciences*, *8*, 1-5.

Fodor, J., Bever, T., & Garrett, M. (1974). *The psychology of language. An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.

Fouts, R. (1973). Acquisition and testing of gestural signs in four young chimpanzees. *Science*, *180*, 978-980.

Fouts, R., Fouts, D., & Van Cantfort, T. (1989). The infant Loulis learns signs from cross-fostered chimpanzees; In R. Gardner, B. Gardner, & T. Van Cantfort (Eds.), *Teaching sign language to chimpanzees* (pp. 178-207). Albany, NY: State University of New York Press.

François, J. (2014). L’émergence et l’évolution du langage humain du point de vue des neurosciences. *Corela* (cognition, représentation, langage), online : doi: 14.4000/corela.3629

Frank, S., Bod, R., & Christiansen, M. (2012). How hierarchical is language use? *Proceedings of the Royal Society of Britain*, online: doi:10.1098/rspb.2012.1741

Freud, S. (1901, 1975). *The psychopathology of everyday life*. Harmondsworth, UK: Penguin.

Freudenthal, D., Pine, J., & Gobet, F. (2005). On the resolution of ambiguities in the extraction of syntactic categories by chunking. *Cognitive Systems Research*, *6*(1), 17-25.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27-52.

Fromkin, V. (1993). Speech production. In J. Berko Gleason & Ratner, N. (Eds.); *Psycholinguistics* (pp. 272-301). Fort Worth, TX: Holt, Rinehart & Winston.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mother’s speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, *6*, 423-442.

Gabriel, A., Meulemans,T., Parisse, C., & Maillart, C. (2014). Procedural learning across modalities in French-speaking children with specific language impairment. *Applied Psycholinguistics*, *36*, 747-769.

Gardner, H. (1985). The centrality of modules. In Précis of the modularity of mind. *Behavioral and Brain Sciences*, *8*, 12-14.

Gardner, R., & Gardner, B. (1975). Early signs of language in child and chimpanzee. *Science*, *187*, 752-753.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol.1). *Speech and talk* (pp. 177-220). London: Academic Press.

Garrett, M. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.

Garrett, M. (1988). Processes in language production. In F. Newmeyer (Ed.), *Linguistics : The Cambridge survey (Vol.3). Psychological and biological aspects* (pp. 69-96). Cambridge, UK: Cambridge University Press.

Garrett, M. (1995). The structure of language processing: Neuropsychological evidence. In M. Gazzaniga (Ed.), *Cognitive neuroscience* (pp. 881-889). Cambridge, MA: MIT Press.

Gathercole, S., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.

Gathercole, S., Pickering, S., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177-190.

Gaux, C., & Gombert, J.E. (1999). La conscience syntaxique chez les préadolescents : question de méthodes. *L’Année Psychologique, 99*, 45-74.

Gazzaniga, M. (2008). *Human*. New York: HarperCollins.

Gleason, J., & Ratner, N. (2009). *The development of language*. Boston, MA: Pearson Education.

Gleitman, L. (1990). The structural sources of verb meaning.  *Language Acquisition, 1*, 3-55.

Gleitman, L., Gleitman, H., & Shipley, E. (1972). The emergence of the child as grammarian. *Cognition*, *1*, 137-164.

Godin, P. (1976). Aspects syntaxiques et sémantiques de la grammaire casuelle appliquée au français. *Cahiers de l’Institut de Linguistique de Louvain*, *3*, 116-164.

Gold, E. (1967). Language identification in the limit. *Information and Control, 10*, 4447-474.

Goldin-Meadow, S. (1984). Gestural communication in deaf children. *Monographs of the Society for Research in Child Development*, Vol. *49*, N°3/4.

Goldin-Meadow, S. (2005). *The resilience of language*. New York : Psychology Press.

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, *391*, 279-281.

Gombert, J.Z. (1990). *Le développement métalinguistique*. Paris : Presses Universitaires de France.

Gomez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431-436.

Gomez, R., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109-135.

Gomez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science*, *4*, 178-186.

Goodale, M., & Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*(1), 20-25.

Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition*, *39*, 1-50.

Gould, S. (1991). Exaptation. *Journal of Social Issues*, *47*, 43-65.

Gowie, J., & Powers, C. (1977, October). *A developmental study of grammatical acceptability in late childhood and early adolescence*. Paper presented at the Annual Boston University Conference on Language Development (unpublished).

Graf, P., & Schacter, D. (1985). Implicit and explicit memory for new associations in normal subjects and amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 164-178.

Graf Estes, K. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology*. Published online. doi:[10.3389/fpsyg.2012.00447](https://dx.doi.org/10.3389/fpsyg.2012.00447)

Grant, J., Valian, V., & Karmiloff-Smith, A. (2002). A study of relative clauses in Williams syndrome. *Journal of Child language*, *29*, 403-416.

Greenbaum, S. (1973). Informant elicitation of data on syntactic variation. *Lingua*, *31*, 201-212.

Greenbaum, S. (1976). Contextual influence on acceptability judgements. *International Journal of Psycholinguistics*, *6*, 5-12.

Greenberg, J. (1963). *Universals of language.* Cambridge, MA: MIT Press.

Grevisse, M., & Goosse, A. (2011). *Le bon usage. Précis de grammaire du français*. Louvain-la-neuve, Belgium : De Boeck Duculot.

Grimaldi, M. (2012). Toward a neural theory of language: Old issues and new perspectives. *Journal of Neurolinguistics*, *25*, 304-327.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. Baker & J. McCarthy (Eds.), *The logical problem of language acquisition* (pp.165-182). Cambridge, MA: MIT Press.

Grodzinsky, Y. (2003). Imaging the grammatical brain. In M. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 551-555). Cambridge, MA: MIT Press.

Grodzinsky, Y., & Amunts, K. (Eds.). ( 2006). *Broca’s region: Mysteries, facts, ideas, and history*. Oxford, UK: Oxford University Press.

Grodzinsky, Y., & Friederici, A. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology*, *16*, 240-246.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct object: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, *41*, 153-195.

Grossberg, S. (1987). Competitive learning. From interactive activation to adaptive resonance. *Cognitive Science*, *11*, 23-63.

Haegeman, L. (1994). *Introduction to Government and Binding theory*. Oxford, UK : Blackwell.

Hagège, C. (1976). *La grammaire générative. Réflexions critiques*. Paris : Presses Universitaires de France.

Hagège, C. (1982). *La structure des langues*. Paris: Presses Universitaires de France.

Hagège, C. (1986). *L’homme de parole*. Paris : Fayard.

Hagoort, P., & Levelt, W. (2009). The speaking brain. *Science*, *324*, 372-373.

Hakes, D. (1980). *The development of metalinguistic abilities in children.* Berlin: Springer Verlag.

Halle, M., Bresnan, J., & Miller, G. (1978. *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Halliday, M. (1985). *An introduction to functional grammar*. London : Arnold.

Harley, T. (2014). *The psychology of language. From data to theory*. Hove, UK: Psychology Press.

Harris, Z. (1951). *Methods in structural linguistics*. Chicago, IL: University of Chicago Press.

Harris, Z. (1970). *Papers in structural and transformational linguistics*. New York: Springer.

Haspelmath, M. (2007). Pre-established categories don’t exist: Consequences for language description and typology. *Linguistic Typology*, *11*, 119-132.

Hassin, R., Bargh, J., Engell, A., & McCulloch, K. (2009). Implicit working memory. *Consciousness and Cognition*, *18*, 665-678.

Hauser, M., & [Bever, T.](http://www.isrl.uiuc.edu/~amag/langev/author/tbever.html) (2008) A Biolinguistic Agenda. [*Science*](http://www.isrl.uiuc.edu/~amag/langev/pubtype/article_Science.html), *322*, 1057-1059.

Hauser, M., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569-1579.

Hauser, M., & Watumull, J. (2017). The universal generative faculty: The source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics*, *43b*, 78-94.

Herman, L., Richards, D., & Wolz, J. (1984). Comprehension of sentences by bottlenose dolphins. *Cognition*, *16*, 129-219.

Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding in the context of a dual-stream model. *Journal of Communication Disorders*, *45*, 393-402.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67-99.

Hills,T. (2013). The company that words keep: Comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language, 2013, 40*, 586-604.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown and Hanlon revisited: Mothers’ sensitivity to ungrammatical forms. *Journal of Child Language*, *11*, 81-88.

Hochberg, J., & Pinker, S. (1985). *Syntax-semantics correspondence in parental speech*. Unpublished manuscript, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA.

Hjelmslev, L. (1928). *Principes de grammaire générale*. Copenhagen: Host.

Hockett, C. (1942). English verb inflection. *Studies in Linguistics*, *1*(2), 1-8.

Hockett, C. (1958). *A course in modern linguistics*. New York: Mac Millan.

Honeck, R. (1971). A study of paraphrases. *Journal of Verbal Learning and Verbal Behavior*, *10*, 367-381.

Hopkin, M. (2005). Sign language reveals fast track to grammar. *Nature.com/news*. doi 050131/full/050131-4

Hopper, P., & Thompson, S. (1980). Transitivity in grammar and discourse. *Language*, *56*, 251-299.

Hsu, A., Chater, N., & Vitányi, P. (2011). The probabilistic analysis of language acquisition : Theoretical, computational, and experimental analysis. *Cognition*, *120*, 380-390.

Hsu, H., & Bishop, D. (2014). Sequence-specific procedural learning deficit in children with specific language impairment. *Developmental Science*, *17*, 352-365.

Huck, G., & Goldsmith, J. (1995). *Ideology and linguistic theory*. London: Routledge.

Hull, C. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

Humboldt, W. von (1836). *Über die Verschiedenheit des Menschen Sprachbaues*. Berlin: Königlichen Academie der Wissenschaften.

Hutchins, W. (1973). The generation of syntactic structures from a semantic base. Amsterdam: North-Holland Publishers.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2003). Language input and child syntax. *Cognitive Psychology*, *45*, 337-374.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Boston, MA: Reidel.

Hyams, N. (1987). The theory of parameters and syntactic development. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 1-22). Boston, MA: Reidel.

Ibbotson, P., & Tomasello, M. (2016). Language in a new key. *Scientific American*, *315*(5), 70-75.

Jackendoff, R. (1977). *X-bar syntax : A study of phrase structure*. Cambridge, MA : MIT Press.

Jackendoff, R. (1978). Grammar as evidence for conceptual structure. In Halle, M., Bresnan, J., & Miller, G. (Eds.), *Linguistic theory and psychological reality* (pp. 201-228). Cambridge, MA: MIT Press.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jaeggli, O., & Safir, K. (1987). *The null subject parameter*. Boston, MA: Reidel.

Jakubowicz, C. (1970). Recherches récentes en psycholinguistique. *L’Année Psychologique*, *70*, 247-293.

James, W. (1892, 1962). *Psychology : Briefer course*. New York: Collier.

Kalouguina, E., & Shi, R. (2013). Abstract rule learning in 11- and 14-month-old infants. *Journal of Psycholinguistic Research, 42*, 71-80.

Kandel, E; (2005). *In search of memory*. New York: Norton.

Kant, I. (1800, 1970). *Logique*. Paris: Vrin.

Kaplan, R., & Bresnan, J. (1985). Lexical-functional grammar: A formal system of grammatical representations. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 173-281). Cambridge, MA: MIT Press.

Karmiloff-Smith, A. (1979). *A functional approach to child language*. Cambridge, UK: Cambridge University Press.

Karmiloff-Smith, A., Brown, J., Grice, S., & Patterson, S. (2003). Dethroning the myth: Cognitive dissociations and innate modularity in Williams syndrome. *Developmental Neuropsychology*, *23*, 227-242.

Katz, J. (1981). *Language and other abstracts objects*. Totowa, NJ: Rowman & Littlefield.

Katz, J., & Fodor, J. (1963). The structure of a semantic theory. *Language*, *39*, 170-210.

Katz, J., & Postal, P. (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Kaynes, R. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony and development* (pp. 178-237). Cambridge, MA: MIT Press.

Kimura, D., & Watson, N. (1989). The relation between oral movement control and speech. *Brain and Language, 37*, 665-690.

Klein, R. (2017). Language and human evolution. *Journal of Neurolinguistics*, *43b*, 204-221.

Kline, M., & Demuth, K. (2010). Factors facilitating implicit learning: The case of the Sesotho passive. *Language acquisition, 17*, 220-234.

Korenberg, J., Chen, X., Hirota, H., Lai, Z., Bellugi, U., Burian, D., Roe, B., & Matsuoka, R. (2000). Genome structure and cognitive map of Williams syndrome. *Journal of Cognitive Neuroscience*, *12*, 89-107.

Krinsky-McHale, S., Devenny, D., Kitter, P., & Silverman, W. (2003). Implicit memory in aging adults with mental retardation with and without Down syndrome. *American Journal of Mental Retardation*, *108*, 219-233.

Kripke, S. (1982). *Wittgenstein, on rules and private language*. Oxford, UK: Blackwell.

Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*, 589-600.

Kuiper, K. (1996). *Smooth talkers. The linguistic performance of auctionneers and sportcasters*. Mahwah, NJ: Erlbaum.

Kupperberg, G., Caplan, D., Sitnikova, T., Eddy, M;, & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, *21*, 489-530.

Kupperberg, G., Holcomb, P., Sitnikova, T., Greve, D., Dale, A., & Caplan, D. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117-129.

Kurzweil, R. (2006). *The singularity is near. When humans transcend biology*. London: Duckworth.

Lagarde, J.P. (1988). Les parties du discours dans la linguistique moderne et contemporaine. *Langages*, *92*, 93-108.

Lahey, M., Liebergott, J., Chesnick, M., Menyuk, P., & Adams, J. (1992). Variability in children’s use of grammatical morphemes. *Applied Psycholinguistics*, *13*, 373-398.

Lai, C., Fisher, S., Hurst, J., Vargha-Khadem, F., & Monaco, A. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, *413*, 519-523.

Lai, J., & Poletiek, F. (2011). The impact of adjacent-dependencies and stage-input on the learnability of center-embedded hierarchical structures. *Cognition*, *18*, 265-273.

Lakoff, G. (1971a). On generative semantics. In D. Steinberg & L. Jakobovitz (Eds.), *Semantics* (pp. 232-296). New York: Cambridge University Press.

Lakoff, G. (1971b). Presupposition and relative well-formedness. In D. Steinberg & L. Jakobovits (Eds.), *Semantics* (pp. 329-340). New York: Cambridge University Press.

Lakoff, R. (1989). The way we were; or, the real actual truth about generative semantics: A memoir. *Journal of Pragmatics*, *13*, 939-988.

Lallot, J. (1988). Origines et développement de la théorie des parties du discours en Grèce. *Langages*, *92*, 11-23.

Lallot, J. (1998). *La grammaire de Denys le Thrace*. Paris : Conseil National de la Recherche.

Langacker, R. (1976). Semantic regulations and the linguistic relativity hypothesis. *Foundations of Language*, *14*, 307-357.

Langacker, R. (1987). *Foundations of cognitive grammar* (Vol. 1). Stanford, CA: Stanford University Press.

Langacker, R. (2008). *Cognitive grammar. A basic introduction*. Oxford, UK: Oxford University Press.

Langacker, R. (2010). Day after day after day. In F. Parrill, V. Tobin, & M. Turner (Eds.), *Meaning, form, and body* (pp. 149-164). Stanford, CA: Center for the Study of Language and Information Publications.

Lashley, K. (1951). The problem of serial order in behavior. In L. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: Wiley.

Lederer, R. (1989). *Crazy English*. New York: Pocket Books.

Legate, J., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, *14*, 315-344.

Leibniz, G. (1703, 1763, 1921). *Nouveaux essais sur l’entendement humain* (reprinted). Paris : Flammarion.

Lenneberg, E. (1964). A biological perspective on language. In E. Lenneberg (Ed.), *New directions in the study of language* (pp. 65-88). Cambridge, MA: MIT Press.

Leonard, L. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, CA: MIT Press.

Levelt, W. (1999). Producing spoken language: A blueprint of the speaker. In A. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). New York: Oxford University Press.

Levelt, W. (2008). *Formal grammars in linguistics and psycholinguistics* (Vol.3). Cambridge, MA: MIT Press.

Levinson, S. (1984). *Pragmatics*. Cambridge, UK: Cambridge University Press.

Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature neuroscience*, *6*, 1230-1237.

Lieven, E. (1994). *Crosslinguistic and crosscultural aspects of language addressed to the children*. In C. Gallaway & B. Richards (Eds.), Input and interaction in language acquisition (pp. 56-73). Cambridge, UK: Cambridge University Press.

Lin, F. (2017). A refutation of universal grammar. *Lingua*, *193*, 1-22.

Linden, L. (2007). *The basic theories of language acquisition*. Norderstedt, Germany: Grin Verlag.

Lorenz, K. (1937). The companion in the bird’s world. *The Auk*, *54*, 245-273.

Lucy, J. (1992a). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. (1992b). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. (1997). The linguistics of color. In C. Hardin & L. Maffi (Eds.), *Color categories in language and thought* (pp. 320-346). Cambridge, UK: Cambridge University Press.

Lum, J., Conti-Ramsden, G., Page, D., & Ullman, M. (2011). Working, declarative and procedural memory in

 specific language impairment. *Cortex*, *48*, 1138-1154.

Luria, A. (1961). *The role of speech in the regulation of normal and abnormal behaviour*. London: Pergamon.

Lyons, J. (1970). *Chomsky*. London: Fontana.

Lyons, J. (1977). *Semantics* (2 vol.). Cambridge, UK: Cambridge University Press.

Maclay, H. (1971). Overview. In D. Steinberg & L. Jakobovitz (Eds.), *Semantics* (pp. 157-182). New York: Cambridge University Press.

MacCorquodale, K. (1970). On Chomsky’s review of Skinner’s verbal behavior. *Journal of the Experimental Analysis of Behavior*, *13*, 83-99

Macnamara, J. (1982). *Names for things: A study of child language.* Cambridge, MA: MIT Press.

MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language*, *19*, 65-77.

MacWhinney, B. (2000). *The CHILDES project. Tools for analyzing talk.* Hillsdale, NJ: Erlbaum.

MacWhinney, B. (2010). Language development. In M. Bornstein & M. Lamb (Eds.), *Developmental science* (pp. 468-508). New York, NY: Psychology Press.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verbal learning model. *Cognition*, *40*, 121-157.

Malotki, E. (1983). *Hopi time: A linguistic analysis of temporal concepts in Hopi language*. Berlin: Mouton.

Mandelbrot, B. (1954). Structure formelle des textes et communication. *Word*, *101*, 1-27.

Mandelbrot, B. (1968). Information theory and psycholinguistics. In R. Oldfield & J. Marshall (Eds.), *Language* (pp. 263-275). Harmondsworth, UK: Penguin Books.

Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Maratsos, M. (1989). Innateness and plasticity in language acquisition. In M. Rice & R. Schiefelbusch (Eds.), *The teachability of language* (Vol.2, pp. 67-98). Baltimore, MD: Brookes.

Maratsos, M., & Chalkley, M. (1981). The internal language of children’s syntax: The ontogenesis of representation of syntactic categories. In K. Nelson (Ed.), *Children’s language* (pp. 127-214). New York: Gardner.

Maratsos, M., Fow, D., Becker, J., & Chalkley, M. (1985). Semantic restrictions on children passives. *Cognition*, *19*, 167-191.

Marchman, V. (1993). Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Science*, *9*, 215-234.

Marcus, G. (1993). Negative evidence in language acquisition. *Cognition, 46*, 53-85.

Marcus, G. (1999). Language acquisition in the absence of explicit negative evidence: Can simple recurrent networks obviate the need for domain-specific learning devices. *Cognition*, *73*, 293-296.

Marcus, G., Vijayan, S., Bandi Rao, S., & Vishton, P. (1999). Rule learning by seven month-old infants. *Science*, *283*, 77-80.

Marks, L, & Miller, G. (1964). The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behavior*, *3*, 1-5.

Martin, E, & Roberts, K. (1966). Grammatical factors in sentence retention. *Journal of Verbal Learning and Verbal Behavior*, *5*, 211-218.

Martin, L. (1986). “Eskimo words for snow:” A case study in the genesis and decay of an anthropological example. *American Anthropologist*, *88*, 418-423.

Martinet, A. (1953). Compte rendu de Knud Togeby, *La structure immanente de la langue française*. *Word*, *9*, 177-178.

Martinet, A. (1965). *La linguistique synchronique.* Paris : Presses Universitaires de France.

Maslen, R., Theakston, A., Lieven, E., & Tomasello, M. (2004). A dense corpus study of past tense and plural over- regularization in English. *Journal of Speech, Language, and Hearing Research*, *47*, 1319-1333.

Matthews, D., & Theakston, A. (2006). Errors of omission in English-speaking children’s production of plural and the past tense: The effects of frequency, phonology, and competition. *Cognitive Science*, *30*, 1027-1052.

McCawley, J. (1968). The role of semantics in a grammar. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 124-169). New York: Holt, Rinehart and Winston.

McCawley, J. (1971). Where do noun phrases come from? In D. Steinberg & L. Jakobovitz (Eds.), *Semantics* (pp. 217-231). New York: Cambridge University Press.

McCawley, J. (1976). *Grammar and meaning*. New York: Academic Press.

McClelland, J., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*, 465-472.

McClelland, J., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, *4*, 287-335.

McKone, E. (1995). Short term implicit memory for words and nonwords. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 1108-1126.

McNeil, D. (1966). Developmental psycholinguistics. In F. Smith & G. Miller (Eds.), *The genesis of language* (pp. 15-84). Cambridge, MA: MIT Press.

McNeil, D. (1971). The capacity for the ontogenesis of grammar. In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 17-40). New York: Academic Press.

Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, *2*, 346-351.

Mehler, J., & Boysson-Bardies, B. (1971). Psycholinguistique: message et codage verbal. *L’Année Psychologique*, *71*, 547-581.

Mehler, J., & Carey, P. (1967). Role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior*, *6*, 335-338.

Mervis, C., & Bertrand, J. (1995). Acquisition of the novel name-nameless category principle. *Child Development*, *65*, 1646-1662.

Mervis, C., Robinson, B., Rowe, M., Becerra, A., & Klein-Tasman, B. (2003). Language abilities in individuals with Williams syndrome. In L. Abbeduto (Ed.), *International review of research in mental retardation* (Vol.27, pp. 35-82). New York: Academic Press.

Meulemans, T. (1998). *L’apprentissage implicite, une approche cognitive, neuropsychologique et développementale*. Paris : Solal.

Miceli, G., Silveri, M., Villa, G., & Caramazza, A. (1984). On the basis for the agrammatics’ difficulty in producing main verbs. *Cortex*, *20*, 207-220.

Miles, H. (1990). The cognitive foundations for reference in a signing orangutan. In S. Taylor-Parker & K. Gibson (Eds.), *Language and intelligence in monkeys and apes. Comparative developmental perspectives* (pp. 511-539). New York: Cambridge University Press.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Miller, G. (1962). Some psychological studies of grammar. *American Psychologist*, *17*, 748-762.

Miller, G. (1965, 1968). Some preliminaries to psycholinguistics. *American Psychologist*, *20*, 15-20. Reprinted in R. Oldfield & J. Marshall (Eds.), *Language* (pp. 202-212), Harmondsworth, UK, Penguin Books.

Miller, G., & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, *2*, 217-228.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plan and the structure of behavior*. New York: Holt.

Miller, G., & McKean, K. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology*, *16*, 297-308.

Miller, G., & Selfridge, J. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology*, *16*, 297-308.

Milner, D., & Goodale, M. (2008). Two visual systems reviewed. *Neuropsychologia*, *46*, 774-785.

Milner, J.C. (1989). *Introduction à une science du langage*. Paris : Seuil.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*, 91-117.

Mittins, W., Salu, M., Edminson, M., & Coyne, S. (1970). *Attitudes to English usage*. London : Oxford University Press.

Moerk, E. (1975). Verbal interaction between children and their mothers during the preschool years. *Developmental Psychology*, *11*, 788-794.

Moerk, E. (1980). Relationships between parental input frequencies and children’s language acquisition: A reanalysis of Brown’s data. *Journal of Psycholinguistic Research*, *7*, 105-118.

Moerk, E. (1983). *The mother of Eve as a first-language teacher*. Norwood, NJ: Ablex.

Moerk, E. (1992). *First language taught and learned*. Baltimore, MD: Brookes.

Moerk, E. (2000). *The guided acquisition of first language skills*. Stamford, CT: Ablex.

Moore, T. (1972). Speeded recognition of ungrammaticality. *Journal of Verbal Learning and Verbal Behavior*, *11*, 550-559.

Moore, T. (1975). Linguistic intuitions of twelve year-olds. *Language and Speech*, *18*, 213-216.

Morgan , J;, & Newport, E. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*, 67-85.

Morin, Y. (1985). On the two French subjectless verbs *voici* and *voilà*. *Language*, *61*, 777-820.

Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, K., McKone, E., & Slee, J. (2003). Dissociations between implicit and explicit memory in children: The role of strategic processing and the knowledge base. *Journal of Experimental Child Psychology, 84*, 124-165.

Muysken, P. (1988). Are creoles a special type of language? In F. Newmeyer (Ed.), *Linguistics: The Cambridge survey* (Vol.2, pp. 285-301). New York: Cambridge University Press.

Nadel, L. (2006). Neuropsychological aspects of Down syndrome. In J.A. Rondal & J. Perera (Eds.), *Down syndrome. Neurobehavioural specificity* (pp. 67-83). London: Whurr.

Nagata, H. (1988). The relativity of linguistic intuition: The effect of repetition on grammaticality judgments. *Journal of Psycholinguistic Research*, 17, 1-17.

Nevins, I., Pesetsky, D., & Rodrigues, C. (2007, March). Pirahan exceptionality: A reassessment. *LingBuzz*, from <http://ling.auf.net/lingBuzz/000411>

Nevins, I., Pesetsky, D., & Rodriguez, C. (2009a). Pirahan exceptionality: A reassessment. *Language*, *85*, 355- 404.

Nevins, I., Pesetsky, D., & Rodrigues, C. (2009b). Evidence and argumentation: A reply to Everett (2009). *Language*, *85*, 671-681.

Newmeyer, F. (2004). Typological evidence and universal grammar. *Studies in Language, 28*, 527-548.

Newmeyer, F. (2005). *Possible and probable languages. A generative perspective on linguistic typology*. Oxford, UK: Oxford University Press.

Newport, E., Gleitman, L., & Gleitman, H. (1977). Mother, I would rather do it myself: Some effects and noneffects of maternal speech style. In C. Snow & C. Ferguson (Eds.), *Talking to children* (pp. 109-150). New York: Cambridge University Press.

Nightengale, E., Yoon, P., Wolter-Warmerdam, K., Dee, D., & Hickey, F. (2017). Understanding hearing and hearing loss in children with Down syndrome. *American Journal of Audiology*, *26*, 301-308.

Nyberg, L., & Cabeza, R. (2000). Brain imaging and memory. In E. Tulving & F. Craik (Eds.), *The Oxford handbook of memory* (pp. 501-519). Oxford, UK: Oxford University Press.

O’Connor, N., & Hermelin, B. (1988). Low intelligence and special abilities. *Journal of Child Psychology and Psychiatry*, *29*, 391-396.

O’Connor, N., & Hermelin, B. (1989). The memory structure of autistic idiot-savant mnemonists. *British Journal of Psychology*, *80*, 97-111.

O’Connor, N., & Hermelin, B. (1991). A specific linguistic ability. *American Journal of Mental Retardation*, *95*, 673-680.

Onnis, L., Waterfall, H., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*, 423-430.

Osgood, C. (1971). Where do sentences come from? In D. Steinberg & L. Jakobovits (Eds.), *Semantics* (pp. 497-529). New York: Cambridge University Press.

Osgood, C., & Sebeok, T. (1954). *Psycholinguistics, a survey of theory and research problems*. Baltimore, MD: Waverly Press.

Owings, D., & Leger, D. (1980). Chatter vocalizations in California ground squirrels. Predator- and social role-specificity. *Zeitscrift für Tierpsychologie*. *54*, 163-184.

Pääbo, S. (2014). *Neanderthal man: In search of lost genomes*. New York: Basic Books.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 80-96.

Pacton, S., Perruchet, P., Cleeremans, A., & Fayol, M. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*, 401-426.

Pacton, S., Perruchet, P., & Fayol, M. (2005). Children’s implicit learning of graphotactic and morpholgical regularities. *Child Development*, *76*, 324-339.

Pacton, S., Sobaco, A., & Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependencies learning valid? *Acta Psychologica*, *157*, 195-199.

Palermo, D. (1971). On learning to talk: Are principles derived from the learning laboratory applicable? In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 41-62). New York: Academic Press.

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 2522-2527.

Papagno, C., & Vallar, G. (2001). Understanding metaphors and idioms: A single-case neuropsychological study in a person with Down syndrome. *Journal of the International Neuropsychological Society*, *7*, 516-528.

Paquier, P;, & Mariën, P. (2005). A synthesis of the role of cerebellum in cognition. *Aphasiology*, *19*, 3-19.

Paradis, M. (2000, April). *Awareness of observable input and output, not of linguistic competence*. Paper presented at The International Symposium on Language Awareness, Ödense University, Denmark (unpublished).

Paradis, M. (2004). *Neurolinguistic theory of bilingualism*. Amsterdam: Benjamins.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.* Amsterdam: Benjamins.

Patel, A., Iversen, J., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca’s aphasia. *Aphasiology*, *22*, 776-789.

Patterson, F. (1980). The novative use of language by a gorilla : A case study. In K. Nelson (Ed.), *Children’s language* (Vol. 2, pp. 497-561). New York: Gardner.

Pechmann, T., Zerbst, D., & Garrett, M. (2000). The time course of recovery for grammatical category information during lexical processing for syntactic construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 723-728.

Perani, D., Cappa, S., Schnur, T., Tettamanti, M., Collina, S., Rosa, M., & Fazio, F. (1999). The neural correlates of verb and noun processing: A PET study. *Brain*, *122*, 2337-2344.

Perfors, A., Tenenbaum, J., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition, 118*, 306-338.

Perrot, J. (1953). *La linguistique*. Paris : Presses Universitaire de France.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language. Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 223-228.

Perruchet, P. (2005). Statistical approaches to language acquisition and the self-organizing consciousness: A reversal of perspective. *Psychological Research*, *69*, 316-329.

Perruchet, P., & Nicolas, S. (1998). L’apprentissage implicite: un débat théorique. *Psychologie Française, 43*(1), 13-25.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*, 233-238.

Perruchet, P., & Pacton, S. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 80-96.

Perruchet, P., & Poulin-Charronnat, B. (2015). The learnability of language. Insights from the implicit learning literature. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 139-165). Amsterdam: Benjamins.

Perruchet, P., Poulin-Charronnat, B., & Pacton, S. (2012). Learning nonadjacent dependencies. In N. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1941-1944). New York: Springer Verlag.

Perruchet, P., & Vinter, A. (1998). Learning and development: The implicit knowledge assumption reconsidered. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495-531). Thousand Oaks, CA: Sage Publications.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25,* 297-330.

Perruchet, P., Vinter, A., & Pacton, S. (2007). La conscience auto-organisatrice : une alternative au modèle dominant de la psychologie cognitive. *Éducation et Didactique*, *1*(3), 105-116.

Peterfalvi, J.M. (1970). *Introduction à la psycholinguistique*. Paris : Presses Universitaires de France.

Peterson, R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production : Evidence for cascade processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 539-557.

Petersson, K., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*, 83-95.

Pezzini, G., Vicari, S., Volterra, V., Milani, I., & Ossella, M. (1999). Children with Williams syndrome. Is there a single neuropsychological profile? *Developmental Neuropsychology*, *15*, 141-155.

Piaget, J. (1963). Le langage et les opérations intellectuelles. In Collectif, *Problèmes de psycholinguistique* (pp. 51-61). *Actes du Symposium de l’Association de Psychologie Scientifique de langue Française*. Paris : Presses Universitaires de France.

Piaget, J. (1968). *Le structuralisme*. Paris : Presses Universitaires de France.

Pickering, M., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*, 377-392.

Pierce, A. (1987, September). *Null subject in the acquisition of French*. Paper presented at the Annual Boston University Conference on Child language, Boston, MA.

Ping, L., Zhen, J., & Li, H.T. (2004). Neural representation of nouns and verbs in Chinese: An fMRI study. *NeuroImage*, *21*, 1533-1541.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1985). A theory of the acquisition of lexical interpretive grammars. In J. Bresnan(Ed.), *The mental representation of grammatical relations* (pp. 655-726). Cambridge, MA: MIT Press.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 399-441). Hillsdale, NJ: Erlbaum.

Pinker, S. (1989). *Learnability and cognition. The acquisition of argument structure.* Cambridge, MA: MIT Press.

Pinker, S. (1994). *The language instinct*. New York: Morrow.

Pinker, S. (1999). *Words and rules*. London: Weidenfeld & Nicolson.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York, NY: Viking Press.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*, 707-727.

Pinker, S., & Jackendoff, R. (2005). The faculty of language. What’s special about it? *Cognition*, *95*, 201-236.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*, 73-193.

Plante, E. (1991). MRI finding in the parents and siblings of specifically language impaired boys. *Brain and Language*, *41*, 67-80.

Plomin, R., & Dale, P. (2000). Genetics and early language development: A UK study of twins. In D. Bishop & L. Leonard (Eds.), *Speech and language impairment in children* (pp. 35-51). Hove, UK: Psychological Press.

Plunkett, K., & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, *23*, 463-490.

Popper, K. (1935, 1992). *The logic of scientific discovery*. London: Routledge.

Postal, P. (1968). Epilogue. In R. Jacobs & P. Rosenbaum, *English transformational grammar* (pp. 267-289). Waltham, MA: Blaisdell.

Postal, P. (1970). On the surface verb ”Remind”. *Linguistic Inquiry*, *1*, 37-120.

Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish nouns from verbs. *Neuroscience Letters*, *197*, 81-83.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. New York: Blackwell.

Principe, A., Calabria, M., Tauste Campo, A., Cruzat, J., & Conesa, G. (2017). Whole network, temporal and parietal lobe contributions to the earliest phases of language production. *Cortex*, *93*, 238-247.

Pulvermüller, F. (1999). Words in the brain’s language. *Behavioral and Brain Sciences*, *22*, 253-336.

Pulvermüller, F., Preissl, H., Lutzenberger, W., & Birbaumer, N. (1996). Brain rhythms of language: Nouns versus verbs. *European Journal of Neuroscience*, *8*, 937-941.

Quine, W. (1960). *Word and object*. Cambridge, MA: MIT Press.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford, UK: Blackwell.

Rao, P., Yadav, N., Vahia, M., Joglekar, H., Adhikari, R., & Mahadevan, I. (2009). Entropic evidence for linguistic structure in the Indus script. *Science*, 324, 1165.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, 45, 45-94.

Reber, A. (1967). Implicit learning of artificial grammars. *Journal of verbal* *Learning and Verbal Behavior, 6*, 855-863.

Reber, A. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81*, 219-235.

Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Reber, A. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York, NY: Oxford University Press.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*, 425-469.

Reeder, P., Newport, E., & Aslin, R. (2017). Distributional learning of subcategories in an artificial grammar: Category generalization and subcategory restrictions. *Journal of Memory and Language*, *97*, 17-29.

Richelle, M. (1976). Formal and functional analysis of verbal behavior: Notes on the debate between Chomsky and Skinner. *Behaviorism*, *4*, 209-221.

Rizzi, L. (1990). *An application of the parametric theory of subject acquisition: Some French data*. Unpublished manuscript, Department of Linguistics, University of Geneva, Geneva, Switserland.

Robertson, D., Davidoff, J., Davies, I., & Shapiro, L; (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General*, *133*, 554-571.

Robertson, D., Davies, I., Corbett, G., & Vandervyver, M. (2005). Freesorting of color across cultures: Are there universal grounds for grouping ? *Journal of Cognition and Culture*, *5*, 349-386.

Robertson, D., Pak, H., & Hanley, J. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*, 752-762.

Rohde, D., & Plaut, D. (1999). Simple recurrent networks can distinguish non-occurring from ungrammatical sentences given appropriate task structure: A reply to Marcus. *Cognition*, *73*, 297-300.

Rondal, J.A. (1978). Maternal speech to normal and Down’s syndrome children matched for mean length of utterance. In E. Myers (Ed.), *Quality of life in severely and profoundly mentally retarded people: Research foundation for improvement* (pp. 193-265). Washington, D.C.: American Association on Mental Deficiency.

Rondal, J.A. (1985). *Adult-child interaction and the process of language acquisition*. New York: Praeger.

Rondal, J.A. (1994a). Exceptional language development in mental retardation: Natural experiments in language modularity. *European Bulletin of Cognitive Psychology*, *13*, 427-467.

Rondal, J.A. (1994b). Computerized longitudinal database concerning one French-speaking child (between 27 and 57 months of age; MLU 2.25 - 4.50). *Childes Language Data Bank*. Pittsburgh, PA: Carnegie Mellon University. <http://childes.talkba>nk.org/data/Romance/French/rondal.zip.

Rondal, J.A. (1995). *Exceptional language development in Down syndrome. Implications for the cognition-language relationship*. New York : Cambridge University Press.

Rondal, J.A. (2000). *Le langage: de l’animal aux origines du langage humain*. Sprimont, Belgium : Mardaga.

Rondal, J.A. (2003). Atypical language development in individuals with mental retardation: Theoretical implications. In L. Abbeduto (Ed.), *International review of research in mental retardation* (Vol. 27, pp. 281-308). New York: Academic Press.

Rondal, J.A. (2006). *Expliquer l’acquisition du langage. Caveats et perspectives*. Sprimont, Belgium : Mardaga.

Rondal, J.A. (2009). *Psycholinguistique du handicap mental*. Marseille, France : Solal.

Rondal, J.A. (2014). *Une théorie du fonctionnement et du développement morphosyntaxique*. Paris : L’Harmattan.

Rondal, J.A. (2015). *Comment l’enfant apprend la grammaire de sa langue sans le savoir*. Paris : Edilivre.

Rondal, J.A. (2016). *D’où vient le langage humain ? Essai de reconstruction évolutive*. Liège, Belgium : Presses de l’Université de Liège.

Rondal, J.A. (2017). Morphosyntactic development and implicit learning in Down syndrome and related conditions. *Journal of Intellectual Disability Diagnosis and Treatment*, *5*, 24-30.

Rondal, J.A., & Cession, A. (1990). Input evidence regarding the semantic bootstrapping hypothesis. *Journal of Child Language*, *17*, 711-717.

Rondal, J.A., & Docquier, L. (2006). Maternal speech to Down syndrome children. An update. *Journal of Speech-language Pathology and Applied Behavior Analysis*, *1*, 218-227.

Rondal, J.A., & Edwards, S. (1997). *Language in mental retardation*. London: Whurr.

Rondal, J.A., Ghiotto, M., Bredart, S., & Bachelet, J.F. (1987). Age-relation, reliability and grammatical validity of measures of utterance length. *Journal of Child Language*, 14, 433-446.

Rondal, J.A., & Guazzo, G.M. (2012). Morphosyntactic difficulties and rehabilitation in persons with Down syndrome. In R. Hodapp (Ed.), *International review of research in developmental disabilities* (Vol. 42, pp. 85-107). New York: Elsevier.

Rondal, J.A., & Perera, J. (Eds.).(2006). *Down syndrome. Neurobehavioural specificity*. Chichester, UK: Wiley.

Rondal, J.A., Ylieff, M., Elbouz, M., & Docquier, L. (2003). Françoise : A fifteen-year follow up. *Down Syndrome Research and Practice*, *8*, 89-99.

Rosch, E. (1974). Linguistic relativity. In A. Silverstein (Ed.), *Human communication: Theoretical perspectives* (pp. 95-121). New York: Halsted Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rosenbaum, D., Cohen, R., Jax, S., Weiss, D., & Van der Well, R. (2007). The problem of serial order in behavior: Lashley’s legacy. *Human Movement Science*, *26*, 525-554.

Rosenberg, S. (1993). Chomsky’s theory of language. Some recent observations. *Psychological Science*, *4*, 15-19.

Rowland, C., & Pine, J. (2000). Subject-auxiliary inversion errors and wh-question acquisition: What children do know? *Journal of Child Language*, *27*, 157-181.

Rowland, C., Pine,J., Lieven, E., & Theakston, A. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech; *Journal of Child Language*, *30*, 609-635.

Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs. In McClelland, J., & Rumelhart, D. (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol.2, pp. 170-215). Cambridge, MA: MIT Press.

Russel, J. (1987). Three kinds of questions about modularity. In S. Modgil & C. Modgil (Eds.), *Chomsky, N., consensus and controversy* (pp. 223-232). New York: Falmer.

Ruwet, N. (1967). *Introduction à la grammaire générative*. Paris : Plon.

Ruwet, N. (1972). *Théorie syntaxique et syntaxe du français*. Paris: Seuil.

Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*, 437-442.

Sachs, J., & Devin, J. (1976). Young children’s use of appropriate speech styles in social interaction and role-playing. *Journal of Child language*, *5*, 17-24.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Sahin, N., Pinker, S., Cash, S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca’s area. *Science*, *326*, 445-449.

Sandler, W., Meir, I, Padden, C., & Aronoff, M. (2005). Spontaneous sign language in Al-Sayyid bedouins. *Proceedings of the National Academy of Sciences of the United States of America* [published online]. doi: 10.1073/pnas. 0405448102

Sapir, E. (1949). *Selected writings of Edward Sapir* (edited by D. Mandelbaum). Berkeley, CA: University of California Press.

Saunders, B. (2000). Revisiting basic color terms. *Journal of the Royal Anthropological Institute*, *6*, 81-99.

Saussure, F. de (1916, 1955). *Cours de linguistique générale* (posthumous publication by C. Bally, A. Sechehaye, & A. Riedlinger). Paris : Payot.

Savin, H., & Perchonock, E. (1965). Grammatical structure and the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, *4*, 348-353.

Saxton, M. (2010). *Child language: Acquisition and development*. Thousand Oaks, CA: Sage Publication.

Schachter, P. (1976). The subject in Philipppine languages: Actor, topic, actor-topic, or none of the above? In C. Li (Ed.), *Subject and topic* (pp. 491-518). New York: Academic Press.

Schellenberg, E., Reber, A., Di Girolamo, K., & Wang P. (2003). Implicit learning in children and adults with Williams syndrome. *Developmental Neuropsychology*, *23*, 201-225.

Schlesinger, I. (1971). Production of utterances and language acquisition. In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 63-102). New York: Academic Press.

Schlesinger, I. (1988). The origin of relational categories. In Y. Levy, I. Schlesinger, & M. Braine (Eds.), *Categories and processes in language acquisition* (pp. 121-178). Hillsdale, NJ: Erlbaum.

Schlesinger, I. (1995). *Cognitive space and linguistic case*. New York: Cambridge University Press.

Schusterman, R., & Gisiner, R. (1988). Artificial language comprehension in dolphins and sea lions. *The Psychological Record*, *38*, 311-348.

Seagoe, M. (1965). Verbal development in mongolism. *Exceptional Children*, *6*, 229-275.

Searle, J. (1985). *Du cerveau au savoir*. Paris: Hermann.

Segui, J., & Kail, M. (1971). Rôle des caractéristiques syntaxiques et du contexte sur la production verbale. *L’Année Psychologique*, *71*, 429-438.

Seidenberg, M., MacDonald, M., & Saffran, J. (2002). Does grammar start when statistics stop? *Science*, *298*, 553-554.

Senghas, A. (2005). Language emergence: Cues from a new Bedouin sign language. *Current Biology*, *15*, 463-465.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*, 1179-1182.

Seyfarth, R., & Cheney, D. (1993). Meaning, reference, and intentionality in the natural vocalizations of monkeys. In H. Roitblat, L. Herman, & P. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 195-219). Hillsdale, NJ: Erlbaum.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379-423.

Shapiro, K., & Caramazza, A. (2003a). Grammatical processing of nouns and verbs in left frontal cortex. *Neuropsychologia*, *41*, 1189-1198.

Shapiro, K., & Caramazza, A. (2003b). The representation of grammatical categories in the brain. *Trends in Cognitive Science*, *7*, 201-206.

Shapiro, K., Pascual-Leone, A., Mottaghy, F., Gangitano, M., & Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *Journal of Cognitive Neuroscience*, *13*, 713-720.

Shapiro, K., Shelton, J., & Caramazza, A. (2000). Grammatical class in lexical production and morphological processing: Evidence from a case of fluent aphasia. *Cognitive Neuropsychology*, *17*, 665-682.

Shatz, M., & Gelman, R. (1973). The development of communication skills: Modification in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, Serial No 38.

Sin, C., Li, H., & Crawford, D. (2015). Transcriptional regulation by FOXP1, FOXP2, and FOXP4 dimerization. *Journal of Neuroscience*, *55*, 437-448.

Sinclair, H. (1971). Sensory-motor action patterns as a condition for the acquisition of syntax. In R. Huxley & D. Ingram (Eds.), *Language acquisition: Models and methods* (pp. 121-135). New York: Academic Press.

Sinclair, H. (1973). Language acquisition and cognitive development; In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 9-25). New York: Academic Press.

Skinner, B. (1938). The behavior of organisms. An experimental analysis. Oxford, UK: Appleton-Century.

Skinner, B. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Slobin, D. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, *5*, 219-227.

Slobin D. (1971). On the learning of morphological rules: A reply to Palermo and Eberhart. In D. Slobin (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 215-223). New York: Academic Press.

Slobin D. (1973). Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 607-619). New York: Holt, Rinehart, & Winston.

Slobin, D. (1975). “The more it changes…” On understanding language by watching it move through time. *Papers and Reports on Child Language Development* (Stanford University), *10*, 1-30.

Slobin, D. (1978). A case study of early language awareness; In A. Sinclair, R. Jarvella, & W. Levelt (Eds.), *The child’s conception of language* (pp. 45-54). Berlin: Springer.

Slobin, D. (1981). L’apprentissage de la langue maternelle. *La Recherche*, *12*, 572-578.

Slobodchikoff, C., Kiriazis, J., Fischer, C., & Creef, E. (1991). Semantic information distinguishing individual predators in the alarm calls of Gunnison’s prairie dogs. *Animal Behaviour*, *42*, 713-719.

Smith, N., & Tsimply, I. (1995). *The mind of a savant. Language learning and modularity*. Oxford, UK : Blackwell.

Smolensky, P. & Legendre, G. (Eds.). (2006). *The harmonic mind* (Vol.1). Cambridge, MA: MIT Press.

Sokolov, A. (1971). *Inner speech and thought*. New York: Plenum Press.

Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology*, *22*, R912-3. Doi:10.1016/j.cub2011.09.049

Sproat, R. (2002). *The linguistic significance of finite state techniques*. Boston, MA: American Association for the Advancement of Science.

Squire, L., & Zola, S. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of National Academy of Sciences of the Unites States of America*, *93*, 13515-13522.

Squire, L., & Zola, S. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, *8*, 205-211.

Staats, A. (1971). Linguistic-mentalistic theory versus explanatory S-R learning theory of language development. In D. Slobin, (Ed.), *The ontogenesis of grammar. A theoretical symposium* (pp. 103-150). New York: Academic Press.

Staats, A. (1974). Behaviorism and cognitive theory in the study of language. A neo-psycholinguistics. In R. Schiefelbusch & L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation, and intervention* (pp. 615-656). Baltimore, MD: University Park Press.

Stadler, M. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 674-685.

Stamatakis, E., Randall, B, & Komisarjevsky Tyler, L. (2007). Grammatical categories in the brain: The role of morphological structure. *Cerebral Cortex*, *17*, 1812-1820.

Steedman, M. (1999). Connectionist sentence processing in perspective. *Cognitive Science*, *23*, 615-634.

Stemmer, N. (1973). *An empiricist theory of language acquisition*. The Hague, The Netherlands: Mouton.

Sternberg, D. (1985). Controlled versus automatic processing. *Behavioral and Brain Sciences*, *8*, 32-33.

Stromswold, K. (2001). The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language, 77*, 647-723.

Taatgen, N., & Anderson, J. (2002). Why do children learn to say « Broke »? A model of learning the past tense without feedback. *Cognition*, *86*, 123-155.

Tattersall, I. (1998). *Becoming human: Evolution and human uniqueness*. New York: Harcourt & Brace.

Teramitsu, I., Kudo, L., London, S., Geschwind, D., & White, S. (2004). Parallel FOXP1 and FOXP2 expression in songbirds and human brain predicts functional interaction. *Journal of Neuroscience*, *24*, 3152-3163.

Tesnière, L. (1966). *Éléments de syntaxe structurale*. Paris: Klincksieck.

Thibaut, J.P., Rondal, J.A., & Kaens, A.M. (1995). Actionality and mental imagery in children’s comprehension of declaratives. *Journal of Child Language*, *22*, 189-209.

Thewis, B. (1991). *Capacité métalinguistique à l’école primaire: évaluation expérimentale des capacités d’analyse syntaxique en 4e et 6e année* (Unpublished maser’s thesis). University of Liège, Belgium.

Thomas, K., Hunt, R., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience, 16*, 1339-1351.

Thomas, M., Grant, J., Barham, Z., Gsödl, M., Laing, E., Lakusta, L., Tyler, L., Grice, S., Patterson, S., & Karmiloff-Smith, A. (2001). Past-tense formation in Williams syndrome. *Language and Cognitive Processes*, *16*, 143-176.

Thorndike, E. (1932). *The fundamentals of learning*. New York: Teachers College Press.

Tinbergen, N. (1951). *The study of instinct*. Oxford, UK: Clarendon Press.

Toley, W., & Van Valin, R. (1984*). Functional syntax and universal grammar.* Cambridge, UK: Cambridge University Press.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, *1,* 131-156.

Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: MIT Press.

Trabant, J. (1995). *Humboldt ou le sens du langage*. Sprimont, Belgium : Mardaga.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tomblin, J., Records, N., Buckwater, P., Zhang, X., Smith, E., & O’Brien, M. (1997). Prevalence of specific language impairment in kindergarten. *Journal of Speech, Language and Hearing Research*, *40*, 1245-1260.

Tranel, D., Adolphs, R., Damasio, H., & damasio, A. (2001). A neural basis for the retrieval of words for action. *Cognitive Neuropsychology*, *18*, 655-670.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Oxford University Press.

Tunmer, W., & Grieve, R. (1984). Syntactic awareness in children. In W. Tunmer, C. Platt, & M. Herriman (Eds.), *Metalinguistic awareness in children* (pp. 144-168). Berlin: Springer Verlag.

Tyler, L., Randall, B., & Stamatakis, E. (2008). Cortical differentiation of nouns and verbs depends on grammatical markers. *Journal of Cognitive Neuroscience*. Advance online publication. doi: 10.1162/jocn.2008.20095

Tyler, L., Russel, R., Fadili, J., & Moss, H. (2001). The neural representation of nouns and verbs: PET studies. *Brain*, *124*, 1619-1634.

Udden, J., Ingvar, M., Hagoort, P., & Petersson, K. (2017). Broca’s region: A causal role in implicit processing of grammar with crossed non-adjacent dependencies. *Cognition*, *164*, 188-198.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270.

Ullman, M., Corkin, S., Coppola, M., Hoickok, G., Growdon, J., Koroshetz, W., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative theory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, *9*, 266-276.

Ullman, M., & Pierpont, E. (2005). Specific language impairment is not specific to language. The procedural deficit hypothesis. *Cortex*, *41*, 399-433.

Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition*, *35*, 105-122.

Valian, V., & Stojak Caplan, J. (1978). What children say when asked “what?” A study of the use of syntactic knowledge. *Journal of Experimental Child Psychology*, *28*, 424-444.

Vallar, G., & Papagno, C. (1993). Preserved vocabulary acquisition in Down’s syndrome: The role of phonological short-term memory. *Cortex*, *29*, 467-483.

Van der Hulst, H. (Ed.).(2009). *Recursion and human language*. Berlin: de Gruyter.

Van der Lely, H., & Pinker, S. (2014). *The biological basis of language: Insight from developmental grammatical impairments* [published online]. Doi: dx.doi.org/101016/j.tics.2014.07.001

Van Valin, R. (1999). Generalized semantic roles and the syntax-semantics interface. In F. Corblin, C. Dobrovic-Sorin, & J. Marandin (Eds.), *Empirical issues in formal syntax and semantics* (Vol. 2, pp. 373-389). The Hague, The Netherlands: Thesus.

Van Valin, R. (2003). *Exploring the syntax-semantics interface*. Cambridge, UK: Cambridge University Press.

Vargha-Khadem, F., Gadian, D., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Review Neuroscience*, *6*, 131-138.

Velichkowsky, B. (2017). Consciousness and working memory: Current trends and research perspectives. *Consciousness and Cognition*, *55*, 35-45.

Vernes, S., Newbury, D., Abrahams, B., Winchester, L., Nicod, J., Groszer, M., Alarcón, M., Oliver, P., Davies, K., Geschwind, D., Monaco, A., & Fisher, S*.*  (2008). A functional genetic link between distinct developmental language disorders*. New England Journal of Medicine, 359,* 2337-2345.

Vias, C., & Dick, A. (2017). Cerebellar contributions to language in typical and atypical development: A review. *Developmental Neuropsychology*, *42*, 404-421.

Vicari, S. (2001). Implicit versus explicit memory function in children with Down and Williams syndrome. *Down Syndrome Research and Practice*, *7*(1), 35-40.

Viglioco, C., Vinson, D., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition*, *94*, B-91-B-100.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Vinter, A., & Detable, C. (2003). Implicit learning in children and adolescents with mental retardation. *American Journal of Mental Retardation*, *108*, 94-107.

Vokey, J., & Brooks, L. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 328-344.

Wada, J., Clarke, R., & Hamm, A. (1975). Cerebral hemispheric asymmetry in humans. *Archives of Neurology*, *32*, 329-346.

Wang, Y., Chang, C., Chen, S., Wang, C., & Wuang, Y. (2015). Profiles of visual perceptual functions in Down syndrome. *Research in Developmental Disabilities*, *37*, 112-118.

Warburton, E., Wise, R., Price, C., Weiler, C., Hadar, U., Ramsay, S., & Frackowiak, R. (1996). Noun and verb retrieval by normal subjects with PET. *Brain*, *119*, 159-179.

Waterfall, H., Sandbank, B., Onnis, L., & Edelman, S. (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language, 37*, 671-703.

Watkins, K, Vargha-Khadem, F., Ashburner, J., Passingham, R., Connelly, A., & Friston, K. (2002). MRI analysis of an inherited speech and language disorder: Structural brain anomalies. *Brain*, *125*, 465-478.

Watt, W. (1970). On two hypotheses concerning psycholinguistics. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 137-220). New York: Wiley.

Waxman, S., & Booth, A. (2001). Seeing pink elephants: Fourteen-month-olds’ interpretations of novel nouns and adjectives. *Cognitive Psychology*, *43*, 217-242.

Weckerly, J., Wulfeck, B., & Reilly, J. (2001). Verbal fluency deficits in children with specific language impairment: Slow rapid naming or slow to name. *Child Neuropsychology*, *73*, 142-152.

Weismer, S., & Hesketh, L. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech and Hearing Research*, *39*, 177-190.

Weissenborn, J. (1992). Null subject in early grammars: Implications for parameter-setting theories. In J. Weissenborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical issues in language acquisition. Continuity and change in development* (pp. 269-299). Hillsdale, NJ: Erlbaum.

Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition.* Cambridge, MA: MIT Press.

Wexler, K., Culicover, P., & Hamburger, H. (1975). Learning-theoretic foundations of linguistic universals. *Theoretical Linguistics*, *2*, 213-253.

Wierzbicka, A. (2006). Sens et grammaire universelle: théorie et constats empiriques. *Linx*, *54*, 181-207.

Wilkins, W., & Wakefield, J. (1995). Brain evolution and neurolinguistics preconditions. *Behavioral and Brain Sciences, 18*, 161-182.

Williams, J., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua*, *118*, 522-553.

Williams, J., & Rebuschat, P. (2012). Statistical learning and syntax: What can be learned, and what difference does meaning make? In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp; 237-264). Berlin: de Gruyter Mouton.

Willis, C. (1991). *Exons, introns, and talking genes: The science behind the human genome project*. New York: Basic Books.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflections in English. *Journal of Child Language*, *30*, 75-115.

Wisniewsky, K., Kida, E., & Brown, T. (1996). Consequences of genetic abnormalities in Down’s syndrome on brain structure and function. In J.A. Rondal, J. Perera, L. Nadel, & A. Comblain (Eds.), *Down’s syndrome. Psychological, psychobiological and socio-educational perspectives* (pp. 3-19). London : Whurr.

Wisniewsky, K., Kida, E., Golabek, A., Walus, M., Rabe, A., Palminiello, S., & Albertini, G. (2006). Down syndrome: From pathology to pathogenesis. In J.A. Rondal & J. Perera (Eds.), *Down syndrome. neurobehavioural specificity* (pp. 17-33. London: Whurr.

Witt, A. (2011) *L’apprentissage implicite d’une grammaire artificielle chez l’enfant avec et sans retard mental: rôle des propriétés du matériel et influence des instructions* (Unpublished doctoral dissertation). Université de Bourgogne, Dijon, France. NNT HAL Open archives: 20010DIJOL019; 2011

Witt, A. & Vinter, A. (2012). Artificial grammar learning in children: Abstraction of rules or sensitivity to perceptual categories. *Psychological Research, 76,* 97-110.

Witt, A., & Vinter, A. (2013a). How explicit and implicit test instructions in an implicit learning task affect performance. PloS One, 8 (1): e53296 doi:10. 1371/journal.pone. 0053296

Witt, A., & Vinter, A. (2013b). Children with intellectual disabilities may be impaired in encoding and recollecting incidental information. *Research in Developmental Disabilities*, *34*,864-871.

Wittgenstein, L. (1961). *Investigations philosophiques*. Paris : Gallimard.

Worf, B. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Worf* (edited by T. Carroll). Cambridge, MA: MIT Press.

Wyatt, B., & Conners, F. (1998). Implicit and explicit memory in individuals with mental retardation. *American Journal of Mental Retardation*, *102*, 511-526.

Yamada, J. (1990). *Laura: A case for the modularity of language*. Cambridge, MA: MIT Press.

Xiaoxue He, A., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, *13*, 335-356.

Yang, C. (2016). *The price of linguistic productivity. How children break the rules of language*. Cambridge, MA: MIT Press.

Yngve, V. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104*, 444-446.

Youssef, V. (1988). The language bioprogram hypothesis revisited. Journal of *Child Language*, *15*, 451-458.

Zipf, G. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: AddisonWesley.

1. The terminology and symbolics of generative grammar have been modified several times with the evolution of the theory. I have kept with the traditional ones, which are more immediately readable than later ones, as for example, those in *Government and Binding* (Chomsky, 1981). [↑](#footnote-ref-1)
2. One must add to the binomial signified-signifier a third term, reference, which was not fully appreciated at the time of Saussure. This is the relationship that a word has with a class of concrete or fictitious objects, events, or abstractions (except for personal names where reference is idiosyncratic). The signified instead is considered to be a purely mental entity. [↑](#footnote-ref-2)
3. Linear refers to something having the appearance of a line, i.e., a series of juxtaposed points (from the Latin *linea*: *thread*, *string*); hence a combination of immediate dependencies. [↑](#footnote-ref-3)
4. I will introduce only those strings of words that may be judged to form a grammatical sentence with a capital letter. [↑](#footnote-ref-4)
5. The idea to represent sentences in the form of a reverse tree comes from Hockett (1958). [↑](#footnote-ref-5)
6. Chomsky (1966) flatly rejects as empty alternative suggestions (for example, Hockett, 1958) explaining the creative aspect of human language in terms of analogies or grammatical patterns. Yet, it can be argued that paradigmatic variation may account for the creativity aspect of human language more convincingly than a metaphorical mechanism simply ascribed to a putative faculty of language. [↑](#footnote-ref-6)
7. It goes back to the work of Roger Bacon in the 13th century and his idea that all languages are built upon a common grammar, though it may undergo incidental variations (cf. Wierzbicka, 2006). [↑](#footnote-ref-7)
8. This repeated Chomskyan pretension is puzzling in a person of his extended cultural and scientific knowledge. [↑](#footnote-ref-8)
9. A biological perspective on language was proposed by Lenneberg (1964, concluding his paper in the following way: “…the basis for language capacity might well be transmitted genetically…” (p.45). As will be discussed in Chapter 3, there is no question that the bases for language are innately given. This is what could be called “neurofunctional innateness” as opposed to “representational innateness.” There is a big difference between the two theories (and that is why Lenneberg’s statement is ambiguous). Representational innateness is the idea that genuine linguistic representations are provided in the human genome and actualized in the brain. Neurofunctional innateness restricts the genetic blueprint to the neural foundation of the procedures and mechanisms involved in language acquisition and functioning. [↑](#footnote-ref-9)
10. In idiomatic expressions, there is often the additional need to decipher the wording in going beyond its literal meaning even when the secondary sense (for example, in metaphors) is conventional. [↑](#footnote-ref-10)
11. The Spanish and Russian examples are particularly telling. They show that the principle of second heavy overcomes a tendency in language pragmatics (in relation to spatial deixis here), that is, the expressive reference is anchored to the speaker. [↑](#footnote-ref-11)
12. Having at one’s disposal a series of items numbered 1…N and a data base with the sequences of these items, one computes for each pair i and j the conditional probability that j immediately follows i in the series. An equation gives the measure of the conditional entropy (defined as a measure of the amount of information needed to describe the outcome of a random variable Y given that the value of another random variable X is known, expressed in logarithmic units, *nats*). The highest value is reached when all the items are sequentially equiprobable and the lowest when the same item is repeated throughout the sample of speech. [↑](#footnote-ref-12)
13. The indication in the last sentence is ambiguous. All semantic relations can be coded morphosyntactically and all morphosyntactic structures correspond to meaning relations. There is not a precise one-to-one correspondence between the two sets in the sense that a given semantic relation may be expressed in sentence surface in a variety of ways. This is probably what Chomsky means by “imperfect.” In other passages, he qualifies the correspondences as “inexact” (1957, p. 101), which is certainly incorrect. The so-called imperfection of the correspondence is a putative consequence of the fact that morphosyntax and semantics are separate dimensions of the language system, each one with its peculiarities, as Chomsky rightly insists. [↑](#footnote-ref-13)
14. The Chomskyan notion of rewriting rule is borrowed from the work of Post in mathematical logic (Davis, 1994). [↑](#footnote-ref-14)
15. The asterisk indicates that the sentence is not grammatical. [↑](#footnote-ref-15)
16. For a psycholinguistic analysis of these works, see my texts (Rondal, 2000, 2016). [↑](#footnote-ref-16)
17. Klein (2017) suggests circa 50,000 ago. He proposes that Neanderthals lacked some genes underpinning language and other cognitive functions (which does not mean that they were necessarily connected to linguistic representations). [↑](#footnote-ref-17)
18. Advanced aspects of morphosyntax are typically not mastered before later in development (i.e., sometimes 9 or 10 years of age). They concern delicate syntactic matters as, for example, in English, the contrast between sentences like *John is eager to see* and *John is easy to see*, in which the adjective is predicated of *John* who is also the subject of the infinitival complement in the first sentence; whereas *John* is the subject in both cases in the second sentence (see C. Chomsky, 1969, and Slobin, 1966, for classical empirical studies on the question). [↑](#footnote-ref-18)
19. Another formal theory involving a notion of well-formedness is the so-called *Optimality theory* developed in more recent years (Prince & Smolensky, 2004; see also Smolensky & Legendre, 2006). It is based on the idea that a language, and in fact any grammar, is based on a system of conflicting forces that can be hierarchized in order of importance. Every language implies some sort of violation of well-formedness. Only the most important ones hierarchy-wise really matter. The surface form of language reflects the resolution of conflicts between competing constraints in the system. Two basic types of constraints (considered to be universal) are distinguished: (1) faithfulness constraints: the surface form has to match the underlying or lexical form in some particular way; and (2) markedness constraints imposing requirements on the structural well-formedness of the surface form. This theorizing has been criticized as being unfalsifiable, hence not allowing empirical testing. Chomsky (1995, p.380), commenting on a previous and unpublished manuscript of Prince and Smolensky on the topic, is not convinced that *Optimality theory* differs from other theories with intermediate levels between sentence depth and surface. [↑](#footnote-ref-19)
20. In a French-German movie realized by Gérard Oury and situated in Nazi Germany of the years 1930, the French actor Jean-Paul Belmondo is shouted out to by a train controller who says to him: *Français vous êtes*? (literal gloss: *French you are?*). Belmondo replies: *En français, on ne dit pas con vous êtes, même si c’est vrai. On dit vous êtes con* (gloss : *In French one does not say stupid you are, even if that is true. One says you are stupid*). In a nutshell, this exchange illustrates the core of the problem addressed in this book and its proposed solution in terms of two morphosyntactic systems (see the following chapters). The language user explicitly knows how to order words correctly (for example, in French). He also knows that meaning is something else (it may be true even if not grammatically expressed). [↑](#footnote-ref-20)
21. Wittgenstein, more generally, is skeptical regarding the possibility of establishing an empirical science of language. His concept of “language game” seems to be alien to the notion of structural integration in a language. But that will not be discussed here. [↑](#footnote-ref-21)
22. Roger Brown, from Harvard University, recorded and analyzed hours of verbal interactions during several months between three children, conveniently named, Adam, Eve, and Sarah, with their respective parents or caregivers. Eve was 18 month-old at the beginning of the study, Adam and Sarah had 27 months. Brown made his transcripts available for other researchers’ analyses. The children came to be referred to as the “Harvard children.” Brown’s (1973) book devoted to the first stages of language acquisition is partially based on a thorough and insightful analysis of these transcripts. [↑](#footnote-ref-22)
23. For example, a sentence with a high transitivity value typically involves a powerful agent, a punctual action, i.e., an action with a clear point of departure and ending, limitation in time, and an individualized patient clearly affected by the action of the agent (for instance, a sentence such as *John cuts a piece of wood* is higher in semantic transitivity than *John sees a girl*). [↑](#footnote-ref-23)
24. In a memoir note, Chafe (2002) recalls some of the reasons for his theoretical endeavor in favor of semantics. He moved from the narrow semantic conception of Bloomfieldian linguistics to developing a fine-grained theory of meaning as the fundamental structure of language. Chafe explains that he never was attracted to generative grammar that he saw as excessively concerned with formal syntax, and in a sense “superficial.” He believes that the base structure of generative grammars is not able to fully account for combinatorial meanings. Chafe (2002) adds that in his view the generative semanticists were right in associating deep structure with semantics but wrong in the way they conceived of semantics itself (for example, the recourse to logical notions such as that of predicate which is insufficiently diversified to account for the variety of relations existing between verbs and nouns). [↑](#footnote-ref-24)
25. The centrality of the verb in morphosyntactic account is a long tradition in structural linguistics; it is already clearly stated in Tesnière (1966). [↑](#footnote-ref-25)
26. This is just a little piece of the Italian system. The pronominal forms vary according to the tonic/atonic contrast and the grammatical (and semantic) function of the pronoun. [↑](#footnote-ref-26)
27. Actually, Sapir’s suggestions were milder than those of Worf who seems to have radicalized the relativity hypothesis to a significant extent (Champagnol, 1993). [↑](#footnote-ref-27)
28. Topicalized sentences have the topic element (new information) appearing at the front as opposed to the canonical position further to the right of the sentence. [↑](#footnote-ref-28)
29. Expressions such as motherese or baby talk should be avoided when referring to the language input to children. As indicated, the particular adaptations are neither specific to mothers nor restricted to babies. [↑](#footnote-ref-29)
30. It is always possible to consider that the implicit apprehension of some formal regularity on a distributional or frequency basis, to the extent it can be generalized, corresponds to a “probabilistic rule.” As Albright and Hayes (2003) also remark, if one accepts the notion of probabilistic rule, the distinction between rue and analogy loses most or all of its substance. [↑](#footnote-ref-30)
31. Natural or Neperain logarithms are based on the irrational number *e* whose value is approximately 2.72. [↑](#footnote-ref-31)
32. In some way. Yang (2016, p.17), for example, claims that children never hear pairs like *walk*-*walked* and *sing*-*sang* in conjunction. May be, but they certainly hear *walk* and *sing* at a given moment and *walked* and *sang* (slightly) later or before in meaningful contexts. Are we to suppose that children depicted by Yang as extraordinarily good at detecting rules and exceptions are not able to relate corresponding verbal forms in nearly related interactive sessions? [↑](#footnote-ref-32)
33. Yang (2016) suggests that his tolerance principle proportionally reduces the weight of universal grammar in accounting for language acquisition to the extent that it is a learning principle. [↑](#footnote-ref-33)
34. Priming refers to an implicit memory effect in which prior exposure to a given stimulus influences the response to a further stimulus. [↑](#footnote-ref-34)
35. The latest developments of the theory are presented in a 2012 opus. They do not relate to the discussion here. [↑](#footnote-ref-35)
36. In FXS, one gene (most often FMR-1, sometimes FMR-2) is mutated, which reduces its transcription and results in a drastic reduction of the production of the developmentally needed proteins fmr-1 or fmr-2 in the brain. [↑](#footnote-ref-36)