Analyticity and Syntheticity

# Empirical Approaches to Language Typology

6

*Editors*

Georg Bossong
Bernard Comrie

# Analyticity and Syntheticity

A Diachronic Perspective with
Special Reference to Romance Languages

*by*
Armin Schwegler

♾ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure
permanence and durability.

The aim of science is to seek the simplest explanation of complex facts. We are apt to fall into the error of thinking that the facts are simple because simplicity is the goal of our quest.

Alfred North Whitehead:
*Concept of Nature*

# Contents

# Preface

Since the early 19th century, linguists have been using the terms "analytic" and "synthetic" to characterize the morphological architecture of languages. The concepts of analyticity and syntheticity are frequently used today, but linguists differ widely as to what they regard as a synthetic or an analytic language. This book explores how it is possible that languages such as French are located by different investigators at opposite extremes on the analytic/synthetic scale. Is this disagreement simply due to the lack of clearly defined typological parameters? Or have languages like French recently moved from one end of the scale to the other, so that the different claims merely refer to different diachronic stages?

A principal aim of this investigation is to demonstrate the validity of the *concepts* of analyticity and syntheticity, notwithstanding the vagueness and ambiguity with which the terms have been used. A clearer understanding and more precise delineation of the terms analytic and synthetic—hence of the concepts themselves—is essential not only to avoid the kinds of confusion and oversimplification which mar many typological studies, but also to grasp better certain evolutionary trends in the histories of Romance as well as of other languages.

A number of well-documented developments from within the histories of Latin and Romance (with occasional glances at Germanic and other Indo-European languages) will be reviewed here to reveal how once essentially unrelated speech segments undergo synthesis to become semantic and/or grammatical units of an entirely different sort. It is hoped that these findings, if corroborated by research from other language families, may serve as a basis for formulating a set of cross-linguistic diachronic principles with respect to the ways in which languages undergo certain morphological changes.

Ever since Sapir's well-known excursion into the question of "drift," there has been considerable interest in and debate about the cyclical movement of languages between analysis and synthesis. This investigation purports to shed new light on this phenomenon, and to show that, contrary to a widespread belief, there is nothing hermetic or mysterious about rythmic cycles of analytic and synthetic constructs.

This book is a thoroughly revised and considerably enlarged version of my 1986 dissertation, completed under the auspices of the Group in Romance Philology at the University of California, Berkeley. I would like to express my deepest appreciation and gratitude to those who have been a source of inspiration and encouragement to me during the past seven years: Yakov Malkiel, Jerry Craddock, Joseph Duggan, Charles Faulhaber, Ruggero Stefanini, and Karl Zimmer. A particular debt of gratitude also goes to Georg Bossong whose insightful comments on an earlier

version of the manuscript have significantly improved both the form and the content of this book. Very special thanks are due to the long list of friends and colleagues whose support, encouragement or cheerful disposition have helped me enjoy the many years that this book was under preparation. Among these let me single out Ella Mae and William Burke, Thomas Büchler, Barbara DeMarco, Virginia Gonzales, Kathryn Klingebiel, John Levy, Albert Muth, Ruedi Müller, Celina Navarro, Salvador Rodezno, Jeff Shimanoff, Victor Simarra, Christopher Stookey, Morteza Tabatabaipour, Norma Thompson, Jeff Turley, and Thomas Walsh.

No one has contributed more to the ideas developed in this work, been more supportive, encouraging, and inspiring than Suzanne Fleischman. Her comments and criticisms—with regard to both content and style—prompted me to recognize certain inconsistencies in earlier versions, and helped sharpen my thinking so that I could ultimately find answers even to the knottiest issues of "analyticity and syntheticity." I dedicate this book to her.

Last, but by no means least, a special note of thanks is due to my parents and my brother Roland who have remained close to me despite living in my far-off native Switzerland.

# Introduction

*Background of the question*

The origins of non-genetic criteria for language classification can be traced to the moment when August Wilhelm von Schlegel, whose brother Friedrich had already proposed a twofold classification of languages into those with affixes and those with inflections, suggested in 1818 that inflectional languages be further subdivided into "synthetic" and "analytic" types. Though the new approach generally met with little enthusiasm, at least during the 19th century, and was at the time vehemently rejected by some of the best-known linguists (e.g., Humboldt), August Wilhelm Schlegel's proposal nonetheless turned out to set the tone for many subsequent approaches to language classification.

The acceptance of the terms analytic and synthetic is due in large part to the publication of Edward Sapir's much celebrated *Language* (1921), in which A. Schlegel's terminology is discussed in considerable detail in the chapter "Types of Linguistic Structure." Recognizing the potential usefulness of the concept, Sapir proposed numerous refinements to A. Schlegel's simple binary classification by incorporating analyticity and syntheticity into a more sophisticated matrix of morphological features. The impact of Sapir's book on subsequent linguistic typology was so profound and long-lasting that the terms analytic and synthetic (as well as Sapir's newly coined "polysynthetic") have since become part of standard linguistic terminology.

The ready acceptance of the terms, and hence the concepts themselves, is somewhat surprising in light of the fact that Sapir himself recognized that his scheme, though a vast improvement over Schlegel's, was still open to criticism. Like those before him, Sapir could do no more than provide a scalar concept of syntheticity which relied heavily upon vague and arbitrary criteria. Qualified labels such as "mildly synthetic," "notably synthetic," etc. could not satisfy those who demanded rigorous, unambiguous, and non-subjective parameters; thus the linguistic community, while accepting the terms, remained undecided about how the concepts should be used.

While most linguists of the period between 1920 and 1950 recognized the general problem of dealing with a useful taxonomic paradigm which, however, lacked the exactitude and rigor expected of modern linguistic analysis, few were willing to tackle the problem head on in order to eliminate the terminological—and conceptual—"softness" inherent in the analytic/synthetic distinction. It was not until 1954 that a radically different approach was undertaken by Joseph Greenberg. Using the analytical tools of contemporary American structuralism, he revised the Sapirian classification, and introduced the MORPHEME as one of the distinctive units of

morphological typology, deriving the degree of syntheticity by a simple mathematical formula: the total number of MORPHEMES divided by the total number of WORDS (thus *lov-ed* [2 morphemes/1 word] had a "synthetic index" of 2.00). By plotting the frequency distribution of morphemes per word in several languages, linguists could now establish what many hailed as truly objective indices of syntheticity. Not surprisingly, there was a quick and positive response to the new method; for while Sapir's classification relied on a fairly loose determination of the frequency and degree of formal morphemic cohesion, Greenberg's classification had the advantage not only of showing differences in syntheticity—now quantified—between languages, but also of providing clear guidelines for obtaining undistorted results in calculations involving large numbers of languages.

Though the structuralist vogue of the thirties and forties might well have led Greenberg to propose a new and promising way of calculating syntheticity, the structuralist enterprise became, ironically, the major obstacle to the application and acceptance of his technique. As the structures of more and more languages were studied, it became apparent that there is no single, entirely satisfactory working definition of the "word," a state of affairs which, in turn, calls into question the usefulness of Greenberg's formula MORPHEME/WORD.

Perhaps in part because Greenberg's approach has a strongly scientific flavor, numerous typologists (e.g., V. Krupa, G. Altmann, E. Slavičkova, W. Lehfeldt, H. Haarmann, etc.) have continued to use his formula, and have in the past two decades reformulated and extended the classic quantitative approach. Alternative ways for calculating the index of synthesis are being suggested, and some important methodological differences have been proposed. Several typologists have argued convincingly that in order to be meaningful the parameters analytic and synthetic must be applied only to *individual* cross-linguistic categories and not to languages as wholes. Yet notwithstanding the originality and soundness of these proposals, the central problem of finding a satisfactory working definition of the "word" has been glossed over; and the lack of a universally applicable definition of the word leads directly to the collapse of what has long been hailed as a promising avenue of research: morphological typology. Some have recognized by now the methodological fallacy of plugging in formulas (e.g., MORPHEME/WORD) in which one of the parameters lacks universal extension. If the concepts analytic and synthetic are to have a place in future linguistic discussion, their definitions must not be based on a methodology whose shortcomings cannot be corrected in the foreseeable future. Unless we can find a formula which does not take the word as one of its basic parameters, the idea of operating with a *quantitatively* based typology of analytic and synthetic languages ought to be laid to rest.

*The aim of this study*

My curiosity about the terms analytic and synthetic was sparked for a number of reasons. First, I noticed that in much of the literature, even recent literature, over-simplifications and confusion occur frequently in the description of morphological systems, as exemplified by the following contradictory quotes:

> ... an analytic language like Vulgar Latin, English, and French ... (Lathrop 1980:22)

> Ce qui est certain, c'est que, même s'il existe un type analytique de langue, ce type n'est en tout cas pas celui du français, qui est une langue synthétique. (Tesnière 1932:64)

Secondly, I observed that in discussing certain aspects of Romance morphosyntax, I myself frequently used expressions such as "analytic (*go-*) future," "analytic person/number marking" etc. without truly understanding the rationale for these labels. Once aware of the conceptual vagueness of the terms, I discovered that many of the so-called analytic constructs (e.g., the "analytic" person/number marking of Fr *je parle* 'I speak' or the "analytic" Sp *vas a hablar* 'you are going to speak') have a considerably tighter morphological cohesion (i.e., are more synthetic) than the label analytic suggests. I realized from the onset of my investigation that a clearer understanding and more precise delineation of the terms analytic/synthetic—and hence of the concepts themselves—is essential, not only to avoid the kinds of confusion and/or oversimplification which have marred many typological studies, but also to understand better certain evolutionary trends in the histories of French, Spanish, Portuguese, Italian, and Romansh, as well as other languages.

Convinced of the potential usefulness of the analytic/synthetic distinction, I have set the following goals for myself in this study:

(1)   to trace the history of these concepts and offer reasons for their remaining relatively vague and ambiguous;

(2)   to demonstrate why the concepts as currently used cannot fruitfully be applied in cross-linguistic analysis;

(3)   to suggest alternative ways in which the analytic/synthetic distinction may profitably be used;

(4)   to rectify a number of oversimplifications in the analysis of typological change, and offer a more nuanced picture of certain morphosyntactic developments in Romance;

(5)   to answer the question of why languages—Romance languages as well as IE and non-IE languages—appear to drift back and forth between analysis and synthesis in a cyclical manner.

*Methodology and results*

This study started as an attempt to answer the above questions, with no preconceived notions as to how the concepts analytic and synthetic would ultimately have to be revised in order to be both meaningful and useful. Hoping to obtain a more balanced picture of how the terms have been applied, both by linguists directly concerned with their definition as well as those who have used them simply as part of standard linguistic terminology, I began by systematically combing a large number of linguistic dictionaries as well as treatises and articles on Romance morphosyntax. Logically, the next step was to determine why there is so little agreement about the morphological typologies of French, Spanish, Italian, etc. How is it possible that the same language is located by different investigators at opposite extremes on the analytic/synthetic scale? Is the disagreement simply due to the lack of clearly defined typological parameters? Or have these languages recently moved from one end of the scale to the other, so that the different claims simply refer to different diachronic stages?

By re-examining developments in Romance predicate negation and in the verbal core from Latin to French—most notably the switch from post- to pre-verbal person/number marking and the rise of forms like *je chanterai* 'I will sing,' *j'ai chanté* 'I sang/have sung,' *je vais chanter* 'I am going to sing,' I have attempted to answer not just these but a number of other pertinent questions as well. First it will be shown that speech units synthesize according to a series of patterned semantic, morphological and/or phonological changes. In the course of the investigation it will become clear that all changes in the analytic/synthetic spectrum have an underlying semantic basis; that is, only mutually "relevant" items (defined in §3.3.1.1) acquire greater morphosyntactic and/or phonological cohesion. Failure to recognize this has led several investigators to the erroneous belief that the main cause for changes in the analytic/synthetic spectrum is the phonological *rapprochement* between syntactically conjoined speech units.

A major difference between the present approach to the question of syntheticity and analyticity and earlier ones lies in the types of data used. In classifying units such as *je parle* 'I speak,' *voy a hablar*, 'I am going to speak,' etc. scholars have generally failed to take into account the changes, crucial for the question of analyticity and syntheticity, which Romance languages have undergone in *familiar* speech. One of the main concerns of this investigation therefore, will be to show how these often very recent developments have come to change the morphological make-up of Romance vernaculars.

Further important differences between this approach and earlier ones will emerge in the course of the discussion. It will become clear that in order to maintain the meaningfulness of the parameters synthetic and analytic:

(1)   their use must be limited to speech units rather than entire languages,
(2)   the "word" should not be regarded as essential to the concepts, and
(3)   they must be understood not as quantifiable absolutes but as the *rough measure of the overall morphemic interdependency of speech units.*

It will be suggested that this measure is best arrived at by taking into account semantic and morphosyntactic as well as phonological criteria. Analyticity will thus be defined as the semantic, syntactic, morphological and phonological *autonomy* of morphemes within a speech unit, while syntheticity will be characterized as the measure of semantic, syntactic, morphological and phonological *interdependency* (or *relatedness*) of morphemes within a speech unit.

An important point which will emerge from my investigation is that no definition of analytic/synthetic—and, consequently, no method of measuring morphemic interdependency—can be entirely objective and/or language independent. But as I hope to show, the impossibility of arriving at a purely objective definition of these concepts is only a minor drawback, since my approach merely involves an attempt to plot the *general direction* (i.e., ANALYTIC → SYNTHETIC, or SYNTHETIC → ANALYTIC) of a speech unit over its recorded history, rather than to calculate the exact location of such a unit along the analytic/synthetic axis at any given point.

One key finding of this study will be that many long-term diachronic changes cannot be grasped appropriately without the notions of analyticity or syntheticity. It will be shown, for instance, that to understand how Lat *ego* 'I' differed from its late OFr descendent *je* it does not suffice to list the individual phonological and morphosyntactic changes that shaped *ego* into a different unit. OFr *je* has a morphosyntactic and semantic profile which in many ways parallels that of its ancestor *ego*, but it differs fundamentally from *ego* by having entered into a tighter relation with the verb on several levels. Similarly, early OFr *pas* (< CL *passum* 'a step'), when used in conjunction with the negator *non*, was on the way to becoming a unit of an entirely different sort roughly around the turn of the first millennium. The important differences between *ego* and *je* on the one hand, and CL *passum* and OFr *pas* on the other, do not, however, become apparent unless we interrelate the individual pertinent diachronic changes which *together* create this new morphological bond. My aim, then, will be to show, through this integrated perspective which is at the heart of the analytic/synthetic concepts, how essentially unrelated developments conspire over time to change certain elements into units of a different sort.

One of the most fascinating and, by the same token, most intriguing aspects of the cyclic move from analysis to synthesis (or vice-versa) is the question of the origin of these changes. Though considerable effort has been devoted to studying this issue in the past few years, it has remained unresolved. The aim of the last chapter of this study is not only to point up some of the weaknesses of earlier in-

vestigations, particularly those which view analysis and/or synthesis as the result of a somewhat mystical, teleological drift, but also to ferret out the motivating factors in the evolution toward new types of morphological structures. Arguing that analysis and synthesis are simply a consequence (rather than a cause) of language change, I will suggest that every innovation ultimately affects the direction a speech unit takes on the analytic/synthetic axis, and that the frequently mentioned drift of IE languages towards greater analyticity merely reflects a statistical trend whose unidirectional momentum has been exaggerated as a result of a widespread failure to recognize signs of synthesis.

# Abbreviations

## A. *Terminology*

| | | | |
|---|---|---|---|
| acc. | accusative | nom. | nominative |
| ART | article | NP | noun phrase |
| adj. | adjective | obl. | oblique |
| adv. | adverb | PART | partitive |
| AUX | auxiliary | pl. | plural |
| c. | century | P/N | person/number |
| COREF | coreferent/coreferential | pop. | popular |
| dat. | dative | PP | past participle |
| EMPH | emphasizer | pres. | present |
| fem. | feminine | refl. | reflexive |
| l. | line | s. | singular |
| loc. | locative | T/A | tense/aspect |
| masc. | masculine | v. | verse |
| mod. | modern | VP | verb phrase |
| NEG | negator/negative | | |

## B. *Languages and Dialects*

Prefixed to an abbreviation, O = old, mid = middle, mod = modern,
W = West, N = Northern, S = Southern, East. = Eastern, Centr. = Central

| | | | |
|---|---|---|---|
| Am. Sp | American Spanish | Lad | Ladin |
| Amp | Ampezzano | Lat | Latin |
| BP | Brazilian Portuguese | Lomb | Lombard |
| Calab | Calabrian | Mil | Milanese |
| Cat | Catalan | Pal | Palenquero |
| CL | Classical Latin | Piedm | Piedmontese |
| Eng | English | Port | Portuguese |
| FSF | Formal Spoken French | Oc | Occitan |
| Fr | French | PR | Proto-Romance |
| Gal | Galician | Rom | Romansh |
| Germ | German | Russ | Russian |
| Gév | Gévaudanais | Sard | Sardinian |
| IE | Indo-European | Sp | Spanish |
| ISF | Informal Spoken French | Surs | Surselvan |
| It | Italian | VL | Vulgar Latin |

# Chapter one

# History of the terms "analytic" and "synthetic"

## 1.1 Introduction

This chapter is divided into four sections. The first will show how the concepts "analyticity" and "syntheticity"[1] arose in the context of 19th century analyses of morphological typology.[2] The second section focuses on Edward Sapir's writings on analyticity and syntheticity, which not only led to the ossification of the already existing terminology, but also produced a thoroughgoing transformation in the application of morphological criteria to language typology. Section three follows the evolution of Sapir's ideas down to our own times. Finally, the fourth section summarizes the main events in the history of the analytic/synthetic dichotomy.

## 1.2 Nineteenth century morphological typology and the rise and fall of the concepts analyticity and syntheticity

### 1.2.1 The Schlegel brothers

Since the beginning of linguistic study, attempts have been made to classify languages into groups. In general, language classification has been approached in four major ways: genetically, areally, sociolinguistically, and typologically.[3] The genetic approach, which classifies languages according to their historical antecedents was, from Dante to the beginning of the last century, the earliest and only approach seriously considered.[4] When the publications of Friedrich von Schlegel (1772-1829),[5] August Wilhelm von Schlegel (1767-1845)[6] and Wilhelm von Humboldt (1767-1835)[7] on morphological classification appeared in the early 19th century, they not only marked a definite break in a long tradition, but also set the tone for the subsequent scientific approach to language classification.[8]

 Among the three protagonists mentioned here, August Wilhelm von Schlegel, whose brother Friedrich had already set up a twofold classification into languages with affixes and languages with inflection (see *n*5), deserves particular attention for being the first to introduce the concepts of analyticity and syntheticity.[9] In his *Observations sur la langue et la littérature provençales* (1818:14) he suggests a classification into three broad categories:[10]

(1) languages without grammatical structure
(2) languages with affixes
(3) languages with inflection

The first type, exemplified by Chinese, is characterized by roots which cannot be developed or modified. Syntax plays the main role in indicating the relationship among words for the languages of this class. The characteristic trait of the second type, exemplified by the then recently discovered Amerindian languages, is that these vernaculars express secondary ideas and relationships by attaching grammatical elements to other words. The third class, affixing languages such as the classical languages and their modern successors, is typified by the modification of "letters"[11] of the roots and by the addition to the roots of derivational suffixes.[12]

August Wilhelm von Schlegel's classification does not, however, end here. He further subdivides the inflectional type (group 3) into two subclasses, "synthetic" and "analytic":

> Les langues à inflexions se subdivisent en deux genres, que j'appelerai les *langues synthétiques* et les *langues analytiques*. J'entends par langues analytiques celles qui sont astreintes à l'emploi de l'article devant les substantifs, des pronoms personnels devant les verbes, qui ont recours aux verbes auxiliaires dans la conjugaison, qui suppléent par des prépositions aux désinences des cas qui leur manquent, qui expriment les degrés de comparaison des adjectifs par des adverbes, et ainsi du reste. Les langues synthétiques sont celles qui se passent de tous ces moyens de circonlocution. (1818:8)

For A. Schlegel, this subdivision aims above all at distinguishing the classical languages (Greek, Latin, Sanskrit) from their modern successors. The former are synthetic because they express the relationship of one word to another by means of the forms of the words themselves. Analytic languages such as the modern Romance tongues, on the other hand, express these same grammatical relationships by means of separate words or simply by word position. A striking problem in Schlegel's model (a difficulty which, as we shall see, remained largely unresolved until Greenberg's contribution (1960 [1954]) more than a hundred years later) is that the division cannot be applied in absolute terms. A. Schlegel is forced to place the Germanic languages (excluding English which he considers analytic) in an intermediate class since they are, according to him, synthetic in origin but tending heavily to analytic forms: 'En Europe, les langues dérivées du latin, et l'anglois, ont une grammaire toute analytique, ... Les langues germaniques forment un classe

intermédiaire: synthétiques dans leur origine et conservant toujours une certaine puissance de synthèse, elles penchent fortement vers les formes analytiques' (p. 17). As a result, the synthetic/analytic dimension, meant to inject rigor into a newly conceived morphological classification of languages, is from its inception beset by the problem of "more or less."[13]

Though it may never be fully known what exactly led A. Schlegel to propose the synthetic/analytic distinction, some scholars (particularly Greenberg 1974:38) see a correlation between the then prevailing view that agglutinative languages (Indo-European and Semitic) were superior to the isolating ones and the rise of the new analytic/synthetic subclassification. Like his brother Friedrich who claimed that the apparent richness of affixing languages was in reality an impoverished system with a 'subjectiv sonderbaren und mangelhaften Charakter' (52), August Wilhelm von Schlegel's writings are characterized by a strong value judgment. In his view, isolating languages are sterile and lack the kind of organic life typical of the more developed agglutinating tongues. Even more important, A. Schlegel sees a direct relationship between morphological structure and collective human intelligence. He, in fact, specifically states that the switch from synthetic to analytic constructs was brought about by the cultural decline of the post-classical era:

> Elles [synthetic languages] appartiennent à une autre phase de l'intelligence humaine: il s'y manifeste une action plus simultanée, une impulsion plus immédiate de toutes les facultés de l'âme que dans nos langues analytiques. A celles-ci préside le raisonnement, agissant plus à part des autres facultés, et se rendant par conséquent mieux compte de ses propres opérations. Je pense qu'en comparant le génie de l'antiquité avec l'esprit des temps modernes, on observera une opposition semblable à celle qui existe entre les langues. Les grandes synthèses créatrices sont dues à la plus haute antiquité; l'analyse perfectionnée étoit réservée aux temps modernes. (27-28)

Greenberg (1974:38) points out that A. Schlegel must have been 'disturbed by the fact that the modern Indo-European languages have tended to lose the inflections found in the older stages,' thus apparently reverting back to a more primitive stage. To account for this difficulty, Greenberg believes that A. Schlegel saw himself forced to introduce a further subclass among the inflectional tongues, namely the synthetic for the older type and the analytic one for the modern languages. Although it would be difficult to disprove altogether Greenberg's thesis, there is at least some evidence that the transition from synthetic to analytic structures was perhaps not as disturbing to A. Schlegel as Greenberg would have it. We must keep

in mind that since, on the one hand, A. Schlegel viewed his own culture as inferior to that of the classical period, and, on the other, he saw morphological structure as directly reflecting the cultural genius of a given age, it only made sense to him that the modern tongues showed a less perfect structure than that of their ancestors.[14] In a revealing passage, A. Schlegel elaborates on the historical events, namely the barbarian invasions, which led to the breakdown of the once ideal synthetic morphology:

> Les conquérans barbares ... trouvant dans les pays conquis une population toute latine, ou, selon l'expression du temps, *romaine*, furent en effet forcés d'apprendre aussi le latin pour se faire entendre, mais ils le parloient en général fort incorrectement; surtout ils ne savoient pas manier ces inflexions savantes, sur lesquelles repose toute la construction latine. Les Romains, c'est-à-dire les habitants des provinces, à force d'entendre mal parler leur langue, en oublièrent à leur tour les règles, et imitèrent le jargon de leurs nouveaux maîtres. Les désinences variables, étant employées arbitrairement, ne servoient plus qu'à embrouiller les phrases; on finit donc par les supprimer et par tronquer les mots .... Mais ces désinences supprimées servoient à marquer d'une manière très sensible la construction des phrases, et la liaison des idées; il falloit donc y substituer une autre méthode, et c'est ce qui donne naissance à la grammaire analytique. (24-25)

According to A. Schlegel, the transition from a synthetic to an analytic language does not, however, always originate in a contact situation as the one he describes above:

> Lorsque les langues synthétiques ont été fixées de bonne heure par des livres qui servoient de modèles, et par une instruction régulière, elles sont restées telles; mais quand elles ont été abandonnées à elles mêmes et soumises aux fluctuations de toutes les choses humaines, elles ont montré une tendance naturelle à devenir analytiques, même sans avoir été modifiées par le mélange d'aucune langue étrangère. (18)

The two passages above reveal that A. Schlegel may well have opted to subdivide the inflectional languages into an analytic and a synthetic group because two parallel processes (the downfall of the Roman empire and the rise of a more analytic morphology) fit very nicely into his preconceived ideas about the relationship between cultural sophistication and language structure. It is, therefore, unlikely that A. Schlegel's invention came as a response to what Greenberg sees as a

"disturbing" fact, namely that inflectional languages—regarded as the acme of excellence and placed at the head of a hierarchy of classes or types of language—had begun to decay. It is much more likely that A. Schlegel took to the new division simply to refine a classification which, in its essence, had already been proposed by his brother Friedrich.[15]

One can rarely grasp the full import of intellectual achievements without considering the social and scientific climate in which they were accomplished. This is particularly the case for certain of the ideas of 19th century philologists. At a time when the concepts of analyticity and syntheticity were being discussed, Darwin's theory of evolution was having a profound effect on all areas of science, linguistics included.[16] Models devised for biological theories were carried over into theories of language; this led to the notion that languages too had a "life cycle" whose individual stages could be described in evolutionary terms and explained by evolutionary laws. Once it was accepted that languages underwent a "growth" and, like animals and humans, could be traced back to a common ancestor by establishing an genealogical tree, it must have seemed only natural to 19th century comparativists to interpret the newly discovered formal differences among the world's languages as attestations of the intermediate stages through which human speech must have passed.

Keeping in mind too the 19th century's sustained fascination with and admiration of Classical Antiquity (and its associated languages), we can understand more easily why a biological model of language was favored in which "primitive," "formless" languages (i.e., those lacking any signs of growth) such as Chinese were placed at the bottom of the scale,[17] while the "formed" (i.e., organically developed) tongues of the IE type were situated at the top.[18] Following A. Schlegel's line of thought, Sanskrit, Greek and Classical Latin now seemed all the more perfect, for—in addition to being the products of high cultures, the words of these languages could be shown to have "roots" from which the synthetic inflections had grown over time.[19] An ideal language, it was thought, was one endowed with words formed "organically," and the more the parts of these forms were intertwined (i.e., synthetic), the more this was interpreted as a sign of healthy growth. Within such a context the "mechanical" agglutination found in languages such as Tibetan or Polynesian naturally represented an intermediate, imperfect stage, while the more analytic systems of the modern IE languages represented the decay of a system once perfect.[20]

Although August Wilhelm von Schlegel clearly perceived that "perfect" synthetic languages could, and in fact do, revert back to more (primitive) analytic modes of expression (cf. his previously cited account of the evolution from Latin to the modern Romance tongues), the overall impression one obtains from reading his

as well as subsequent works of the 19th century, is that, much like biological evo-
lution, language evolution, and in particular the development "analysis → synthe-
sis," is nonetheless viewed as an essentially *unidirectional* and *irreversible* process.
For as long as the Classical cultures prospered, languages such as Sanskrit, Greek,
or Classical Latin could not help but gravitate uniformly towards the more com-
plete synthetic type. Conversely, during the "post-Classical period of cultural
decline," Romance and other IE tongues were "naturally" pulled down towards more
analytic (and, therefore, less perfect) modes of expression. Such a scheme left no
place for language-internal (or other non-cultural) factors which might stop or even
reverse the momentum of a language towards greater analysis or synthesis.

    As the following chapters will show, A. Schlegel's theory of language change
contains at least three assumptions which are of prime importance for an under-
standing of the subsequent history of the concepts of analyticity and syntheticity.
He claimed, first, that *entire* languages (rather than individual structures) undergo
the typological changes associated with analysis → synthesis; second, that analysis
precludes synthesis (and vice versa), i.e., the two processes do not take place simul-
taneously within a single language; and, third, that the change "analysis → synthe-
sis" is non-cyclical, i.e., languages are not seen as moving back and forth between
these two extremes. I would point out that these assumptions still underlie many
contemporary studies dealing with morphosyntactic typology. The uncritical ab-
sorption of these ideas into modern linguistics has, as we shall see, significantly
added to the confusion surrounding the terms analytic/synthetic ever since the con-
cepts were first introduced.

### 1.2.2 Reactions to Schlegel's analytic/synthetic division

The next significant treatment of our topic comes from the pen of Wilhelm von
Humboldt. Though explicitly rejecting any Schlegelian theory of historical evolu-
tion in which more advanced (i.e., inflectional) languages evolve out of the so-
called primitive ones (i.e., those without structure and/or languages with affixes),
Humboldt, in the spirit characteristic of German Romanticism, continues in the
Schlegelian vein by relating morphological typology to the *Volksgeist*.[21] Most
notably, however, he speaks out against the strict classificatory divisions introduced
by A. Schlegel. In his *Ueber das Entstehen der grammatischen Formen und ihren
Einfluß auf die Ideenentwicklung* (1822), he stresses the fact that in most languages
there exist tendencies for inflection (*Beugung*) as well as agglutination (*Anfügung*)
and that a systematic separation between them cannot be maintained (1963:46, esp.
334ff).[22] Even more important for our purpose are his serious objections to
Schlegel's analytic/synthetic breakdown.

A. W. v. Schlegel hat diese Gattung der Sprachen mit dem Namen der
ANALYTISCHEN, so wie die eines vollständigen organischen und beu-
gungsreichen Baues mit dem der SYNTHETISCHEN belegt, und diese letz-
tere Benennung vorzüglich ist in andere Schriften übergegangen [sic,
A.S.]. Ich glaube mit einigen Worten angeben zu müßen, warum ich mich
derselben absichtlich nicht bediene. Der Name der synthetischen soll zwar
den Unterschied von agglutinirenden bezeichnen, daß die Synthese die ein-
zelnen Theile in Eins verschmelzt, aber jede Synthese setzt immer ein zu
verbindendes Mehreres voraus, und wo ist dies, wenn z.B. aus *binden* ich
*band* wird? eine Lautbeugung, die gerade den feinsten Sprachorganismus
vorzugsweise charakterisirt. Die Zusammenschmelzung in Eins läßt sich
auch nur gradweise unterscheiden. Man kann nicht sagen, daß sie da sey,
oder fehle, sie ist in gewissem Verstande immer vorhanden, nur mehr oder
weniger innig. Der in jede feinste Abschattung der Ideen eingehende Ur-
heber jener Benennungen bemerkt bei den synthetischen und analytischen
Sprachen selbst, daß die Gränzlinie nicht scharf zu ziehen ist, und es paßt
dies noch mehr auf die synthetischen und affigierenden. ... Der Ausdruck
ANALYTISCHE Sprachen scheint mir noch weniger passend. Es geht in den
hier genannten Sprachen nicht sowohl eine Auflösung der synthetischen
Formen vor, als daß man durch Verbindung einiger, unaufgelöst bleiben-
der, andre entbehrlich macht.
    (Humboldt, *Ueber die Verschiedenheit des menschlichen*
        *Sprachbaus* [1827-29], 1963:317-318)[23]

By reading further in Humboldt's extensive work, one discerns that the synthet-
ic/analytic division is rejected not simply because it is too vague to be useful but
also because, in his view, non-genetic language classification is inherently im-
possible:

Allein einer andren und solchen Classification, wo auch die gar nicht
stammverwandten Sprachen nach allgemeinen Aehnlichkeiten ihres Baues
zusammengestellt würden, wiederstrebt .... Die einzelnen Sprachen sind
nicht als Gattungen, sondern als Individuen verschieden, ihr Charakter ist
kein Gattungscharakter, sondern ein individueller. Das Individuum, als
solches genommen, füllt allemal eine Classe für sich. (p. 189)

Whether Humboldt's rejection of Schlegel's scheme was instrumental in the
subsequent work of such prominent linguists as Franz Bopp and August Pott may
only be guessed at today. What is striking, though, is that both Bopp (1833) and

Pott (1833) spent considerable effort in making further contributions to the typologization of languages without incorporating the concepts of syntheticity and analyticity into their models.[24]

A minor detour is in order at this juncture in order to mention the term "polysynthetic," which in later linguistics frequently appears in conjunction with the Schlegelian pair synthetic-analytic.[25] One would perhaps expect that this neologism was the outgrowth of a further refinement of Schlegel's division. This was not, however, the case. The proposal was made in 1819 (i.e., one year after Schlegel's *Observation*) by the American-based linguist Du Ponceau. Arguing that it was yet premature to attempt a classification of the world's languages, Du Ponceau (1838:447) introduced the term polysynthetic not to establish a more refined classificatory system but merely to propose a general label for the Amerindian languages familiar to him:[26]

> Partout où j'ai eu des renseignemens suffisans pour m'assurer de leur caractère, j'ai trouvé que ces langues appartiennent à la classe que j'ai nommée *polysynthétique*, dans la vue seulement de la désigner et sans y attacher autrement aucune importance.

The next one to propose a distribution of languages according to their morphological structure is August Schleicher. Best known for his *Stammbaum* theory of genetic classification, Schleicher reintroduces the 'old' Schlegelian terminology of synthetic and analytic, not, however, without expanding their scope. As a refinement of the basic tripartite division into isolating, agglutinating, and inflectional languages, Schleicher (1850:9-10, 18) proposes to apply the synthetic/analytic subdivision not only to the flexional but also to agglutinative types. It is interesting to note, however, that Schleicher is never explicit about what leads him to extend the subclassification to the agglutinative type; nowhere in his other writings on language classification (Schleicher 1848, 1859, 1871) does he reiterate these ideas.[27] Schleicher's half-heartedness towards the notions of syntheticity and analyticity appears to have had an impact on his contemporaries concerned with language classification. Thus H. Steinthal (1850, 1860), for instance, who, after first being favorably disposed to the Schlegelian concept (1850:11-12), later refrains from implementing it in his own scheme (see his summary table "System der Sprachen als die Entwicklung der Sprachidee" [p. 82]). Others, like M. Müller, simply reject the new ideas, arguing that it is not 'necessary to distinguish between *synthetic* and *analytic* languages' (1862:325) since morphologically, phrases like Fr *j'aimerai* (for *j'ai à aimer*) are still inflectional.[28]

Whatever the value of the various criticisms, in reviewing pre-1875 works concerned with morphological typology, one cannot fail to notice that Schlegel's

synthetic/analytic distinction was given a rough ride.[29] Based on the "fragmentary character" of word structure, the idea of dividing languages into genetically unrelated groups convinced few, and a once debated idea and its terminology seemed to be well on the way out. The lack of scientific precision in Schlegel's "more or less analytic" did not convince the Neogrammarians, who quickly began to make headway in the 1870's; they attempted to found a genuine science based on rigorous methods and exceptionless laws, whence they surely could see no point in establishing a classification into which some languages (e.g., modern Germanic in Schlegel's model) could not be fitted. Furthermore, as Greenberg (1974:40) succinctly points out, in the Neogrammarians' 'concentration on the specific phenomenon of languages ... there was no room for the classification of languages as wholes.' Morphological typology was, to be sure, quickly becoming a part of the past, and was left to a few loners such as F. Misteli (1893)[30] and F. Fink (1901 and 1910),[31] neither of whom embraced the Schlegel-Schleicher breakdown into analytic/synthetic languages. The net result was that the *word* and its morphemes, at the center of interest for over half a century, lost its once unchallenged position. Those like Schlegel who had toyed with a possible refinement of the basic tri- or quadri-partite division into synthetic and analytic subgroups, had apparently failed to provide a program that could secure for these concepts a spot in future linguistic discussion: analyticity and syntheticity, never firmly grounded from the day of their inception, were being forced into a quick and—at least from a late 19th century perspective—irrevocable retirement.[32]

### 1.2.3  The use of "synthetic" and "analytic" by Georg von der Gabelentz

Gabelentz (1881/1891 [1901]) uses "synthetic" and "analytic" not as classificatory concepts for the description of different language types, but rather as a complementary distinction between two existing systems of grammar coexisting within a single language.[33] The analytic system is fairly traditional and resembles modern structural grammar in that it takes the sentence as the basic unit and segments it into successively smaller constituents, ultimately arriving at the smallest elements of meaning. The synthetic system, on the other hand, rests on what was then an entirely different concept: it sought to establish, *inter alia*, how a given thought is formulated by the speaker. Put in Saussurian terms, Gabelentz's synthetic grammar thus takes the *signifié* as point of departure and attempts to show with the help of a series of quasi-transformational rules how and under what circumstances one arrives at a given *signifiant*.[34] In dealing with parts of speech, for instance, such a synthetic grammar would typically try to elucidate the following questions: Through what means are certain segments of a sentence replaced by other segments? Under

what circumstances can certain segments be dropped? What determines the choice between alternations of the type Fr *hier était le vingt deux* vs. *le vingt-deux était hier* 'yesterday was the twenty-second'? How do such (nearly) synonymous constructions differ from each other? (1878:660-661, 1901:101-103)

The author never attempts to present his synthetic grammar as a new theory of language. His motivation for positing the analytic/synthetic division is above-all practical for it is meant to facilitate both the description and acquisition of non-IE languages. In his *Chinesische Grammatik*, where the analytic/synthetic division finds its first large-scale application, he stresses the utilitarian value of his approach, and in so doing explicitly denies any scientific value to his synthetic method:

> Neu ist dagegen mein synthetisches System. Man kann den praktischen Werth eines solchen sehr wohl empfinden, ohne seine wissenschaftliche Berechtigung anzuerkennen. (1881:xi)

Gabelentz acknowledges that the perspectives of both the analytic and synthetic grammars are essential to understanding and teaching the mechanics of speech production (pp. 479-481), but underlines the importance of the synthetic grammar for handling the finer points of a language:

> Es gilt, sich in der fremden Sprache richtig, womöglich gewandt aus-drücken zu lernen. So muß man über eine möglichst vollständige Syn-onymik verfügen und die Unterschiede zwischen den Synonymen möglichst genau kennen. So aufgefaßt ist das zweite grammatische Sys-tem [i.e., synthetic grammar] geradezu die hohe Schule der Sprachlehre,— sollte sie wenigstens sein. (1888:xii)

Gabelentz argues that in addition to the didactic advantages of his system of grammar, its unique division into analytic and synthetic parts provides an alternative and improved way of describing those languages (e.g., Chinese) for which the traditional grammatical model, based on the structure of Latin and Greek, is ill suited.

Gabelentz's innovative suggestions, in particular his peculiar use of "analytic" and "synthetic," did not impress his contemporaries, nor was his framework ever adopted by 20th century linguists. Although his idiosyncratic application of the analytic/synthetic distinction is ultimately of little import to the overall history of these concepts, Gabelentz's work is nonetheless important to an understanding of their history. For their use as something other than taxonomic terms makes clear that by the end of the 19th century Schlegel's division of languages into an analytic and synthetic type was all but forgotten.

## 1.3 Edward Sapir and the renaissance of an old idea

The Neogrammarian school had such a profound, long-lasting and negative impact on typological studies that the Schlegel-Schleicher division of languages into synthetic and analytic types was abandoned for over half a century. A new, radically different, and powerful movement which would sweep across linguistic studies— namely structuralism—stood as yet beyond the horizon. With respect to these upcoming developments and the Neogrammarian school, Sapir, in his much celebrated book *Language* (1921), became a transitional figure between the old and new. Transitional because, inter alia, he returns to an old and by then forgotten idea (analyticity/syntheticity) while introducing hitherto unknown techniques (e.g., the matrix presentation of language features [see below]), which in their basic conception announced future structuralism. Though obviously drawing his analytic/synthetic scheme from Schlegel and Schleicher, he rejects the subjective evolutionary aspects which typified his predecessors.[35] His classification of language types is complex and complicated,[36] undoubtedly because he realizes that the simple traditional tri- or quadri-partite morphology-based division of the 19th century typologists is a strait-jacket fitting only a few, if any, better known languages.[37] He, therefore, distinguishes not one but three separate axes of classification: (1) grammatical concepts, (2) grammatical processes, and (3) firmness of affixation. This third, and most important distinction in our discussion, concerns 'the *relative* firmness with which the affixed elements are united with the core of the word' (p. 135, emph. mine). Sapir (pp. 135-136) establishes the following three degrees of affixation:

(1) ANALYTIC:   A language that does not combine concepts into single words at all (Chinese) or does so economically (English, French). In an analytic language the sentence is always of prime importance, the word is of minor interest.

(2) SYNTHETIC:   In a synthetic language (Latin, Arabic, Finnish) the concepts cluster more thickly, the words are more richly chambered, but there is a tendency, on the whole, to keep the range of concrete significance in the single word down to a moderate compass.

(3)   POLYSYNTHETIC: More than ordinarily synthetic. The elaboration of the word is extreme. Concepts which we should never dream of treating in subordinate fashion are symbolized by derivational affixes or "symbolic" changes in the radical element, while the more abstract notions including the syntactic relations, may also be conveyed by the word. A polysynthetic language illustrates no principles that are not already exemplified in the more familiar synthetic languages. It is related to them very much as a synthetic language is related to our own analytic English.

Sapir (p. 136) stresses that 'the three terms are purely quantitative—and relative, that is, a language may be "analytic" from one standpoint, "synthetic" from another.' To illustrate the relativity of his classification, Sapir (136 *n*11) proceeds to explain that English, typed as analytic (see above), is analytic only in tendency since, relative to French—at least in certain aspects—it is still fairly synthetic. Sapir, in other words, does not aim at establishing a rigid classification, i.e., one in which a language must fit into either the synthetic or the analytic compartment, but rather seeks to highlight *structural tendencies* (the famous "drifts")[38] within a given language.[39]

By employing three sets of distinctions (conceptual type, technique, and degree of synthesis) Sapir offers a wide variety of possibilities for language characterization.[40] His matrix (p. 151, reproduced here as Table 1), with all its complex possibilities of variation,[41] contains an unprecedented combination of features, some of which are mutually exclusive. Thus the languages given under his type A are necessarily analytic. Languages of type C are also predominantly analytic and are, according to Sapir (p. 148), not likely to develop beyond the synthetic stage.

Sapir does not limit his attention to merely establishing parameters for his morphology-based matrix. He clearly goes beyond such considerations when he attempts to draw inferences from the relationship between diachrony and structural type of language structure. He argues, for instance:

it is interesting ... to note that of the three intercrossing classifications represented in our table (conceptual type, technique, and degree of synthesis), it is the degree of synthesis that seems to change most readily, that the technique is modifiable but far less readily so, and that the conceptual type tends to persist the longest of all. (154)

Earlier he comments:

> it is often illuminating to point out that a language has been becoming
> more and more analytic in the course of its history or that it shows signs
> of having crystallized from a simple analytic base into a highly synthetic
> form. (136)[42]

Despite Sapir's vastly improved methodology of basing language classification
on non-genetic criteria, and despite the sophisticated and original incorporation of
analyticity and syntheticity into a multi-layered matrix, it is easy to see why his
scheme might be open to serious criticism. In the tradition of Schlegel and
Schleicher, Sapir provided a sliding scale of syntheticity which relied heavily upon
vague and arbitrary criteria. His highly subjective qualifying notations of "mildly
synthetic," "mildly polysynthetic," "notably synthetic," "very nearly complex pure-
relational," etc. simply could not satisfy those who continued to demand rigorous
and unambiguous taxonomic parameters.[43] Moreover, Sapir's strange rejection of
his own thoughts on the analytic/synthetic concept on the pages immediately fol-
lowing his exposition thereof, help to show that Sapir himself eventually discarded
his initial enthusiasm for the analytic/synthetic division.[44]

Less obvious shortcomings of Sapir's chapter "Types of Linguistic Structures"
came to light in the years following publication of his book. Velten (1935) pointed
out that it was never entirely clear whether semantic, syntactic, or morphological
criteria (or a combination of these) were to be applied in determining the degree of
analyticity and/or syntheticity.[45] Moreover, he objected to Sapir's analytic/synthe-
tic division because it would be wrong to distinguish a synthetic from an analytic
language based on the following juxtaposed expressions: *reiten – equo vehi*,
*patiner – Schlittschuh laufen, Schimmel – cheval blanc, naschen – to eat dainties
surreptitiously* (p. 6). For Velten, a language classification based on such principles
is inadequate, all the more so because the most fundamental criterion for determin-
ing the degree of synthesis and/or analysis is based on a concept, the word, for
which a satisfactory definition had never been given (p. 21).[46]

Whatever the apparent shortcomings of Sapir's work on language typology, and
however vague his use of the terms analyticity and syntheticity, the Schlegelian
terminology did, as we shall see in the following section, once again make its way
into linguistic discussion; there can be no doubt that the immense popularity of
Sapir's *Language* contributed significantly to the spread of a now very popular ter-
minology which might otherwise have fallen into oblivion.[47]