



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

”Syntactic Conspiracies in a Diachronic Perspective“

Verfasser

Mag. rer. nat. Dr. rer. nat. Wolfgang Brozek

angestrebter akademischer Grad

Magister der Philosophie (Mag. phil.)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 328

Studienrichtung lt. Studienblatt:

Allgemeine/Angewandte Sprachwissenschaft

Betreuer:

Prof. Mag. Dr. Hans Martin Prinzhorn

Acknowledgment

I am first and foremost indebted to Ass.-Prof. Mag. Dr. Hans Martin Prinzhorn for supervision of the present Thesis and excellent support with relevant literature.

Moreover, I am thankful to emer. O. Univ.-Prof. Mag. Dr. Wolfgang U. Dressler, Ao. Univ.-Prof. Dr. Chris Schaner-Wolles, and Ao. Prof. Dr. John Rennison for helpful discussions on the topics of the present work. Prof. Dressler in particular, through his critical and constructive comments, contributed substantially to improvement of this Thesis.

Further thanks go to Ass.-Prof. Mag. Dr. Hans-Christian Luschützky and Mag. Lic. Dr. Carlos-Jorge Lettner for valuable recommendations and ideas that were well included in the present work.

For years of vital support whatsoever, I finally wish to express my deepest gratefulness to my mother.

Table of Contents

Abstract	1
1. Preview and outline	2
2. Acquisition-based approaches to diachronic change	4
2.1. Past and present concepts of acquisition-based diachronic change	4
2.2. Language acquisition-based computational models of diachronic change	7
2.2.1. <i>Evolutionary model of learnability (Clark and Roberts, 1993)</i>	7
2.2.2. <i>Evolutionary systems model (Niyogi and Berwick, 1995, 1997, 1998)</i>	10
2.2.3. <i>A variational model of grammar competition (Yang, 2000)</i>	12
2.2.4. <i>Models of iterated learning</i>	15
2.2.5. <i>Evaluation of acquisition-based computational approaches to diachronic change</i>	17
3. From language universals to conspiracies	19
3.1. Concepts of language universals and linguistic typology	19
3.2. Syntax-related implicational universals of language: Greenberg and the aftermath	25
3.3. Adding the diachronic dimension: the notion of conspiracies	32
4. Language acquisition and the emergence of conspiracies	36
5. Syntactic conspiracies exemplified	39
5.1. Conspiracies of VO and OV word order	39
5.2. Transitive expletives and object shift in Germanic	46
5.3. V2 word order and <i>pro</i> -drop: a negative conspiracy in Western Romance	49
6. Discussion and conclusions	53
6.1. Classification of conspiracies	53
6.2. Factors explaining (syntactic) conspiracies	55
6.3. The role of language acquisition and conspiracies in diachrony	61
6.4. Concluding remarks	63
7. References	65
Appendix I: Deutsche Zusammenfassung (German abstract)	71
Appendix II: Scientific CV	

Abstract

Empirical, mainly synchronic, studies have repeatedly established cross-linguistic co-occurrences or negative associations of two or more features, notably also in syntax. Joseph H. Greenberg's publication on predominantly implicational language universals from the 1960s figures amongst the most influential of such works for syntax, besides exerting considerable impact on language typology. The present work claims one principal explanation for syntactic co-occurrences in regularities during the process of language acquisition that shapes diachronic progression and change. In this regard, various important quantitative statistical and computational models that capture both language acquisition and change are presented and discussed. Following up the importance hence attributed to the study of diachronic change, associations found between linguistic, specifically syntactic, properties are labeled "conspiracies" if evidence demonstrates their affinity (or incompatibility) not only in synchronic analysis, but also throughout diachronic development. Several synchronic and diachronic examples mainly from Indo-European languages illustrate this concept, including correlations with either VO or OV word order, transitive expletive constructions and object shift, and the mutual exclusion of V2 word order and the occurrence of null-subjects as a negative conspiracy. Own statistical analysis adds to the value of presented results. Furthermore, a detailed discussion on probable and conceivable causes for the emergence of conspiracies highlights the role of cognitive-psychological factors operative during language learning. This is exemplified by an experimental artificial language learning study in adults that yielded results predicted by one of Greenberg's implicational universals and by a principled tendency of harmony across syntactic categories, the latter of which was described to underlie the notion of language "drift" put forth by Edward Sapir in the 1920s. However, also other causes for diachronic change and emergence or absence of (expected) conspiracies are discussed, including language contact, inheritance of properties amongst related languages, and socio-cultural factors, and the importance of population dynamics for effecting diachronic change is clarified.

Key words: syntactic conspiracies – language acquisition – language learning – diachronic change – language universals – language typology – implications – syntactic structure – computational simulations

1. Preview and outline

Based on observations of recurrent pairs or clusters of syntactic properties throughout languages of the world, major goals of the present work are to explore causes of such associations with a focus on language acquisition and implications for diachronic evolution, and to trace the historic developments of some syntactic features that appear to be “tied together”.

The section immediately following this outline is concerned with explanations for diachronic change only in brief, before laying emphasis on influential concepts of the past one-hundred and thirty years, embarking on ideas by neo-grammarians Hermann Paul, that attempt(ed) to link language acquisition and change, and including the generative concept of parameter setting during language acquisition. A further subsection is dedicated to the introduction to several statistical and computational models that simulate diachronic change through language learning, notably all more or less inspired by evolutionary biology, amongst those, iterated learning models influenced by population genetics in particular.

Section 3, in the first subsection, will provide an overview on different concepts of language universals and, to a lesser extent, language typology, placing Joseph H. Greenberg’s notion of implicational universals (be they statistical or exceptionless), published in the 1960s, at center stage. Greenberg’s findings and postulations for syntax-related linguistic universals are presented in the subsequent subsection, and their aftermath up to the present time, including studies stimulated by his work and modifications of his universal statements, is presented and discussed. A further subsection that introduces the concept of conspiracies as conspired traits or elements in the guise of associations or correlations harboring, next to the synchronic dimension, also the aspect of diachronic correlation, concludes this section.

Section 4 is designed as a first glance at the intricate relation between language learning and the emergence of regularities such as linguistic conspiracies. As evidence for the existence of such a connection, work is presented and discussed from the creation of creole and sign languages and the emergence of linguistic, more specifically syntactic, structure in artificial language learning paradigms as simulations of language evolution.

A selection of syntax-related conspiracies exemplified with synchronic and diachronic data alike from published studies represents the content of section 5, supplemented with own statistical analyses conducted in SPSS, versions 11.5 and 19. As one example, a report on the synchronic correlation of syntactic phenomena with either VO or OV word order precedes diachronic analysis of two of such phenomena, *i.e.* syntactic position of the auxiliary and the adposition, in the development of Romance, English, and Icelandic. Furthermore, a

conspiracy of object raising with transitive expletives, traced to movement of the main verb and the subject out of VP in Germanic languages, will be discussed. Finally, the incompatibility of V2 word order and the omission of pronominal subjects (*pro-drop*) in the diachrony of Western Romance will be demonstrated, including particularly data from French, and explanatory approaches will be presented.

The last Section 6 is reserved to general discussion and conclusions, at the outset providing a tentatively detailed classification of conspiracies, expanding on the general statements about them in Section 3. The following subsection proceeds towards a detailed discussion of possible causes for the occurrence and emergence of conspiracies, which include factors that facilitate language processing, and linguistic relatedness and contact. Moreover, explanations given for syntactic principles described by John Hawkins, *i.e.* Heaviness Serialization and Mobility Principles and Cross-Category Harmony, will be debated as bearing relevance for causes of conspiracies, encompassing psycho-linguistic, syntax-internal and syntactic-semantic, and language historical factors, as well as analogy. In order to underscore the role of cognitive-psychological factors embracing language learning, one recent experimental study is introduced in detail that demonstrated the formation of syntactic structure in congruence with both the prediction of the Cross-Category Harmony principle and one of the language universals put forth by Greenberg. The subsequent subsection relates the notion of linguistic drift as defined by Edward Sapir in the 1920s to the tendency for syntactic harmony embodied by the Cross-Category Harmony principle. Further aspects will be the role of confounding criteria like socio-cultural factors and language contact interfering with this principle of harmony, thus shaping diachronic development, and a discussion of whether individual mismatch learning suffices to be a relevant factor driving diachronic change. A concluding paragraph representing the last subsection summarizes the main aspects of the present work.

2. Acquisition-based approaches to diachronic change

2.1. Past and present concepts of acquisition-based linguistic change

In order to interpret divergence and convergence of linguistic features across languages and dialects to the end of explaining linguistic variation, sound knowledge on how development occurred over time, both descriptively and functionally, is indispensable, being the chief goal of diachronic investigation. In attempts to reveal the driving forces that govern language change, a number of factors were proposed to be involved: Crucially, language and dialectal contact as well as sociological phenomena internal to a society were shown to undoubtedly play their salient parts (*e.g.*, Labov, 1994). But also processes internal to a linguistic system or its speakers have been implicated in why and how languages change with time, in particular related to conflicts amongst linguistic subdisciplines as accounted for by the polycentristic approach (Dressler, 1977) and to psycholinguistic cognitive mechanisms like language acquisition: Originally, the formulation of the concept to causally link two seemingly unrelated processes, *i.e.* language change which takes place at the community or population level and (typically first) language acquisition, as accomplished by the individual language learner, dates back to Hermann Paul (1886) who claimed that events during language learning are of utmost importance to changes in the use of language occurring from one epoch to another.¹

In the 20th century, one of the most concise articulations of the idea that diachronic alterations are shaped through the process of language acquisition was put forward by Traugott (1972: 9) stating that “the fact that each generation, or rather each child, learns the language anew and makes its own hypotheses about the patterns of that language is the main cause for language change”. The principle of this concept was similarly taken up by Andersen (1973) in an article on changes in Czech dialectal phonology, sketched in a popular schematic representation, according to which speakers of the preceding generation’s grammar G_1 produce a defined output that serves as blueprint for the next generation to infer from it a grammar G_2 that might differ from G_1 , giving in turn rise to output which differs from that of

¹ *cf.* Paul (1886: 31): „[...] Aber die hauptperiode der beeinflussung ist doch die zeit der ersten aufnahme, die spracherlernung. Diese ist principiell von der sonstigen beeinflussung nicht zu sondern, erfolgt auch im allgemeinen auf die gleiche weise; es lässt sich auch im leben des einzelnen nicht wol ein bestimmter punkt angeben, von dem man sagen könnte, dass jetzt die spracherlernung abgeschlossen sei. Aber der graduelle unterschied ist doch ein enormer. Es liegt auf der hand, dass die vorgänge bei der spracherlernung von der allerhöchsten wichtigkeit für die erklärung der veränderungen des sprachusus sind, dass sie die wichtigste ursache für diese veränderungen abgeben. Wenn wir, zwei durch einen längeren zwischenraum von einander getrennte epochen vergleichend, sagen, die sprache habe sich in den und den punkten verändert, so geben wir ja damit nicht den wirklichen tatbestand an, sondern es verhält sich vielmehr so: die sprache hat sich ganz neu erzeugt und diese neuschöpfung ist nicht völlig übereinstimmend mit dem früheren, jetzt untergegangenen ausgefallen.“

the preceding generation. Next to, quite remarkably in this model scheme of phonological change, placing the focus on grammar² in the transmission of language across generations,³ Andersen defined as mechanism an abductive mode of inference of the target grammar utilized by the language learners: Invoking a distinction described by scientist and philosopher Charles S. Peirce (*cf.* Reilly, 1970), in contrast to deductive and inductive inference, an abductive process concludes a case based on an observed result and a given law. According to Roberts's interpretation (2007: 124) then, provided that the law is Universal Grammar (UG) and the result is an input utterance from the corpus, thus the linguistic environment, it is the learner's abductive task to attain the target grammar, as a case of UG. What is more, this abductive mechanism of grammar acquisition, as opposed to inductive and even more to deductive learning, is apparently highly prone to errors and mismatches, thus being claimed an important source of creative innovation.⁴ This approach of mis-learning as prime source of language change also figures in other notions at about the same time: Kiparsky (1974), *e.g.*, invokes imperfection in learning to account for this idea, whereas Parker (1976) views as viable mechanism for linguistic change the individual learner's assignment of different structural descriptions to input utterances, *i.e.* the learner's mis-assignment of constituent structure.

The idea of diachronic change through language learning was furthermore comprehensively expounded by Lightfoot (1979) who notably defined grammar change as the prime trigger of language change, more accurately, he advocated reanalysis as the essential mechanism involved.⁵ In brief, Lightfoot traces grammatical change to opacities, giving rise to ambiguities in the structure of a language, that undergo re-modeling to render them perceptively transparent by the process of reanalysis, being the essence of what he termed the

² Lightfoot (1991: 4) provides an account of what is generally conceived of as "grammar": "It has long been generally agreed that linguistic expressions are made up of subunits and have an internal hierarchical structure. It is also generally agreed that a grammar (in the sense defined) is not just a list of expressions but is a finite algebraic system that can "generate" an infinite range of expressions." Alternatively, Jackendoff (2002: 125) defines "grammar" as a tripartite organization constituted by three domains, *i.e.* phonological, syntactic, and conceptual (the latter corresponding to semantic) structure.

³ "What is needed is a model of phonological change which recognizes, on the one hand, that the verbal output of any speaker is determined by the grammar he has internalized, and on the other, that any speaker's internalized grammar is determined by the verbal output from which it has been inferred." (Andersen, 1973: 767)

⁴ However, also inductive and deductive processes play a role in grammar acquisition: Once the learner has abductively guessed a target grammar, (s)he inductively checks its compatibility against input utterances, just as the learner's own utterances are put to the test of conformity with grammatical laws by deduction (*cf.* Andersen, 1973: 776-778).

⁵ Reanalysis is defined by Campbell (2004: 284) as a mechanism that, without modifying surface manifestation, changes the underlying structure of a syntactic construction, embracing constituency, hierarchical structure, grammatical categories and relations, and cohesion; reanalysis importantly depends upon the speakers' possibility to assign more than one analysis to a given construction.

Transparency Principle. In this regard, the author claims that, in the attempt to discover possible changes, the focus of attention in historical linguistics must shift from a theory of change to a theory of grammar that imposes restrictions on such possible changes (*cf.* Lightfoot, 1979: 152, 153). In a later work, the same author advocates a link between the setting of parameters during grammar acquisition and language change (Lightfoot, 1991): In this view, along the lines of up-to-date linguistic theory, Universal Grammar is deemed to be constituted by principles and parameters that are to be set by an appropriate environmental linguistic input during language acquisition. Pivotal for this process is the selection of binary values pertaining each to corresponding parameters that are provided *a priori* by UG. This concept represents the Principles and Parameters (P&P) approach that was elaborated in detail elsewhere (*cf.* Chomsky and Lasnik, 1993; Chomsky, 1995). While most work on P&P is concerned with defining which is a valid parameter for UG and which is not, Lightfoot (1991) attempts to pinpoint the mechanisms at work in setting of (undisputed) parameters by drawing evidence from diachronic material from which he concludes on the triggers from the linguistic environment, *i.e.* the primary linguistic data, that must have elicited the setting of certain parameter values. In a nutshell, Lightfoot advocates simplicity and robustness, the latter comprising saliency and frequency, as key properties of such relevant triggers that be devoid of complex and embedded structure (but may offer access to embedded clauses if these are bound by a matrix clause), lack information on negative expressions, and come sufficiently discernible, either qualitatively (*i.e.*, salient) or quantitatively (*i.e.*, frequent), or both. Several examples of diachronic changes, primarily from throughout the Middle English period, serve to corroborate Lightfoot's claims about the nature of parameter setting during language acquisition, encompassing syntactic phenomena⁶ like verb-order change towards SVO, the rise of exceptional case marking structures,⁷ and the emergence of auxiliary verbs, as well as alterations related to morphology like loss of dative case and the development of a structural case system (all from Middle English), and ongoing dialectal loss of inflection markers suffixed to infinitives in present-day Brazilian Portuguese.

To-date, in accordance with aforementioned ideas, acquisition of the first language grammar, *i.e.* “internal language” or “I-language” in the sense of Chomsky (1986b), is considered the critical step for language to undergo change (van Kemenade, 2007). In this vein, during the last decades, a number of works in support of this concept have emerged that

⁶ According to Patel (2003), syntax is defined as a set of principles governing the combination of discrete structural elements into sequences.

⁷ Exceptional case marking (ECM) describes the case of a subject of an infinitival subordinate clause marked with object morphology, much akin to what is known as Acl (accusativus cum infinitivo) in Latin (Chomsky, 1981; Chomsky, 1986a: 85-87).

seek to capture the relationship between language acquisition and change as evidenced by diachronic variation not only descriptively but intriguingly also by quantitative modeling including stochastic and computational models and simulations (Clark and Roberts, 1993; Niyogi and Berwick, 1995, 1997, 1998; Yang, 2000; Griffiths and Kalish, 2007; Kalish *et al.*, 2007; for reviews, *cf.* Nowak *et al.*, 2002, van Kemenade, 2007; Kirby *et al.*, 2004). Importantly, all such models are faced with the dilemma to bridge the logical problem of language acquisition and the logical problem of language change, seemingly a paradox: On one side, the former relates to the challenge to describe properly how the learner, utilizing a certain algorithm, converges correctly to the target input grammar, *i.e.* the concept of learnability. On the other side, the latter, as a direct consequence, takes account of the fact that if language learners always flawlessly copied the parent generation's grammar, no language change would be possible at all (*cf.* Clark and Roberts, 1993; Niyogi and Berwick, 1995, 1997, 1998). The following section addresses how such simulations cope with this paradox, concluded by a comparative evaluation of the presented models.

2.2. Language acquisition-based computational models of diachronic change

2.2.1. Evolutionary model of learnability (Clark and Roberts, 1993)

Considerable inspiration for learning-based models of diachronic variation has been repeatedly drawn from evolutionary biology spearheaded by Wallace's and Darwin's groundbreaking ideas (Darwin and Wallace, 1858; Darwin, 1859) on mechanisms underlying the great variety and evolutionary relationship of biological taxa. One of the pioneering accounts that combined language learning with diachronic change including aspects of natural selection was put forth by Clark and Roberts (1993). Therein, borrowing methodologically from evolutionary biology, a genetic algorithm was employed that filters out genetic material (corresponding to parameter settings) that adapts best to the environment (*i.e.*, the linguistic input), thus mimicking natural selection. Central to this model is the attribution of "fitness metrics" to the learners' linguistic hypotheses represented by strings of parameter settings that determine the success in analyzing the input text and hence the chances of such strings to "reproduce" and "survive" in a given linguistic environment. Beyond the fitness metric and parameter settings encoded by strings of binary numbers, the algorithm also accounts for reproduction operators that gauge success of the learners' hypotheses in relation to the linguistic input. Two principal processes pertaining to reproduction are at work which recombine and change existing learners' hypotheses, *i.e.*, respectively, crossover entailing the

combination of a new hypothesis out of parts of two parent hypotheses, and mutation resulting in newly and randomly generated hypotheses similar but not identical to the old ones the process has operated on (this again in direct analogy to molecular genetic mechanisms operating in genomes thus creating variability susceptible to the forces of evolution). Functionally then, values obtained from the fitness metric weigh the learners' hypotheses and provide guidance in applying the reproduction operators.

Specifically, a fitness metric can be calculated for each hypothesis (expressed as a string of binary figures) by which the learner parses the input data. Such hypotheses, termed “parser” or “parsing device”, consist of modules like case, binding, theta-roles, and X-bar theory, that can sustain violations in response to the linguistic input. The number of violations detected by a given parser v_i are summed up and subtracted from all violations detected in all parsers $\sum_{j=1}^n v_j$, related to the total of violations $\sum_{j=1}^n v_j$. Essentially, two additional weighed variables are introduced accounting for the subset condition (*cf.* Berwick, 1985) and the measure of “elegance”, respectively: The subset condition, which posits that the learner, provided with no negative evidence from the linguistic environment, guesses the smallest possible (subset) grammar rather than its corresponding superset grammar at each step in the learning process, is represented by the difference of all counts of superset settings $\sum_{j=1}^n s_j$ and the number of such superset settings s_i of a given hypothesis, divided by the total $\sum_{j=1}^n s_j$ and weighed by the superset penalty constant b . Likewise, the “elegance” criterion, which postulates the learner's preference of “compact” representations featured by low numbers of nodes to cover input strings involving short chains of arguments and operators in each parse, is allowed for by the difference of the total count of nodes in all parsers $\sum_{j=1}^n e_j$ and the number of nodes in a given parser e_i , weighed by a scaling factor for elegance of representation, the constant c . Finally, to yield the fitness metric for a given hypothesis, the variables are combined by summation and stratified for the total number of parsers save the one in question ($n - 1$):

$$\frac{(\sum_{j=1}^n v_j + b \sum_{j=1}^n s_j + c \sum_{j=1}^n e_j) - (v_i + b s_i + c e_i)}{(n - 1) (\sum_{j=1}^n v_j + b \sum_{j=1}^n s_j + c \sum_{j=1}^n e_j)} \quad (1)$$

Estimates of typical values for the constants b and c indicate their being very small, consistent with their disproportionately little weight relative to parsing failures as represented by violations to the linguistic input. From empirical work on parametric change in the history of French, the authors furthermore conclude that elegance seems to be more important to the

learner than superset settings, thus $c > b$. However, in cases of contradictory input data manifested by ambiguous grammatical violations, the model predicts that it is the elegance measure and the subset condition criteria that urge the learner towards a preferred setting of parameters. On the other hand, more than one parsing device can successfully and correctly analyze a given input sentence with a set of parameters. Therefore, several grammars expressed as binary parameter strings might be compatible with that sentence. These can be summarized in one string that retains the binary numbers common to all strings that successfully parse the input sentence, accompanied by asterisks on the variable positions. Such strings are referred to as P-encodings for a given sentence representing its parameters and their corresponding values. P-encodings serve furthermore as a useful notation for the determination of parameter value frequencies in texts; if a correct value makes frequent appearance in the input data, it has fair chances for a high fitness value in the population.

Collectively, the mechanism by which the fitness metric works can best be summarized by quoting the authors (Clark and Roberts, 1993: 316): “[...] relative to an input text, the fitness metric drives the learner toward a hypothesis that minimizes the number of violations and the number of superset settings and that generates the most elegant syntactic representations possible, given that grammatical violations are avoided.” Stated otherwise, the fitness metric, by assigning fitness values to parameter settings, bestows a learner’s hypothesis about a grammar with a certain probability to be selected for “reproduction” in a given linguistic environment, with respective bearings on the fitness of linguistic hypotheses generated by the learner: While hypotheses displaying incorrect parameter values exhibit lower reproduction rates which concurrently entails eventual elimination of such parameter values from the population, hypotheses showing the correct parameter settings are propagated more readily resulting in the prevalence of these parameter values in the population. Authors conceive of the application of a genetic algorithm to diachronic data as the solution to the logical problem of language change. To be sure, in a case study of language change during Old and Middle French, authors demonstrate, by constructing P-encodings from five parameters (the same ones as referred to in the model of Yang, 2000), the applicability of the fitness metric to explain loss of three syntactic features in these periods, *i.e.* simple inversion in interrogatives, null subjects (*i.e.*, the dropping of subject pronouns), and verb-second (V2)⁸ word order. Importantly, these features did not vanish independently from each other, as will be discussed in more detail below (*cf.* section 5.3.).

⁸ V2 (“verb-second”) denotes the appearance of a verbal element in exactly the second position of a sentence (*cf.* Niyogi and Berwick, 1997: 716)

2.2.2. Evolutionary systems model (Niyogi and Berwick, 1995, 1997, 1998)

According to the model put forward by Niyogi and Berwick (1995, 1997, 1998), the paradox of the logical problem of language change is tackled in a systems model likewise drawing on evolutionary biology: Therein, language acquisition and change are related in an analogous manner to the biological development of the individual, *i.e.* ontogeny, and, respectively, to the evolution of a whole population or species, *i.e.* evolutionary population biology and phylogeny. A further analogy of this concept related to biology posits that initial individual grammatical knowledge corresponds to an individual's genotype, whereas finally attained states of the grammar reflect an individual's phenotype. Along these lines, diachronic language change is formally modeled based upon the assumption of slight variation amongst the outcomes of individual grammar acquisition from one generation to the next. The consequences entail a heterogeneous distribution of linguistic knowledge in the offspring generation. If to a set of parameters prone to linguistic change, distribution probabilities are assigned in the initial generation, resultant scenarios and change patterns in favor of certain directions of development can so be calculated for subsequent generations.

To construct a computational approach that maps language acquisition to change, three main components are requisite for the model (summarized in Niyogi and Berwick, 1998: 194): First of all, there is a set of grammars G of which the language learner selects one, g , based on input sentences; furthermore, an algorithm is needed for implementing this choice; and lastly, a probability distribution P describing the likelihood with which input sentences appear to the learner is required. More into detail, Niyogi and Berwick (1997) employ the triggering learning algorithm (TLA) as previously described in Gibson and Wexler (1994), a memoryless learning algorithm that converges to the target grammar by constantly updating its grammatical knowledge (*i.e.*, its grammatical hypothesis) with every sentence. For calculation of the probability to acquire each combination of grammatical parameters, the TLA is modeled by a first order Markov chain that can assume 2^n states, each representing a distinct grammar, with n the number of parameters. Alternatively, a Maximum Likelihood learning algorithm can be suited equally well as statistical technique for probability estimation (Niyogi and Berwick, 1995, 1998). If then only one (binary) parameter is accounted for that takes on the one or the other value in two grammars g_1 and g_2 , corresponding to languages L_1 and L_2 , respectively, the proportion of the population that speaks L_1 is denoted by p , whereas $1-p$ equals the fraction speaking L_2 . In this mathematical framework, the linguistic composition of the subsequent generation, captured in p_{n+1} following the composition of the start population which is p_n , can be derived from different amounts of input sentences N that

shape the grammatical hypothesis of language learners. Clearly enough, the number of N rises in parallel with maturation time. The following equations (1) for $N = 2$ and (2) for $N = 3$ thus describe development of L_1 from distribution p_n to p_{n+1} with a and b denoting the proportion of sentences common to both L_1 and L_2 , used by learners of either L_1 representing target grammar g_1 , or L_2 representing target grammar g_2 , respectively:

$$p_{n+1} = \frac{p_n^2}{2} ((1 - b)^2 - (1 - a)^2) + p_n (b(1 - b) + (1 - a)) + \frac{b^2}{2} \quad (1)$$

$$p_{n+1} = p_n^3 (a - b)^2 (2 - a - b) + p_n^2 (a - b) (2 - 2a + 4b - ab - 3b^2) + 2p_n (1 + b) (1 - a) + p_n b^2 (2 - 3b + a) + b^3 \quad (2)$$

Eventually, for cases where $N = \infty$ describing the condition of infinite learning time, the model boils down to the following formula:

$$p_{n+1} = \frac{p_n (1 - a)}{(1 - b) + p_n (b - a)} \quad (3)$$

For $a = b$ it can thus be shown for all n that $p_{n+1} = p_n$, corresponding to an unchanging system. Other outcomes are yielded in cases where $a \neq b$: For $a < b$ implying that $(1 - a) > 0.5$, p_n converges on 1, as a consequence all learners end up with acquisition of L_1 . Inversely, if $a > b$ which entails that $(1 - b) > 0.5$, p_n tends to 0, with the result that only L_2 learners remain. As corollary of this model it can therefore be concluded that, provided that learning time is not confined, in a community either the two grammars stably co-exist, or one grammar takes over completely, in which case the other will eventually be eradicated entirely.

In an attempt to approximate attested “real-world” diachronic evolutions, Niyogi and Berwick (1997, 1998) proceed beyond the one-parameter model to simulations with increased numbers of (binary) parameters: These include 3-parameter models as well as 5-parameter simulations for the loss of the verb-second (V2-) grammar in the transition from Old to Modern French (Niyogi and Berwick, 1997). Interestingly, application of the 3-parameter model which accounts for occurrence or absence of V2 as well as specifier and head positions, V2 supersedes non-V2 populations which is against the historical facts; however, the model accomodating five parameters (*i.e.*, occurrence or presence of: null subject, V2, clitic nominative pronouns, assignment of nominative case by the inflected verb via

government, assignment of nominative case by the inflected verb via agreement, *cf.* Clark and Roberts, 1993: 322) correctly predicts the loss of V2. Additionally, the authors demonstrate that it is possible to precisely model the speed of transition from Old to Modern French, if, in accordance with the historically attested facts, a heterogeneous initial condition with a mixed population of V2 and non-V2 speakers is assumed. Beyond this analysis, Niyogi and Berwick (1998) demonstrate how their model can fit diachronic data on the shift of the direct object clitic position relative to the host verb from Classical to European Portuguese, contingent upon the choice of the algorithm: While the TLA yielded an incorrect outcome, application of the Maximum Likelihood estimator, under certain assumptions, made a valid prediction.

Nonetheless, as admitted by the authors, several idealizing over-simplifications render the presented system model simulations vulnerable to limitations: These include the assumption of non-overlapping generations whose linguistic make-up is shaped only by adult speakers who transmit linguistic input to the subsequent generation of language learners, in a one-way direction, with no linguistic exchange amongst the learners. Furthermore, it is assumed that all (first) language learners are exposed to example sentences from the same linguistic distribution (*i.e.*, lack of neighborhood effects), that members of the adult generation exhibit a stable, unchanging grammar use throughout life (corresponding to a monolingual maturation hypothesis), and that accordingly (first) language learners acquire only one grammar rather than more than one simultaneously within a limited time frame. Arguably the two latter points in particular, *i.e.* whether individuals either bear or lack multiple grammars, have a considerable impact on language change. Such possible drawbacks notwithstanding, authors' ideas represent valuable contributions towards bridging the gap between the individual language learner and population-based linguistic change.

2.2.3. A variational model of grammar competition (Yang, 2000)

Pinpointing the interplay of inborn linguistic capacities as epitomized by UG (universal grammar) and environmental input factors, Yang (2000) gave an account of a computational model linking language learning and change, which likewise borrows from concepts of evolutionary biology: Internal and external forces, *i.e.* genetic endowment and environmental circumstances, interact to determine an individual's phenotype, just as, correspondingly, UG and linguistic input alike, interactively mould the language eventually attained; this interaction over time moreover sets the frame for linguistic change. However, UG can only assume a finite number of states, thus the array of possible observable features of the inborn component is constrained. Relying on these postulates, Yang (2000) developed a

language acquisition model which underlies directly a further model of language change, the former maintaining that the individual language learner harbors multiple distinct grammars whose distribution is shaped by environmental input, thus favoring or punishing one or the other of such grammatical systems. In this regard, cross-linguistic variation reflects the distribution of all these grammars displayed by language learners, which varies due to continuous adaptation to environmental input evidence.

By this variational approach to language acquisition, the author formalizes the aforementioned concepts by employing as learning algorithm the linear reward-penalty scheme (Bush and Mosteller, 1958) unlike Niyogi and Berwick in their approach (1997, 1998) who utilize the triggering learning algorithm (TLA) or a maximum likelihood estimation. However, as remarked by Yang (2000: 233, 248), successful convergence to a target grammar is not always warranted by using the TLA, and discrete learning of parameters, elicited by environmental triggers according to the TLA, is contra-factual in view of gradual language development evidenced in first language learners. Thus in order to capture the gradualness of learning and competition of grammars during language acquisition, for one thing, each grammar G_i is assigned a probability p_i with which the learner has access to G_i . Furthermore, if a sentence s can be part of a linguistic environment E , it is a possible utterance in E , therefore $s \in E$. If in turn s can be analyzed (or, put differently, parsed) by a grammar G_i , with altogether N different grammars in the population, learning obeys a model where the language learner selects G_i with the probabilities p_i^t and p_i^{t+1} at times t and $t+1$, respectively. In case of successful analysis of a sentence s by G_i , reward ensues resulting in a higher value of p_i^{t+1} , cf. formula (4). In parallel, all other grammars that are not selected, e.g. G_j ($i \neq j$), receive punishment by lowering the value of p_j^{t+1} compared with p_j^t (5). Conversely, if G_i is unable to parse s , it receives punishment as formalized in (6), whereas the other grammars, e.g. G_j , are rewarded (7). Both reward and punishment depend linearly on a learning parameter γ :

$$p_i^{t+1} = p_i^t + \gamma (1 - p_i^t) \quad (4)$$

$$p_j^{t+1} = (1 - \gamma) p_j^t \quad (5)$$

$$p_i^{t+1} = (1 - \gamma) p_i^t \quad (6)$$

$$p_j^{t+1} = \frac{\gamma}{N-1} + (1 - \gamma) p_j^t \quad (7)$$

In addition, a penalty probability c_i is associated with a corresponding grammar G_i , which accounts for the “fitness” of G_i in a linguistic environment E : In detail, c_i describes the probability with which a sentence s uttered in E cannot be parsed by G_i , or stated differently, the proportion of sentences of E that G_i is incompatible with; c_i can therefore be regarded as fitness measure for its respective grammar in a given environment of input data, its numerical estimation is possible by thorough inspection of historical and contemporary text corpora. In the language acquisition model, it can be demonstrated that grammars with low penalty probabilities, *i.e.* grammars featuring high compatibility with input sentences from E , are prevailing in the population over those grammars with high penalty probabilities and that the learner converges to a stable distribution pattern of grammars which are a function of c_i : For the simplified case of solely two grammars G_1 and G_2 , under the assumption of infinite time for learning, the probabilities p_1 and p_2 for G_1 and G_2 to be selected, respectively, approach the following values:

$$p_1 = \frac{c_2}{c_1 + c_2} \quad p_2 = \frac{c_1}{c_1 + c_2} \quad (8)$$

Accordingly, the model accounts for variation and the gradual gain of predominance of the target grammar, both evidenced in first language acquisition, caused by competition of numerous (but, following the rules of UG, finite and discrete) grammatical systems. Importantly, since in the course of learning a certain distribution of more than merely one grammar is attained, one linguistic expression s can as a consequence be parsed by more than one grammar, a condition which creates susceptibility to the forces of language change at the population level. This situation can be further translated into a formalized model of acquisition-based diachronic change, whereby at this point the author underscores that for initial induction of linguistic change, the model relies upon factors other than the process of language learning *per se*: Migration, linguistic innovation, and socio-cultural factors may all alter the linguistic evidence for the subsequent generation of language learners, while the relevance of individual mislearning is rejected (on grounds of evidence for disparate stabilities of the same syntactic phenomena in different languages), provided that all learners are exposed to a homogeneous linguistic environment E .

Under the assumption of a linguistic environment consisting of two grammars G_1 and G_2 , *i.e.* a two grammar system, penalty probabilities c_1 and c_2 (defined as shares of utterances incompatible with their corresponding grammars G_1 and G_2 in E) can be re-expressed as proportions α and β of sentences of G_1 and G_2 , respectively, that are incompatible with each

other, multiplied by a corresponding proportion value (or weight) p or q for G_1 and G_2 , respectively, describing the share of a grammar of all utterances in E :

$$c_1 = \beta q \qquad c_2 = \alpha p \qquad p + q = 1 \qquad (9)$$

The proportion metrics of mutual incompatibility α and β are synonymously termed advantage of G_1 and G_2 , respectively. Beyond that, since penalty probabilities αp and βq are outcome predictors of language acquisition, the ratio of p over q can be determined for each subsequent generation $n + 1$:

$$\frac{p_{n+1}}{q_{n+1}} = \frac{\frac{\alpha p}{(\alpha p + \beta q)}}{\frac{\beta q}{(\alpha p + \beta q)}} = \frac{\alpha p}{\beta q} = \frac{c_2}{c_1} \qquad (10)$$

It therefore turns out that either the weight p of G_1 or that of G_2 , which is q , constantly increases from generation to generation at the expense of the other grammar, resulting in a time-dependent ever-increasing or ever-decreasing ratio of p over q . Concomitantly, the mutually exclusive proportions of each grammar α and β , which, according to Yang (2000), are presumably constants that, unlike penalty probabilities, cannot be derived from historical corpora, determine the result of grammar competition: In case that $\alpha > \beta$, G_1 eventually replaces G_2 , whereas if $\beta > \alpha$, G_2 wins over G_1 . Conclusively therefore, the model predicts that “once a grammar is on the rise, it is unstoppable“ (Yang, 2000: 239). In summary, the work demonstrates that a variational model of language acquisition can be translated into a variational model of language change, postulating that heterogeneity in the linguistic environment, caused by factors external to language learning, instigates competition of grammars internally represented to the individual learners by UG, a process which overtly manifests itself in different distributions of grammars in every subsequent generation, leading the way to language change.

2.2.4. Models of iterated learning

Iterated learning models have been applied, besides in models of cultural transmission, also in simulations of language evolution and emergence of linguistic structure using artificial language learning paradigms (*cf.* Kirby and Hurford, 2002; Kirby *et al.*, 2004; Griffiths and Kalish, 2007; Kalish *et al.*, 2007; Real and Griffiths, 2009, 2010). Such models use Bayes's

law to simulate learning processes over generations of learners from an initial input state, implying that intergenerational changes depend upon the learners' inductive biases to solve problems such as generating hypotheses with respect to grammar from a restricted set of examples. It turns out that over generations knowledge is not transmitted flawlessly, an outcome which is due to Bayesian learners who infer a target grammar from a set of hypotheses guided by their biases. Each hypothesis h is assigned a probability $p(h)$ by the learner or Bayesian agent before being exposed to data, thus a prior probability distribution describes the whole set of hypotheses, reflecting all relevant *a priori* biases of a learner. Faced with input data d , a learner in turn assigns to each hypothesis a posterior probability in the context of this data, *i.e.* $p(h | d)$. Given furthermore the probability $p(d | h)$ indicating how likely data is produced on part of a speaker, who used to be a previous generation's learner, adopting hypothesis h , a relation between all probabilities can be established as follows:

$$p(h|d) = \frac{p(d|h) p(h)}{\sum_h p(d|h) p(h)} \quad (11)$$

Hence the probability a learner assigns to a hypothesis h in the context of data d is a function of the probability of data generated by the previous generation under hypothesis h and the learning bias, stratified by the average over all hypotheses. In an iterated Bayesian model simulating transmission of knowledge over generations (*cf.* Kalish *et al.*, 2007), the probability that a learner of the n^{th} generation selects hypothesis i if the previous learner of generation $n-1$ chose hypothesis j , which is $p(h_n = i | h_{n-1} = j)$, is dependent upon all probabilities assigned to hypotheses on part of the learner as given in (11) as well as the likelihood of data produced by the previous generation, expressed in the following equation:

$$p(h_n = i | h_{n-1} = j) = \sum_d p(h | d) p(d | h_{n-1} = j) \quad (12)$$

Accordingly, the hypothesis selected by a learner is dependent on that chosen by the previous learner only. A Markov chain model utilizing the transition matrix defined by equation (12) predicts the convergence on the distribution of prior learning biases $p(h)$, hence it is only the process of biased learning that determines the result of iterated learning. As iterated learning models of linguistic evolution have demonstrated, a random language

equipped with context-free grammar at the outset of the simulation is utterly unstable and, within a few generations, undergoes change towards a stable system exhibiting syntactic structure (Kirby *et al.*, 2004). Kalish *et al.* (2007) furthermore showed that not only Bayesian agents in simulations but also human learners in an experimental study stably converged to the probability distribution $p(h)$, a process reflecting inductive learning biases. From this, it can be hypothesized that intergenerational transmission of knowledge by iterated learning could shed light on the structure of human mind, and moreover that all cultural transmission processes are reflections of underlying cognitive biases that cultural manifestations such as religions and social norms as well as languages are adapted to (*cf.*, Griffiths and Kalish, 2007; Kalish *et al.*, 2007). The process of iterated learning therefore serves as a possible explanation for the emergence of structure in linguistic systems.

2.2.5. Evaluation of acquisition-based computational approaches to diachronic change

Taken together, applying a comparative view on the presented models of diachronic change through language acquisition, one striking common characteristic is the analogy with concepts of evolutionary biology, though only Clark and Roberts (1993) take advantage of a genetic algorithm, whereas Niyogi and Berwick (1995, 1997, 1998) and Yang (2000) alike employ algorithms that are usually applied in psychology. As for models of iterated learning, though pointed out to share analogies with evolutionary biology insofar as universal grammar embodied in generalized inductive learning biases reflects the environment for a language to adapt to by a process of selection, this process is mechanistically distinct from biological systems: While the phenotype is in general considered not to be able to change the genotype directly in biology, languages are altered on part of the learners through a process of inferring grammatical competence, *i.e.* the “language genotype”, from utterances performed by the speakers, *i.e.* the “language phenotype” (Kirby *et al.*, 2004). However, the algorithm applied in iterated learning is fully equivalent with the Wright-Fisher model describing genetic drift upon mutation in population biology (Real and Griffiths, 2010). Though as yet utilized for simulations of language evolution with artificial “corpora”, mechanisms of iterated learning bear strong relevance for principles of diachrony.

While the approach of Clark and Roberts (1993) is devoid of predictions concerning the trajectories of grammar evolution over time, Yang’s account (2000), in line with precedent work on diachronic change (*e.g.*, Bailey, 1973; Kroch, 1989), observes a sigmoid (logistic) growth pattern, and both sigmoid (logistic) and exponential curve shapes are yielded in the framework of Niyogi and Berwick (1995, 1997, 1998). Both trajectory patterns imply a

short time course of rapid changes followed by a long-lasting period characterized by the gradual elimination of residual forms. Similarly, models of iterated learning, displaying rapid emergence of linguistic structure that converges to a stable state, are claimed to follow a sigmoid trajectory of frequencies over time, like in models of genetic drift (Real and Griffiths, 2010). Importantly, while Clark and Roberts (1993) use no more than a single grammar to model the attainment of the final outcome of language acquisition, and according to Niyogi and Berwick (1995, 1997, 1998) a proportion of learners reach one single grammar whereas the remainder converges on another, Yang's approach (2000) accounts for the possibility that a learner converges to multiple grammars.⁹ The probabilistic premises of iterated learning models concerning a learner's hypothesis and the learning outcome distribution is in principle compatible with both views of convergence to single or multiple grammars, even though in the end a certain grammar, contingent upon the data at the outset, is predicted to be most stable, *i.e.* most likely.

Worthwhile enough, individual mismatches on the way to convergence on the target grammar are conceived as a viable source of linguistic change in the frameworks proposed by Clark and Roberts (1993) as well as Niyogi and Berwick (1995, 1997, 1998), however, a scenario explicitly ruled out by Yang (2000) who not only denies an effect of individual mislearning events on the population level and therefore diachronic variation, neither does he regard the process of language acquisition as causative for change: This role in the model is attributed to factors outside grammar acquisition, encompassing migration, linguistic innovation owing to language contact, and socio-cultural circumstances. Further discussion on this issue is offered in section 6.3.

⁹ This latter assumption is compatible with a report by Kroch (1989) who contends that transition periods during language development represent grammar competition, entailing that learners of such periods attain more than one grammar, a condition designated as "syntactic diglossia": In order to back up this claim, the author therein adduces as examples, amongst others, the rise of the periphrastic auxiliary "do" during late Middle and early Modern English just as the phenomenon of intra-sentential code-switching (*e.g.*, MacSwan, 2000), both cases indicative of evidence that language learners acquire proficiency in more than one grammatical system.

3. From language universals to conspiracies

3.1. Concepts of language universals and linguistic typology

Diverse concepts and approaches have emerged in the history of linguistics that tackle the issue of language universals, frequently resulting in abstract but also implicational cross-linguistic generalizations (for review, *cf.* Fitch, 2011). However, given the bewildering diversity of linguistic systems representing possible human grammars, the exploration of common design features of language has not always been a central paradigm in linguistic and cognitive research as it is to-date. In fact, reservations to the idea of linguistic universality are articulated still nowadays, for instance by Evans and Levinson (2009) who argue that exceptions can be found to many linguistic universals held absolute (exceptionless), that only a minority of languages is thoroughly explored, and that universals might not reflect independent features due to relatedness and contact among examined languages. Admittedly, throughout a long developmental history that can be ascribed to most languages, diversification takes place resulting in the huge variation found in human communication systems, yet linguistic diversity is just one side of the coin, the formulation of linguistic universals, be they absolute (exceptionless) or statistical (representing tendencies), is the other. More recently, common features of unrelated creole languages are often adduced as evidence for the existence of such universals, since virgin human languages *in statu nascendi* are believed to unveil such universal traits most clearly, as will be an important topic of section 4. Historically, in the 18th and 19th centuries, the terms “universal grammar“ and “philosophical grammar“ used to denote features of a given language so obvious that there was no need for them to be mentioned in respective grammars and dictionaries of that language, including statements like “Utterances express meanings“ and “Languages contain words“, *i.e.* properties deemed common to all languages (*cf.* Fitch, 2011). Remarkably, in this era, Charles Darwin (1871) maintained that all human beings are born with an instinctual desire to learn language and the neural equipment to do so, ideas which roughly anticipated modern concepts stating interplay of innate UG and language learning, relying upon the neuronal substrate and environmental input alike.

By the first half of the 20th century, the school of relativism was at its zenith in cultural and social anthropology as well as linguistics, above all in the United States of America, entailing lowered interest in the search for universals, since the uniqueness of every culture and language under examination was stressed, as exemplified by works of Franz Boas, Alfred L. Kroeber, Edward Sapir, and Benjamin L. Whorf (Stocking, 1974; Kroeber, 1939; Sapir, 1949; Whorf, 1956). Historically, the rise of relativism, which in anthropology occurred *via*

diffusionist ideas advocating the importance of regional research in distinct “culture areas” (e.g., Graebner, 1911; Kroeber, 1939), can be understood as counter-movement against evolutionism. This latter movement had classified societies and languages as “less developed” vs. “highly developed”, with a teleologically directed evolution from the former to the latter, *i.e.* towards westernized societies and inflectional Indo-European languages (e.g., Morgan, 1877; Finck, 1899). Yet another highly influential contemporary school, emerging in the early 20th century, was structuralism whose founder F. de Saussure (*cf.* de Saussure, 1916) placed focus on synchronic investigation stressing form over content contrary to the preceding century’s era of predominant diachronic research. Structuralism initially was concerned with description of languages and cultures rather than cross-cultural and cross-linguistic comparisons, moreover, Leonard Bloomfield, the most prominent proponent of linguistic structuralism in the United States, adopted a behaviorist approach (e.g., Bloomfield, 1933), an attitude which for linguistics was jettisoned by Noam Chomsky (1959) by his notion of an innate grammar learning ability. Henceforth elevated scientific focus was directed to the exploration of universal tendencies and universal properties common to all human languages, in conjunction with investigations into possible underlying inborn cognitive abilities. It was Chomsky who in the following years and decades placed the creative, productive aspect of language at center stage, re-defining the old term „Universal Grammar“ as innate, biological endowment to learn the complete grammar of a language (e.g., Chomsky, 1965): He thereby distinguished substantive universals that relate to the substance of a language (*i.e.*, nouns and verbs as constituents of every language, furthermore expressions designating persons, items, feelings, and behaviors) from formal universals relating to grammatical rules and regularities including their interactions (e.g., binary constituent structure of phrases and sentences, and syntactic rules operating in this framework). Thereby, according to Chomsky, the notion of formal universals need not be restricted to syntax, in fact not even to language, as exemplified by the statement “color designations must divide the color spectrum into continuous segments“ which is classified as a formal semantic universal. Importantly, the notion of Universal Grammar (UG) does not supply a list of universal features of languages spoken by adults, it rather describes the toolkit and the mechanisms that are available to first language learners, as clarified by Jackendoff (2002: 75). This implies that not every such mechanism available by UG must necessarily be at work in every language, or in other terms, not every tool needs to be used from the provided toolkit.¹⁰ The Principles and Parameters (P&P) theory

¹⁰ “When you have a toolkit, you are not obliged to use every tool for every job. Thus we might expect that not every grammatical mechanism provided by Universal Grammar appears in every language. For instance, some languages make heavy use of case marking, and others don’t; some languages make heavy use of fixed word

as introduced briefly in section 2.1. is one important approach to the toolkit principle (Jackendoff, 2002: 75).

A different approach to the identification of language universals was adopted by Charles Hockett who, as a result of comparative work on animal communication systems, described “design features of human language” as distinctive hallmarks relative to other biological species (Hockett, 1960, 1966). Originally, Hockett reported on thirteen design features (1960), amending three additional ones to yield sixteen altogether (1966). In detail, Hockett claimed ten of those features as necessary and sufficient to define any communication system as language, comprising productivity or openness, displacement, duality of patterning, arbitrariness, discreteness, interchangeability, total feedback, specialization, rapid fading or transitoriness, and broadcast transmission and directional reception.¹¹ Furthermore, the remainder of the originally published set of design features include the existence of a vocal-auditory channel through which acoustic signals are transmitted and received, semanticity implying the stipulation of a relation between the acoustic signal and its denotation, and traditional transmission describing the non-existence of pure genetically coded innateness of any linguistic system thus entailing the necessity of teaching and learning in the transmission from generation to generation. Further additions by Hockett in 1966 include the possibility to express wrong and meaningless utterances like lies as denoted as prevarication, the possibility of meta-linguistic utterances as termed reflexiveness, and the possibility to learn a language other than one’s own first language as described as learnability. All of these characteristics are deemed valid by and large still to-date, in particular as for spoken language, with possibly some modifications appropriate (Fitch, 2011). Current doctrine maintains that the combination of all these features is found in no other species outside *Homo sapiens*.

Based on these design features of any communicative system, Hockett (1966) proceeded to deduce universals of (human) language grouped into three categories, namely generalizations, grammatical universals, and phonological universals. Contained in Hockett’s

order, and others don’t. We would like to say that Universal Grammar makes both these possibilities available to the child; but only the possibilities actually present in the environment come to realization in the child’s developing grammar.” (Jackendoff, 2002: 75)

¹¹ *productivity* or *openness*: the potentially limitless production of different utterances by a finite set of linguistic signs; *displacement*: the possibility to refer to spatially and temporally remote meanings; *duality of patterning*: the dichotomy of a phonological and semantic system; *arbitrariness*: the property that confers unlimited possibilities of denotation, also with respect to abstract meanings; *discreteness*: the property of linguistic systems to be composed of discrete rather than continuous elements; *interchangeability*: the property of language signals to be both received and transmitted; *total feedback*: the property of language signals to be both transmitted and received at the same time by the same speaker; *specialization*: the exclusiveness of language signals to subserve only communication tasks; *rapid fading* or *transitoriness*: the temporariness of emitted language signals; *broadcast transmission* and *directional reception*: the non-directional property of signal transmission and directedness of signal reception

universal generalizations are statements as evident as “Every human community has a language” and “Every human language has the vocal-auditory channel, tradition, and learnability”, aside from rather elusive claims about generally higher degrees of variation in phonetics and phonology as compared with syntax and at small size-levels than at large, at least in the syntactic domain (the latter statement refers to large-scale patterns of syntax common to all languages, with diversity displayed by the details). Comprised in grammatical universals are the ubiquitous occurrence of proper names, of markers devoid of denotation (like “and” and “or”), and of expressions for 1st and 2nd persons singular in particular and deictic elements in general, just like a clause type constituted by “topic” and “comment” parts and predicators with one-referent and two-referent structures. Lastly, redundancy and symmetry in the phonological system, and the universality of sound change, the latter an inherent feature of diachronic development, figure, amongst others, as phonological universals.

In line with the growing interest in common grounds for all languages which emerged at the beginning of the second half of the 20th century, it was Greenberg (1966) in the 1960s by his notion of implicational universals to exert a substantial impact on the research areas of language universals and linguistic typology alike, foremost as much as morphology and syntax is concerned, as will be discussed extensively in the following section. Worthwhile enough, the formulation of phonological universal implications is ascribed to Jakobson (1974: 51ff., in a translation of his original monograph from 1941) who referred to the conjunction of two elements, such as stops in any phonetic system as prerequisites for fricatives, as “solidarity”. Noteworthy concepts of language universals after Greenberg (1966) include that presented by Eugenio Coseriu (1974) and, equally embracing typology, the approach put forth by Seiler (1995). As for the former, Coseriu (1974) drew a distinction between three primary types of linguistic universals, *i.e.* possible, essential, and empirical universals, and two secondary types which derive from the primary ones. Beyond those, alternative categories are suggested including classification according to formal criteria, degree of generality, activity, semiotic level, and formulation. Specifically, possible universals comprise all conceivable linguistic categories, be they attested or even unattested, that are not at odds with the concept of language. Furthermore, essential universals designate imperative properties of language pertaining to linguistic competence and capacity. And third, empirical universals contain only attested properties of language, be they absolute, *i.e.* found in all languages, or relative, *i.e.* attested in the bulk of languages but not in all. Concerning the secondary types, selective universals, by derivation from possible and empirical universals, describe the fact that every

language displays a restricted array of all possible processes or elements. Lastly, implicational universals, combining possible and essential universals, reflect Greenberg's notion of the same topic. Coseriu concludes that all languages, their considerable diversity notwithstanding, are constructed according to the same principles, rendering them identical or at least analogous in this sense. Though languages differ in their semantic and material organization, differences do not concern the same level as principal analogies.

Typology, a linguistic subdiscipline that received immense boost by Greenberg's study (1966), aims at classifying languages according to structural similarities, irrespective of their genealogic relatedness. Traditional gross typological classification of languages applies several criteria according to which languages are categorized, encompassing most importantly morphological, syntactic, and morpho-syntactic criteria (for overview, *cf.* Song, 2001). By morphological criteria, languages can be classified as isolating, agglutinative, inflecting, or polysynthetic, whereas syntactic criteria account for the position of verb, subject, and object in relation to one another giving rise to the notions of SVO, SOV, VSO, V2 *et c.* word order patterns that languages typically (or predominantly) follow. Beside these, morpho-syntactic classification categorizes languages along the expression of syntactic concepts on the morphological surface: Following five logical combinations of how semantic roles of agent (A), patient (P), and the argument of an intransitive clause (S) are treated morphologically, a distinction between case marking systems ensues, *i.e.* nominative-accusative (AS/P; patient marked differently from A and S), ergative-absolutive (A/SP; agent marked differently from S and P), AP/S, tripartite (A/S/P; no overlap of case markers), and neutral (ASP; no overt case marking distinction). Since typologists search for traits that delimitate languages from one another so to define distinct language types, an enterprise that relies upon diversity, a potential contradiction exists in regard to the goals of language universalists. However, the notion of implications in universal statements brought about the aspect of variation, since typology uses combinations of linguistic features to characterize language types, with usually more than one possibility how two or more elements combine as parts of an implication (Comrie, 1981: 30; Hawkins, 1983: 50-51). Which of those combinations are possible or more likely than others is a task of comparative linguistic work, sometimes comprising analyses on corpus data from large numbers of languages from different families. Classical textbooks that refer to typology in conjunction with linguistic universality include those by Comrie (1981) and Mallinson and Blake (1981), the latter covering exclusively syntactic aspects and advocating enhanced efforts for cross-linguistic investigations.

One notable approach to combine linguistic universality along with typologically relevant implications within a cognitive framework is the “Language Universals Research and Language Typology” (UNITYP) program initiated and coordinated by Hansjakob Seiler (*cf.* Seiler, 1995). Research activities within this program are guided by the mindset that universality is not found in the substance of language, *i.e.* certain elements, features or constructions, but rather in processes and operations, an attitude which is rooted in the Humboldtian notion of language as activity (*enérgeia*, *Thätigkeit* [sic]) rather than as a product (*érgon*, *Werk*) (Humboldt, 1836: 41; Seiler, 1995: 299-301). These activities have lead to the identification of several linguistic dimensions¹² that are each reflected by a corresponding cognitive-conceptual domain, which can be subdivided into subdimensions and subdomains which concrete linguistic structures with their elementary units termed “parameters” (in the subdimensions) or “properties” (correspondingly, in the subdomains) are assigned to. Importantly, in this framework, it is possible for one particular grammatical construction to be plurifunctional, *i.e.* pertain to more than one (sub)domain. Whereas on the one side, dimensions, subdimensions and parameters are conceived of as empirical universals of language *sensu* Coseriu (1974), Seiler furthermore proposes three operational/functional principles according to which (sub)dimensions and parameters are organized, representing essential universals as described by Coseriu (1974): The first one, indicativity, which is associated with discourse-relatedness and metalinguistic reference, is in a converse and complementary relationship with the second more marked principle, *i.e.* predicativity that embodies the linguistic explicitness of a concept as expressed, for example, by strong use of morpho-syntactic marking. While both indicativity and predicativity are mandatory universals, the third one, iconicity, does not obligatorily occur in every language. In brief, the notion of iconicity conveys the concept of image representation and intuitive straightforwardness but also confers ambiguity or even indeterminacy. Within the overarching dimensional framework of the UNITYP program, contributions to language typology have been and are published on a regular basis.

In living up to the need for more cross-linguistic investigations as articulated by Mallinson and Blake (1981), extensive comparative typological studies from the previous decades comprising very large cross-linguistic samples, aside from Greenberg’s classical report (1966), have been conducted. These include Hawkins’s study on word order universals (1983) and Dryer’s notion of phenomena correlated with relative order of verb and object (1992). More recently, Haspelmath (1997) published a study on indefinite pronouns in a core

¹² including apprehension, concomitance, determination, localization, nomination, numeration, participation, possession, and situation (Seiler, 1995: 275ff.)

sample of detailed data from 40 languages and in an expanded dataset of limited information from an additional 100 languages. An extensive survey by Haspelmath *et al.* from 2005 in more than 2500 languages that incorporated grammatical as well as lexical features was ever since expanded to an exhaustive online database (Dryer and Haspelmath, 2011). Relevant findings of some of these large-scale typological studies will, embarking on Greenberg's report (1966), recur in the following sections.

3.2. Syntax-related implicational universals of language: Greenberg and the aftermath

In the quest for universals that characterize all human languages, Joseph H. Greenberg's contribution to the Dobbs Ferry conference in 1961, elaborated into a congress report (Greenberg, 1966, 2nd edition; 1st edition: 1963), undoubtedly marked a significant cornerstone:¹³ From corpora of thirty languages spread over five continents, representing a variety of unrelated linguistic families,¹⁴ Greenberg extracted 45 universals that he classified according to basic (word) order typology, syntax, and morphology. Aside from general ("unrestricted") statements obeying the scheme "Every language has *x*", Greenberg also formulated a large proportion of absolute and statistical implications, following the patterns "If in a particular language there is *x*, then there is also *y*" (with the converse statement, *i.e.* "Given *y*, then *x* is always found", usually not holding true) and "If in a language there is *x*, then *y* is more probable than *z*", respectively. Remarkably, Greenberg was the first author to explicitly formulate such implicational universals (Haspelmath, 2008), at least with respect to syntax and morphology, with a tremendous influence on subsequent work in language universals and typology. However, without doing injustice to this pioneering work, at this point to indicate limitations of the study is in order that are pointed out by the author himself: The selection of languages covering large areas of the world and diverse linguistic families notwithstanding, Greenberg admits a bias by choosing only such ones he was familiar with. Moreover, given the restricted number of sample languages Greenberg drew his conclusions from, universal statements are to be considered no more than tentative and suggestive. In this regard, it must be considered that implicational conclusions were sometimes drawn only from a subset of the whole sample of thirty languages, for example when a certain word order type

¹³ As Osgood (1966: 299 f.) points out, the Dobbs Ferry conference of 1961 constituted a major turning point, in fact a revolution, in modern linguistics and psychology alike that henceforth directed enhanced focus on generality and structural analogies.

¹⁴ Yorubá, Nubian, Swahili, Fulani, Masai, Songhai, Berber (Africa); Maya, Zapotec, Quechua, Chibcha, Guaraní (Amerindian); Turkish, Hebrew, Burushaski, Hindi, Kannada, Japanese, Thai, Burmese, Malay (Asia); Basque, Serbian, Norwegian, Modern Greek, Italian, Welsh, Finnish (Europe); Maori, Loritja (Oceania).

only was analyzed. For illustration, listed below are the seven word order-related (1. through 7.) and eighteen syntax-related (8. through 25.) Greenbergian universals (Greenberg, 1966):

1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.
2. In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.
3. Languages with dominant VSO order are always prepositional.
4. With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.
5. If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.
6. All languages with dominant VSO order have SVO as an alternative or as the only alternative basic order.
7. If in a language with dominant SOV order, there is no alternative basic order, or only OSV as the alternative, then all adverbial modifiers of the verb likewise precede the verb.
8. When a yes-no question is differentiated from the corresponding assertion by an intonational pattern, the distinctive intonational features of each of these patterns are reckoned from the end of the sentence rather than from the beginning.
9. With well more than chance frequency, when question particles or affixes are specified in position by reference to the sentence as a whole, if initial, such elements are found in prepositional languages, and, if final, in postpositional.
10. Question particles or affixes, when specified in position by reference to a particular word in the sentence, almost always follow that word. Such particles do not occur in languages with dominant order VSO.
11. Inversion of statement order so that verb precedes subject occurs only in languages where the question word or phrase is normally initial. This same inversion occurs in yes-no questions only if it also occurs in interrogative word questions.
12. If a language has dominant order VSO in declarative sentences, it always puts interrogative words or phrases first in interrogative word questions; if it has dominant order SOV in declarative sentences, there is never such an invariant rule.
13. If the nominal object always precedes the verb, then verb forms subordinate to the main verb also precede it.
14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.
15. In expressions of volition and purpose, a subordinate verbal form always follows the main verb as the normal order except in those languages in which the nominal object always precedes the verb.
16. In languages with dominant order VSO, an inflected auxiliary always precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb.
17. With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.
18. When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.
19. When the general rule is that the descriptive adjective follows, there may be a minority of adjectives which usually precede, but when the general rule is that descriptive adjectives precede, there are no exceptions.

20. When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.
21. If some or all adverbs follow the adjective they modify, then the language is one in which the qualifying adjective follows the noun and the verb precedes its nominal object as the dominant order.
22. If in comparisons of superiority the only order, or one of the alternative orders, is standard-marker-adjective, then the language is postpositional. With overwhelmingly more than chance frequency if the only order is adjective-marker-standard, the language is prepositional.
23. If in apposition the proper noun usually precedes the common noun, then the language is one in which the governing noun precedes its dependent genitive. With much better than chance frequency, if the common noun usually precedes the proper noun, the dependent genitive precedes its governing noun.
24. If the relative expression precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun or both.
25. If the pronominal object follows the verb, so does the nominal object.

Likewise appearing in a conference proceedings edition, Coseriu (1974) presented his view, a rather conceptual one and by no means based on empirical data like Greenberg's report, on the issue of language universals. Amongst other types mentioned in the preceding section, he therein described implicational universals of language (implications) that can be either theoretical (deducible by conceptual analysis) or empirical as found in reality, thus taking up Greenberg's concept of implications. These are, according to Coseriu, subject to further differentiation along formal criteria, since they can assume positive ("if *x*, then also *y*") or negative ("if *x*, then not *y*"), as well as, drawing on a discrimination introduced by Husserl (1913: 264-265), unilateral ("*x* implies *y*, but not vice versa"), or bilateral (or reciprocal; "*x* implies *y*, and vice versa") properties.¹⁵ By a similar token, Aitchison (1996) proposed a comprehensive classification system of language universals along the dimensions of firmness and independence, straightforwardly in relation to the way suggested by Greenberg (1966): Whereas on the firmness axis, absolute and statistical statements can be distinguished according to the scheme "Languages always have *x*" and "Languages usually have *x*", respectively, the independence dimension sorts unrestricted ("*x* occurs everywhere") from implicational ("if *x*, then *y*") universals. So a unifying synthesis of Greenberg's (1966), Coseriu's (1974), and Aitchison's (1996) notions applied to Greenberg's first 25 universals of

¹⁵ cf. Coseriu (1974: 54): "Les implications peuvent être *théoriques* (dédites au moyen de l'analyse conceptuelle des possibilités considérées) ou bien *empiriques* (constatées). Ainsi, p.ex., l'implication: « Le terme neutre d'une opposition sémantique binaire a deux signifiés de langue » est une implication théorique, tandis que: « Si dans une langue il y a flexion, il y a aussi dérivation » est une implication empirique. Du point de vue de leur forme, les implications peuvent être *unilatérales* (*x* implique *y* [mais *y* n'implique pas *x*]), ou bien *bilatérales* ou *réciroques* (*x* implique *y* et *y* implique *x*); *positives* (si *x*, alors *y*), ou bien *négatives* (si *x*, alors non *y*)."

language yields the following terminological distinction:¹⁶ 1.) Absolute positive implications are by far the most abundant statements, embracing as many as seventeen, *i.e.* universals 3, 5, 6, 7, 8, 11, 12, 13, 15, 16, 19, 20, 21, the first part of 22 and 23 each, 24, and 25. Not surprisingly, in view of the sample size confined to thirty languages, other authors denied the absolute nature of these implicational statements, considering them “merely” statistical universals, as for instance Vennemann (1974a,b). 2.) Statistical positive implications are represented by eight Greenbergian universals, namely 2, 4, 9, the first part of 10, 17, 18, and the second part of 22 and 23 each. 3.) Absolute negative implications include only (the second part of) universal 10, and 4.) universals 1 and 14 are not implications but rather general statements according to the pattern “In a surrounding *S*, every language has *x*”, corresponding to non-implicational unrestricted universals within a defined context. Taken together, the overwhelming majority of the Greenbergian syntax-related universals are implications. In this vein, Hawkins (1983: 19) emphasizes the significance of Greenberg’s findings insofar as he fundamentally established the validity and importance of implicational universals.

To some of Greenberg’s universal statements, criticism was raised, and proposals for emendations by modification and amendment were voiced partly based on larger sample sets in typological studies following up his work. One such proposal was based upon the observation that Greenberg’s implications do not contain statements about SVO word order (*cf.* Hawkins, 1983: 29-31), save as alternative to VSO stated in universal 6, rendering the feature of SVO devoid of correlation with other syntactic properties in Greenberg’s sample: Drawing on ideas articulated by Lehmann (1973), Vennemann (1974a,b) thus conflated the categories of VSO and SVO languages to yield one single group of VO word order languages with concomitant re-labeling of SOV as OV languages. Beyond that, he claimed a distinction between operand elements like verb *V* and operator elements like object *O* according to the syntactic category that is obtained after merging the two elements: Since both constituent structures [OV] and [VO] represent the same syntactic category as *V*, *V* hence corresponds to the operand and *O* to the operator (*cf.* Vennemann, 1974b: 347). Likewise, Vennemann (1974a,b) applied this principle to other constituent pairs in which one element is assigned the role of the operand and the other that of the operator, including (operand followed by operator) adverbial and verb, adjective and noun, main verb and auxiliary, and genitive and noun. Depending on the sequence in which the elements in each pair appear, *i.e.* how they are serialized, being the essence of Vennemann’s Natural Serialization Principle (NSP),

¹⁶ Greenberg (1966) remarks that reciprocity is usually ruled out. According to the formulation *per se*, whether either unilaterality or reciprocity would apply to a universal is not always clear.

languages can be classified typologically, thus giving rise to Vennemann's re-formulation of Greenberg's (1966) statements. Notably, these re-stated formulations differ from the original statements in two important aspects: Whereas, like in Greenberg's account, implications abound, re-formulated statements are all statistical, whereas the bulk of Greenberg's universals are claimed to be absolute in nature, and implications are maintained as bilateral-reciprocal in stark contrast to Greenberg's notion of explicit uni-laterality (see above).

In conducting a re-evaluation of Greenberg's original sample comprising thirty thoroughly investigated languages plus an appended 142 languages for monitoring distribution patterns of basic word order types, enlarged by an expanded database of some 350 more languages from his own research, Hawkins (1983) extracted twenty-three implicational word order universals that are listed in Table 1. Stated implications were all demonstrated to be unilateral and in many instances exceptionless for the investigated sample. As an example, implicational universal III reads "If a prepositional language has any verb position other than SVO, then if the adjective follows the noun, the genitive follows the noun likewise." Moreover, by inference from other universals, Hawkins identified two comprehensive exceptionless rules applying to prepositional and postpositional languages, *i.e.* the Prepositional Noun Modifier Hierarchy (PrNMH) as stated in universal XIV and the Postpositional Noun Modifier Hierarchy (PoNMH) as expressed by universal XVIII, respectively, which describe the positional stableness of serialized phrases in adposition to noun modifiers like demonstrative, numeral, adjective, genitive, and relative clause. Despite the occurrence of variant orders that do depart from the predicted subtypes, Hawkins argues that these variants can all be attributed to one of the basic subtypes, insofar rendering both the PrNMH and the PoNMH exceptionless.

Next to the identification of these implicational universals of word order, Hawkins (1983: 117-120) proposed a revision of Greenberg's universal 20, which originally stated that the serialization of modifiers preceding the noun is implemented in the order demonstrative, numeral, and adjective, and in the same or the exact opposite order if they follow the noun (see above): In the face of data evidencing postnominal orders not predicted by Greenberg's universal 20, Hawkins concluded that solely the notion on prenominal order of modifiers holds whereas reliable prediction of postnominal word order is not possible for attested languages, therefore proposing a revised version of this universal.¹⁷

¹⁷ "When any or all of the modifiers (demonstrative, numeral, and descriptive adjective) precede the noun, they (*i.e.*, those that do precede) are always found in that order. For those that follow, no predictions are made, though the most frequent order is the mirror-image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow." (Hawkins, 1983:119–120)

Table 1. Summary of Hawkins's twenty-three implicational word order universals (1983)

universal	absolute version	statistical version	remarks
I	SOV \supset (A-N \supset G-N) OV \supset (A-N \supset G-N)		
II	VSO \supset (N-A \supset N-G) V1 \supset (N-A \supset N-G)		
III	Prep & -SVO \supset (N-A \supset N-G)	Prep \supset (N-A \supset N-G)	
IV	Postp \supset (A-N \supset G-N)		
V	N-Dem \supset N-A		
VI	N-Num \supset N-A		
VII	Prep & -SVO \supset (N-Dem \supset N-G)	Prep \supset (N-Dem \supset N-G)	derived from universals III & V
VIII	Prep & -SVO \supset (N-Num \supset N-G)	Prep \supset (N-Num \supset N-G)	derived from universals III & VI
IX	N-G \supset N-Rel		
X	Prep \supset (N-A \supset N-Rel)		derived from universals III & IX
XI	Prep \supset (N-Dem \supset N-Rel)	N-Dem \supset N-Rel	derived from universals III & V
XII	Prep \supset (N-Num \supset N-Rel)	N-Num \supset N-Rel	derived from universals III & VI
XIII	Prep \supset (-SOV \supset N-Rel)		
XIV	Prep \supset ((N-Dem \vee N-Num \vee N-Poss \supset N-A) & (N-A \supset N-G) & (N-G \supset N-Rel))		PrNMH ^a derived from universals III, V, VI, IX, XXI
XV	Postp \supset (Dem-N \supset G-N)		
XVI	Postp \supset (Num-N \supset G-N)		
XVII	Postp \supset ((A-N \vee N-A) & (Rel-N \vee N-Rel))		cf. universal X for prepositional languages
XVIII	Postp \supset ((A-N \vee Rel-N \supset Dem-N & Num-N & Poss-N) & (Dem-N \vee Num-N \vee Poss-N \supset G-N))		PoNMH ^b derived from universals IV, V, VI, IX, XI, XII, XV, XVI, XVII, XXII, XXIII
XIX	Prep \supset (A-Adv \supset A-M-St)		
XX	Postp \supset (Adv-A \supset St-M-A)		
XXI		N-Poss \supset N-A	

Table 1 continued

XXII	Postp \supset (Poss-N \supset G-N)
XXIII	N-Poss \supset N-Rel

abbreviations: A (adjective), Adv (adverb), Dem (demonstrative), G (genitive), M (marker), N (noun), Num (numeral), Poss (possessive), Postp (postposition), Prep (preposition), Rel (relative clause), St (standard); word orders (S)OV, SVO, VSO, V1 (verb first)

^a The Prepositional Noun Modifier Hierarchy (PrNMH)

^b The Postpositional Noun Modifier Hierarchy (PoNMH)

By contrast, Cinque (1996, 2000, 2005) put forth an alternative analysis which saves Greenberg's original statement as a whole by advocating that all attested orders of noun modifiers are base generated in and derived from the prenominal sequence of Dem (demonstrative) – Num (numeral) – A (adjective) – N (noun). From this basic prenominal position, all attested postnominal orders, as he contends 14 out of $4! = 24$ logically possible orders, are derived by movement of the noun phrase adhering to movement parameters of varied markedness including pied-piping effects that are ascribed less marked properties (*cf.* Cinque, 2005: 321ff.). However, Cinque's construal has recently been challenged by Dryer (2009) based on a sample of 341 languages attesting to three additional word orders that are ruled out by Cinque (2005). To account for these findings, Dryer (2009) put forward an alternative analysis that considers semantic categories of demonstrative, numeral, and adjective irrespective of the syntactic phrases where they are realized, extending this notion also to Greenberg's universal 18 (see above). Conversely, support in favor of Cinque's interpretation was recently voiced by Abels (2011) in an analogous study on Germanic verb clusters who refuted Dryer's notion as too permissive. Eventually, evidence for revision of another Greenbergian universal was presented by Bennett (1979) who demonstrated that the formulation of implicational universal 23, which asserts that the order of the proper noun preceding the common noun usually corresponds to the order of noun preceding the genitive and vice versa, should be the exact opposite way, *i.e.* that the position of the common noun relative to the proper noun corresponds to the position of the noun relative to the genitive. A modification of this notion was suggested recently by Cinque (2009) who observed that head-initial SVO languages that are associated with a "noun preceding genitive" pattern display great heterogeneity as far as their ordering of proper and common noun is concerned, notably also intra-linguistically. Likewise heavily inspired by Greenberg's findings (1966), Dryer (1992) explored the order of paired grammatical elements in relation to the order of verb and object in OV and VO languages. Details of this comprehensive empirical study will be presented in section 5.1.

3.3. Adding the diachronic dimension: the notion of conspiracies

Not only did Greenberg make claims about synchronic universals of language expressed as implications (1966), in an essay on diachronic typology, he proposed the existence of diachronic universals “as well as law-like implicational relations, whether in the form of unrestricted or implicational universals in synchronic typology” (Greenberg, 1995: 146). He therein went on to state that, in the context of unrestricted universals and exceptions as transitional states, “the synchronic regularities are here the result of the interaction, often conspiratorial, of diachronic factors” (Greenberg, 1995: 153). In this regard, it is conspicuous that linguistic traits and with them languages, in the course of history sometimes take directions that either converge or diverge in a non-stochastic manner, even in unrelated languages with no attested contact influence that could account for such commonalities. The results of such processes, if stable over the longitudinal time axis, are seen as (positive or negative) correlations, which can be conceived of as having originated from change in one linguistic element, *e.g.* a different parameter setting, that entailed another one or several other changes, yielding, at a given time point, a certain cross-sectional pattern of synchronic correlations or implications: Such a mechanism throughout diachrony is in congruence with ideas described by Aitchison (1989) who contended that one linguistic change might trigger subsequent alterations on preferred routes in language evolution. Likewise in this vein, she later on (1996) on the one hand invokes notions brought up by Chomsky that changing the setting of a single parameter may elicit “complex effects with proliferating consequences” (Chomsky, 1981: 6) and that typologically different languages can be the effect of merely a few changes in parameters (Chomsky, 1986b: 152). On the other hand, she makes clear that any natural language system features constraints encompassing real constraints (prohibitions or filters) and pseudo-constraints (preferences) that surface as implicational universals. By the same token, Lightfoot (1991: 167) remarks that new parameter settings, which are typically manifested and accompanied by clusters of changes, may set off chain reactions. Thus relying upon these notions, it is understood that one element of a correlation appearing as implicational universal was cause whereas the other element(s) is/are effect of the diachronic change bringing forth the correlation seen in the synchronic perspective.

In extension of this, any association between linguistic traits which is proven significant by appropriate statistical testing is to be regarded a correlation or more specifically, an implication which can further be subdivided into implications exhibiting either mutual or one-way dependencies, a distinction which straightforwardly connects to Coseriu’s notion of reciprocal and unilateral implications (1974), respectively. Beyond that,

for a correlation or implication to be termed a conspiracy, it is hence adequate to claim that it further require, besides the establishment of cross-linguistic robustness in the cross-sectional synchronic dimension, demonstration of its longitudinal diachronic development, if possible likewise in more than one language depending on the availability of reliable historical data. Stated differently, an underlying conspiracy can surface as correlation of certain linguistic traits in a synchronic investigation. Conceptually, conspiracies can so be defined as accommodating both synchronic conspiracies (equaling implications that are usually confined to the synchronic aspect) and diachronic conspiracies (see section 6.1.). What is more, only the perspective of the longitudinal time dimension permits deeper insights for one of the traits that form part of a conspiracy to be revealed as trigger entailing as effect the alteration in the other correlated element(s). As an example, one such thoroughly studied conspiracy encompasses syntactic and morphological phenomena associated with the V2 word order in Germanic languages including unisententiation, strengthening of verb-final in embedded clauses, rise of subordinating conjunctions and of lexical subjects, and loss of verbal and nominal morphology: To be sure, such phenomena correlated with V2 were described in both the synchronic and the diachronic dimensions in members of Germanic languages, and, consistent with the above-stated definition, were labeled the “V2 conspiracy” (*cf.* Weerman, 1989; Platzack, 1990).

Arguably, if a language is analyzed for correlated phenomena at different stages in its history along its diachronic trajectory, conspiracies might be covered up, since they exhibit varying extents of (statistical) correlation at distinct synchronic cross-sections, contingent upon the distribution in syntactic contexts like clause or phrase type, until eventually reaching the maximum state of (perhaps complete) correlation in all syntactic contexts at a certain time point. While this development on the one hand reflects the concept of statistical and absolute implications, respectively, projected onto the diachronic dimension, it can on the other straightforwardly be accounted for by the notion of competing grammars (Yang, 2000) and competing grammatical subsystems (Kroch, 1989) that become manifest in transitional states that reflect more than a single grammar, thus blurring correlations that might be stronger or even absolute in case of only one prevailing grammar. In a similar vein, Vennemann (1974b) held the view that inconsistencies in the typological classification of VO and OV language types mirror evolution from one type to the other. And Greenberg (1995: 150) remarked that a diachronically-oriented typology might aid in overcoming the challenges posed to synchronic typology by languages that do not fit into well-defined typological categories, by providing explanations in terms of stages gone through in the process of change from one frequent and

stable type to another such type. This is also in line with Croft's (1990: 209) point of view that exceptions to synchronic universals should be explained by diachrony, since these exceptions are indicative of unstable, transitional states.

What is more, the observed variance in statistical correlations can be ascribed to the nature of the linguistic traits (and their acquisition) themselves: In detail, it cannot be ruled out that one linguistic change in trait (or parameter) A, corresponding to the re-setting of one parameter to another value, elicits another change in trait (or parameter) B, corresponding to the subsequent change of the value pertaining to another parameter, but the same alteration in B does not trigger A to undergo change. Hence the result of change in A should only be seen in the context of the change in B it entails (except in a very early state of the process), whereas the alteration in B could feasibly occur devoid of the change in A if alternative pathways from the change in A to generate change B exist and/or if erosion of the "more archaic" trait A has taken place. An alternative view was expressed by the principle of Universal Consistency in History (UCH) (Hawkins, 1983: 210-212) advocating that synchronic implicational universals display consistency throughout the history of languages. This notion, according to Hawkins (1983), implies that if in the synchronic dimension a parameter value in trait A occurs only in the context of a parameter value of another trait B but not vice versa, it must have been either a change in trait B that elicited the change in A or a simultaneous event in A and B. To resolve such obscurities, diachronic investigation is potentially apt to differentiate the causal change from the ensuing effect(s). Furthermore, the (originally synchronic) concept of unilaterality *vs.* reciprocity can appropriately be applied to conspiracies also along the diachronic axis, just as the notion of positivity *vs.* negativity of implications, desirably corroborated by cross-linguistic evidence. Consequently, reciprocal relations should be especially well-suited to surface as conspiracies, or stated otherwise, associated or conspired linguistic features that display high correlation might bear a fair likelihood to reflect a mutually dependent diachronic development, *i.e.* reciprocity.

The analysis of diachronic data thus bears enormous relevance for elucidating structure of linguistic system(s), including universals of language which conspiracies might be a part of. By this token, Kroch (1989: 200) underscores the necessity of diachronic study supplementing the synchronic dimension for understanding the evolution of as complex a system as a linguistic grammar, which alone permits the establishment of sound linguistic

theory, since historic data contain information that is not offered by analysis of solely synchronic material.¹⁸

¹⁸ “In historical materials, we find a kind of information that is necessarily absent in synchronic data and that offers the prospect of an important contribution to general linguistics from history, that is, information about the time course of language change. With such process information, we may hope to learn how the grammars of languages change from one state to another over time and, from an understanding of the process by which they change, to learn more about their principles of organization. After all, perturbing a complex system and observing its subsequent evolution is often an excellent way of inferring internal structure. In addition, since the features of any language at a given point in time are the result of a complex interweaving of general principles of language and particular historical developments, knowledge of the historical process by which a language has reached a given state may be important to the proper assignment of responsibility to historical and general factors and so to the proper formulation of linguistic theory.” (Kroch, 1989: 200)

4. Language acquisition and the emergence of conspiracies

Taken together, linguistic traits like syntactic phenomena that display correlation are of outstanding interpretative value for linguistic theory exploring possible structures of natural language, language typology, and the study of language acquisition alike, this even more if their existence is attested cross-linguistically in synchrony and throughout diachronic change, which, in accordance with the gross distinction provided in section 3.3., warrants their designation as conspiracies. Interestingly, such correlated (or conspired) phenomena not only represent regular patterns in grammatical systems, but might also mirror regularities in the process by which linguistic knowledge is attained and varied over time, *i.e.* in language acquisition. This concept is borne out *e.g.* by Lightfoot (1991: 167) who observed that chain reactions set off by new parameter settings can be understood through the acquisition process. Stated conversely, cognitive principles of language learning might be held responsible for certain linguistic traits or parameter settings to occur along with each other or avoid each other. Along this line, it is conceivable that one change in a parameter value, whatever the cause, could elicit one or several further changes because of improved ease of cognitive processing during first language learning entailed by these innovations in the context of the first change. One important mechanism by which such intra-linguistic changes are implemented arguably is, as mentioned above in section 2.1., the minimizing of opaque syntactic structures by reanalysis (*cf.* Lightfoot, 1979).

The postulation of conspiracies reflecting (cognitive) principles in first language acquisition seems furthermore warranted in the light of regularities witnessed in the process of creolization: In circumstances where pidgin languages with impoverished linguistic structure arise as means of communication amongst (usually immigrant) speakers with diverse linguistic backgrounds and limited knowledge of the language spoken by the majority of the population, featuring elements from the pidgin speakers' original substrate languages and the superstrate language of the community alike, the pidgin speakers' offspring develop creoles that display both important differences from the parents' pidgin and noteworthy commonalities (*cf.* Bickerton, 1999): One core finding in every process of creolization is that grammatical structure where there was none in the parents' pidgin is created rapidly in creole languages, reflecting a marked contrast between the linguistic input and the competence eventually attained. A further intriguing commonality of all creoles is that some of these newly developed structures can be traced neither to the substrate nor the superstrate languages. And to add one more aspect, some of these newly formed grammatical structures, irrespective of the substrate and superstrate languages, recur throughout creoles over the

world. Such general properties of creole languages importantly encompass the occurrence of pre-verbal particles, simple inflectional morphology, SVO word order (*cf.* Muysken, 1988), and double negation (negative concord) which was evidenced in many but not all creoles, notably also in instances where neither substrate nor superstrate could have conveyed this property (Holm, 2000: 195ff.). Beyond that, the poverty of inflectional markers generally entails absence of syntactic properties like inversion of subject and verb and the optional omission of subject pronouns (Rizzi, 1982), which is also evidenced in creoles with Spanish and Portuguese as superstrates that both do display these syntactic features (Muysken, 1988). Such grammatical regularities that arise first of all counter the input from the original pidgin and moreover within a strikingly short period of time (one generation off the pidgin language usually creates the core features of a creole language) hence mirror general cognitive principles of language acquisition that are guided by and form part of UG. Consistently, the role of first language acquisition in the creolization process was underscored by Sankoff and Laberge (1973) and Bickerton (1984). Beyond evidence provided by creole studies for learning regularities leading to common cross-linguistic structure and typological profiles, Senghas *et al.* (2004) demonstrated the emergence and development of syntactic structure in a study on Nicaraguan sign language. Similarly, Sandler *et al.* (2005), in a sign language used by an isolated community in Israel, found regular grammatical structures which were absent in both spoken and other sign languages around. Along these lines, Hudson Kam and Newport (2009) observed regularization of linguistic input inconsistencies amongst both adult learners and even more child learners in experiments using an artificial language. These results are borne out by another experimental study (Real and Griffiths, 2009) employing an artificial language paradigm that reported strong regularization effects in an iterated model (*cf.* Kirby *et al.*, 2004) simulating adult learning over generations. Thus demonstration of the existence of a grammatical regularization bias amongst (not only first) language learners provides strong indication for a link between learning-related cognitive processes and recurrent cross-linguistic patterns. The notion that such learning biases indeed underly also linguistic implications is lent support to in a further experimental work that links adult learners' biases to one of Greenberg's implicational universals (Culbertson *et al.*, 2012). Section 6.2. will get back to this issue in more detail.

Conspired linguistic traits therefore arguably provide insight not only into the design of human linguistic systems, but also into general cognitive foundations like those of (not only first language) learning. In this regard, the following section 5 and its subsections are dedicated to a selection of striking phenomena where syntactic traits go along with or avoid

each other, with due emphasis on their diachronic development, thus justifying their designation as conspiracies. Specifically, a selection of syntactic features conspired cross-linguistically with either VO or OV word order will be presented, as well as the conspiracy of syntactic phenomena related to the movement of the verb to TP in Germanic languages, and the mutual exclusiveness of V2 and drop of subject during the diachrony of Western Romance languages. Notwithstanding the fact that the demonstration of diachronic evidence for a given conspiracy will be confined to a selection of languages or even to only one language, universal validity can nevertheless be extrapolated or at least be hypothesized on basis of such findings.

5. Syntactic conspiracies exemplified

5.1. Conspiracies of VO and OV word order

Not only is basic word order a useful feature of classification for language typology, more specifically the position of the verb relative to its object(s) in a sentence is regarded as one of the parameters pertaining to the framework of principles and parameters (P&P) (*cf.* Chomsky and Lasnik, 1993; Chomsky, 1995) that are assumed to be set in an early phase of language acquisition (Lightfoot, 1991). As reported in section 3.2. from works of Greenberg (1966), Vennemann (1974a,b), and Hawkins (1983), a number of syntactic features have been shown to associate with languages of the (S)VO type (*i.e.*, the verb predominantly precedes its object(s)) or of the (S)OV type (*i.e.*, the verb predominantly follows its object(s)). In this vein, expanding on Greenberg's findings (1966), Dryer (1992), in a comprehensive synchronic study including data from 625 languages, pinpointed pairs of syntactic elements whose relative order correlated with that of the verb and object (thus, either the OV or the VO type). By categorizing the languages under investigation into clusters of related "genera" (corresponding roughly to the subfamilies of Indo-European) and assigning those to six large geographic areas, Dryer (1992) in his data set examined word order parameters and their relationship with the order of verb and object. Accordingly, following Dryer, elements of such correlation pairs that go along with the verb position are termed "verb patterners", whereas "object patterners" tend to associate with the object position.¹⁹ Table 2 provides a summary of the findings in Dryer (1992) including evaluation of results by chi-square testing that are not part of the original article. Even though *p*-values for two of the word order parameters classified as non-correlation pairs are <0.05 indicating statistical significance, Dryer reports both parameters, *i.e.* position of demonstrative and numeral relative to the noun, as not correlated with verb and object order. Conversely, the complementizer position which is not amenable to chi-square testing due to scarcity of data, is nevertheless regarded as being in correlation with the relative position of verb and object (Table 2). Notably, explanations for such observed word order correlations include the head-dependent theory and the head-complement theory which both state that the head usually takes the same position as the verb, and the complement, which is dependent on the head, that of the object (Dryer, 1992), much akin to Vennemann's NSP (1974a,b; *cf.* section 3.2).

¹⁹ *cf.* Dryer (1992: 87): "If a pair of elements X and Y is such that X tends to precede Y significantly more often in VO languages than in OV languages, then (X, Y) is a CORRELATION PAIR, and X is a VERB PATTERNER and Y an OBJECT PATTERNER with respect to this pair."

Table 2. Tabulation of word order parameters and their correlation with the relative order of object and verb based on data from language “genera” representing sets of related languages. The chi-square (χ^2) test was employed to test for correlation pairs reported in Dryer (1992). Word order parameters are presented in decreasing order of their corresponding χ^2 values.

word order parameter	order	association with OV [%]	association with VO [%]	CC ^a	χ^2	p ^b
correlation pairs						
adposition	postposition	93.86	14.63	0.625	125.507	3.942E-29
	preposition	6.14	85.37			
verb (V) and adpositional phrase (PP)	PP-V	87.50	1.67	0.650	96.534	8.772E-23
	V-PP	12.50	98.33			
noun (N) and genitive (Gen)	Gen-N	90.32	32.26	0.517	79.220	5.556E-19
	N-Gen	9.68	67.74			
verb (V) and manner adverb (Adv)	Adv-V	91.43	24.14	0.566	60.337	7.995E-15
	V-Adv	8.57	75.86			
verb (V) and subject (S)	S-V	97.78	57.94	0.444	59.565	1.183E-14
	V-S	2.22	42.06			
copula (Cop) and predicate (Pred)	Pred-Cop	88.89	21.82	0.561	58.369	2.173E-14
	Cop-Pred	11.11	78.18			
adverbial subordinator position	clause-final	69.09	1.67	0.580	58.206	2.361E-14
	clause-initial	30.91	98.33			
content verb (VP) and tense/aspect auxiliary verb (Aux)	VP-Aux	92.31	12.50	0.625	45.514	1.516E-11
	Aux-VP	7.69	87.50			
adjective (Adj) and standard of comparison (St)	St-Adj	80.56	3.13	0.614	41.198	1.375E-10
	Adj-St	19.44	96.88			
"want" (Want) + subordinate verb (V)	V-Want	74.36	8.70	0.557	38.315	6.020E-10
	Want-V	25.64	91.30			
noun (N) and relative clause (Rel)	Rel-N	41.27	1.64	0.433	28.577	9.005E-08
	N-Rel	58.73	98.36			
tense-aspect affix position on verb	suffix	91.60	63.77	0.326	22.324	2.303E-06
	prefix	8.40	36.23			
plural word position relative to noun (N)	following N	100.00	34.78	0.550	16.957	3.825E-05
	preceding N	0.00	65.22			
content verb (VP) and negative auxiliary verb (NegAux)	VP-NegAux	72.73	7.14	0.561	11.500	6.959E-04
	NegAux-VP	27.27	92.86			

Table 2 continued

question particle position	final	71.11	40.43	0.295	8.762	0.00308
	initial	28.89	59.57			
article (Art) and noun (N)	N-Art	59.38	28.85	0.289	7.663	0.00564
	Art-N	40.63	71.15			
complementizer position	final	no numbers reported	not attested in data			
	initial	no numbers reported	100.00			
non-correlation pairs and unclassified cases						
demonstrative (Dem) and noun (N) ^c	Dem-N	71.17	55.21	0.163	5.677	0.017
	N-Dem	28.83	44.79			
numeral (Num) and noun (N) ^d	Num-N	50.00	67.09	0.169	5.361	0.021
	N-Num	50.00	32.91			
verb (V) and negative particle (Neg)	V-Neg	26.19	14.00	0.151	2.155	0.142
	Neg-V	73.81	86.00			
possessive affix position on noun	suffix	38.96	50.00	0.109	1.574	0.210
	prefix	61.04	50.00			
verb (V) and tense/aspect particle (T/A)	V-T/A	25.00	13.79	0.141	1.077	0.299
	T/A-V	75.00	86.21			
intensifier (Intens) and adjective (Adj)	Intens-Adj	60.94	51.92	0.090	0.951	0.330
	Adj-Intens	39.06	48.08			
noun (N) and adjective (Adj)	Adj-N	42.64	42.11	0.005	0.006	0.937
	N-Adj	57.36	57.89			

^a contingency coefficient

^b numbers in bold indicate statistical significance at $p < 0.05$

^c $p < 0.05$ notwithstanding, Dryer assumes no statistically significant correlation, since averaged proportions of genera of either OV or VO type that associate with either parameter order are virtually identical over all geographic areas

^d $p < 0.05$ notwithstanding, Dryer assumes no statistically significant correlation, since data from Africa run counter to the overall trend

However, since in many instances of the presented data there is no support for the notion conferred by the head-dependent and the head-complement theory, Dryer (1992) proposed as alternative explanation the branching direction theory that asserts a strong tendency for non-phrasal, syntactically non-branching, categories to pattern like verbs, and phrasal, *i.e.* branching, categories to pattern like objects. A corollary of such a dichotomy is encountered in the right- or left-branching structure of languages.

After demonstration of the synchronic situation, emphasis henceforth in this section

will be placed on the diachronic development of two of the syntactic parameters coupled with the relative order of object and verb listed in Table 2, *i.e.* the position of the VP relative to the auxiliary (Aux), and, to a minor extent, that of the adposition. Judging from Table 2, one huge conspiracy involving all correlation pairs displaying statistical significance could be suspected, however, the diachronic situation does not always give rise to a straightforward interpretation, as will be apparent from the following examples.

Syntactic position of the VP relative to the auxiliary in conspiracy with relative order of verb and object is evidenced in the transition from Classical Latin - the language used by the educated classes of Rome in the first century BC (according to Vincent, 1988: 27), the period of Cicero and Caesar - to modern Romance languages. Even though Classical Latin word order was rather unrestricted, the default sequence is considered to have been preferentially OV (Vincent, 1988: 59ff.). Concomitantly and in line with synchronic observations by both Greenberg (1966) and Dryer (1992), the auxiliary verb usually followed the main verb as demonstrated in the examples below where the auxiliary appears underlined and in bold whereas the main verb is underlined only:

*Quantum facinus ad vos delatum **sit**, videtis.*²⁰

How much/big - crime - to - you (pl.) - reported - was (conj.) - you (pl.) see

‘You see what a big crime was reported to you.’

*Latius opinione disseminatum **est** hoc malum;*²⁰

Further - than the opinion - spread (part.) - has - this - evil

‘This evil has spread further than one thinks;’

*[...] quacumque ratione placet, celeriter vobis vindicandum **est**.*²⁰

by whichever - way - it is liked - quickly - you (pl., ind. obj.) - to take action - is

‘whichever way you like, you must take action quickly.’

However, counter-evidence where the auxiliary precedes the main verb, albeit compatible with the notion that classical Latin used to be not strictly OV, occurs in the same text:

*[...] vitae tantam laudem, quanta vos me vestris decretis honestastis, nemo **est** adsecutus.*²¹

in life - as much - glory - as that by which - you (pl.) - me - with your (pl.) - decrees - honored - no one - has - achieved

‘no one has achieved in life as much glory as that by which you honored me with your decrees.’

²⁰ M. Tullius Cicero, “Oratio in Catilinam Quarta in Senatu Habita” (“4th oration against Catilina, delivered in the senate”), §6

²¹ M. Tullius Cicero, “Oratio in Catilinam Quarta in Senatu Habita” (“4th oration against Catilina, delivered in the senate”), §21

Romance languages of to-date, all of which have their origin in Latin, obey the VO pattern accompanied by strict adherence to the sequence of the auxiliary foregoing the main verb, as exemplified by the following Spanish and Portuguese sentences (main verb underlined, object in bold, auxiliary bold and underlined):

*Durante varios meses se empeñó en demostrar el **acierto** de sus conjeturas.*²²

During - several - months - 3rd pers. refl. pron.-insisted - on - prove - the - correctness - of - his - conjectures

‘During several months he insisted on proving the correctness of his conjectures.’

» *En el mundo están ocurriendo cosas increíbles* « – *le decía a Úrsula.*²²

In - the - world - are - happening - things - incredible - (ind. obj. clitic) - used to say - to - Úrsula

‘“Incredible things are happening in the world”, he used to say to Úrsula.’

*El laboratorio había sido desempolvado.*²²

The - laboratory - had - been - freed from dust

‘The laboratory had been freed from dust.’

*Em sua cidade devia haver um pipoqueiro.*²³

In - his - town - should - there be - a - popcorn seller

‘In his town there should be a popcorn seller.’

Presumably, the transition of (predominantly) OV and V-Aux to VO and Aux-V occurred during the period of Vulgar Latin (4th or 5th century AD) as stated by Harris (1978: 19).

An analogous development is observable in the transition from Old English (appr. 450 – 1066; cf. Campbell, 2004: 9) to Middle English (1066 – 1476; cf. Campbell, 2004: 8), when a change took place from V2 word order in main clauses and (predominantly) OV word order in subordinate clauses, a situation that exists in modern High German, to a consistent VO word order throughout. The next example sentence demonstrates that the object (in bold) was able to precede the verb (underlined) in Old English subordinate clauses (Pintzuk, 1991: 77; Pintzuk, 1999):

*[...] þæt he his **stefne** up ahof.*²⁴

that - he - his - voice - up - raised

‘that he raised up his voice.’

²² Gabriel García Márquez, “Cien Años de Soledad” (“One Hundred Years of Solitude”)

²³ Paulo Coelho, “O Alquimista” (“The Alchemist”)

²⁴ Bede 154.28

In Old English subordinate clauses with OV order, the auxiliary, underlined and in bold in the following example, used to follow the main verb, underlined only, which is what in principle applied to Classical Latin (in all clause types) as well (Pintzuk, 1991: 117; Pintzuk, 1999):

[...] *forþon of Breotone nædran on scippe lædde **wæron**.*²⁵

because - from - Britain - adders - on - ships - brought - were

‘because vipers were brought on ships from Britain.’

At the same time, however, not always did the nominal object (in bold, not underlined) precede the verb in Old English subordinate clauses, even though the verb did precede its auxiliary (Fischer *et al.*, 2000: 144):

[...] *þæt hi urum godum geoffrian **magon** dancwurde **onsægednysse**.*²⁶

that - they - our - gods - offer - may - grateful - sacrifice

‘that they may offer a grateful sacrifice to our gods.’

Conversely, in cases where the object preceded the verb, the auxiliary could do so likewise (Fischer *et al.*, 2000: 52):

[...] *se ðe nan ðing **nele** on ðissum **life** ðrowian.*²⁷

he - who - no - thing - not-wants - in - this - life - suffer

‘he who will suffer nothing in this life.’

In an attempt to account for these word order inconsistencies, in Old English such variation has been discussed to reflect competing grammars (*cf.* Kroch, 1989; Yang, 2000; sections 2.2.3. and 3.3.), eventually resulting in change to the VO type grammar associated with auxiliaries preceding the verb, which is evidenced from Middle English up to nowadays (Pintzuk, 2002). Also OV type classical Latin, as demonstrated above, used to be inconsistent in its (predominant) ordering of the auxiliary following its dependent verb.

A similar pattern is found in Old Norse, the predecessor of modern Scandinavian languages which was spoken in Norway and Iceland up to the 14th and even 16th century AD, respectively (Barnes, 2008: 1-2). Old Norse is typified by OV word order, which changed to VO in all modern Scandinavian languages. As for Icelandic, Hróarsdóttir (2000: 140, 160) analyzed text corpora from several centuries, observing a marked decline in the percentage of auxiliaries following the main verb accompanied by the rise of VO order from the 14th to 19th century (Table 3): Nominal objects preceding the verb (OV word order) and following it (VO

²⁵ Bede 30.1-2

²⁶ ÆCHom I, 38.592.31

²⁷ ÆCHom I, 10.164.22

word order) were used in approximately equal frequencies up to the 16th century in all syntactic contexts, followed by a decrease of OV occurrence to roughly 12% by the 19th century ($\chi^2=361.08$, $p=7.19\times10^{-76}$, contingency coefficient: 0.342). In parallel, a drop in frequency of auxiliaries following the main verb is observed over all syntactic contexts, reaching likewise 12% by the 19th century ($\chi^2=104.12$, $p=7.14\times10^{-21}$, contingency coefficient: 0.491). A firm link between the two phenomena is established by a statistically significant correlation of % OV and % V-Aux over all centuries (Pearson's correlation coefficient: 0.852, $p=0.031$).

Table 3. Combined syntactic change of OV to VO and V-Aux (main verb preceding the auxiliary) to Aux-V (main verb following the auxiliary) word order in the history of Icelandic. Figures represent either numbers of cases encountered in a corpus from the corresponding century or percentage of all cases.

century	OV	VO	% OV	V-Aux	Aux-V	% V-Aux
14 th	98	120	44.95	25	7	78.13
15 th	67	77	46.53	7	8	46.67
16 th	64	64	50	16	11	59.26
17 th	148	209	41.46	17	22	43.59
18 th	31	105	22.79	6	13	31.58
19 th	202	1544	11.57	36	264	12

For illustration, the following sentences from Old Norse show that the main verb (underlined) used to follow the object (in bold), both in main clauses like in the first example (Barnes, 2008: 225) and in subordinate clauses like in the second example (Hróarsdóttir, 2000: 159), and that furthermore an auxiliary (bold and underlined) used to follow the main verb as demonstrated in the second sentence:

*Hann hafði **heit** strengt.*

He - had - oath - sworn

‘He had sworn an oath.’

*[...] að eg mundi **hann** sigrad geta.*

that - I - would - him - defeat - can

‘that I would be able to defeat him.’

A sentence from contemporary Icelandic (Zaenen *et al.*, 1990: 111), by contrast, shows the association of VO word order with all auxiliaries (in bold and underlined) preceding the main verb (underlined only) that is followed by the object (bold only):

*Ég tel henni **hafa verið skilað** peningunum.*

I - believe - her - to have - been - returned - the money

‘I believe her to have been given back the money.’

As for adpositions contingent upon word order type, neither Old English nor Classical Latin nor Old Norse were predominantly postpositional languages, as would however be expected in the light of synchronic data like Greenbergian universal 4 (*cf.* section 3.2.) and Dryer’s findings (Table 2). A deeper look into history, however, reveals that the reconstructed proto-Indo-European language was beyond doubt of the OV type (Lehmann, 1993: 202), with sufficient evidence that it used to be postpositional too (Lehmann, 1993: 203ff.). As a consequence, it is plausible to assume that Old Norse, Classical Latin, and Old English alike represent transitional states on the way to prepositional languages of the VO and AuxV type from a common origin that used to be postpositional with concomitant OV and VAux word orders. From the available evidence of the adduced examples, it can furthermore be concluded that still in an OV context the postpositional property was lost first, and that the auxiliary was raised above the verb later concurrently with the transition to VO type languages. This finding is somewhat unexpected in light of the synchronic data, because in Dryer’s study (1992) both correlation pairs exhibit the same contingency coefficient of 0.625, and the *p*-value is even lower for the adposition parameter (Table 2).

5.2. Transitive expletives and object shift in Germanic

Two syntactic phenomena are well described from Germanic languages, *i.e.* the occurrence of transitive expletive (non-referential) constructions (Bobaljik and Jonas, 1996), and object shift, describing overt raising of the object out of VP, the latter known as Holmberg’s Generalization (*e.g.*, Holmberg, 1999). Importantly, these two phenomena were reported to occur along with each other in Germanic when object shift is restricted to NP nominal objects, an association which is labeled Bures’s correlation (Bures, 1992). Expanding on this notion, Bobaljik and Jonas (1996) demonstrated that it is movement of the verb from the VP up to the TP in finite clauses in conjunction with movement of the subject to Spec-TP that facilitates these phenomena. What the authors further showed for contemporary Germanic languages is that Icelandic, German, Dutch, Yiddish, and Frisian, which are all characterized by movement of the verb to the TP, allow of transitive expletive constructions and full NP nominal object shift, whereas English, Norwegian, Swedish, Danish, all of which do not permit expression of the verb in TP, do not (but pronominal objects do shift in

Mainland Scandinavian). Moreover, Afrikaans, albeit displaying nominal object shift, does not permit transitive expletives, whereas in Faroese nominal object shift is ruled out while transitive expletives are, except for one dialect, grammatical. Since both Afrikaans and Faroese show V to T raising and subject movement to Spec-TP, it is therefore clear that a conspiracy with transitive expletives and object shift is unilateral, also because languages outside Germanic that express the verb in TP *via* movement from VP do not always display both phenomena, like *e.g.* French that is lacking transitive expletives. At this point, a transitive expletive construction figures as an example sentence from Faroese in the following (Bobaljik and Jonas, 1996: 209):

Tað bygdu nakrir íslendingar hús í Havn.

There - built - some - Icelanders - houses - in - Torshavn

‘Some Icelanders built houses in Torshavn.’

In addition, a further example presents a shifted object in Icelandic, including an analysis of the relevant syntactic structure (Bobaljik and Jonas, 1996: 202):

Jólasveinninn bordaði [_{AgRP} *hattinn* _i [_{VP} *ekki* _{t_v} _{t_i}]].

The Christmas troll - ate - the hat - not

‘The Christmas troll did not eat the hat.’

Modern Icelandic shall likewise serve to illustrate the detailed syntactic situation encountered in a sentence with both a transitive expletive construction and a shifted object in Fig. 1 as analyzed by Bobaljik and Jonas (1996: 217). It becomes clear that the object moves out from VP and is raised to the specifier position of the object agreement phrase. While the subject moves from Spec-VP to Spec-TP, the verb is raised even further than to T, *i.e.* to the subject agreement phrase (Fig. 1).

A situation where, when movement of the verb to T disappeared, transitive expletives and object shift also vanished, is found in the history of English: As the following example from 16th century Early Modern English (full period: 1476 – 1700, *cf.* Campbell, 2004: 7) demonstrates, negation (underlined and in bold) could come between the main verb (underlined only) and its direct object (so could also adverbs), indicative of movement of the main verb out of the VP up to the TP (Roberts, 1999: 290):

[...] if I **gave** **not** this accompt to you.²⁸

‘if I did not give this account to you.’

²⁸ J. Cheke (Letter to Hoby, 1557)

Representing the same period, other texts attest to the existence of transitive expletives, as shown in the first of the following examples (Jonas, 1996: 151), and object shift, exemplified with a pronomial object (underlined) that is raised out of the VP above the negation (Roberts, 1995: 274):

*Within my soul there doth conduce a fight [...]*²⁹

‘Within my soul, a fight does begin [...]

*[...] if you knew them not.*³⁰

‘if you did not know them.’

It is needless to remark that none of those constructions are grammatical to-date, since in modern English the main verb is not allowed to move out of its VP, results of which are that adverbs cannot separate main verb and object (**I know very well him*), whereas auxiliaries including periphrastic “do” are raised to TP.

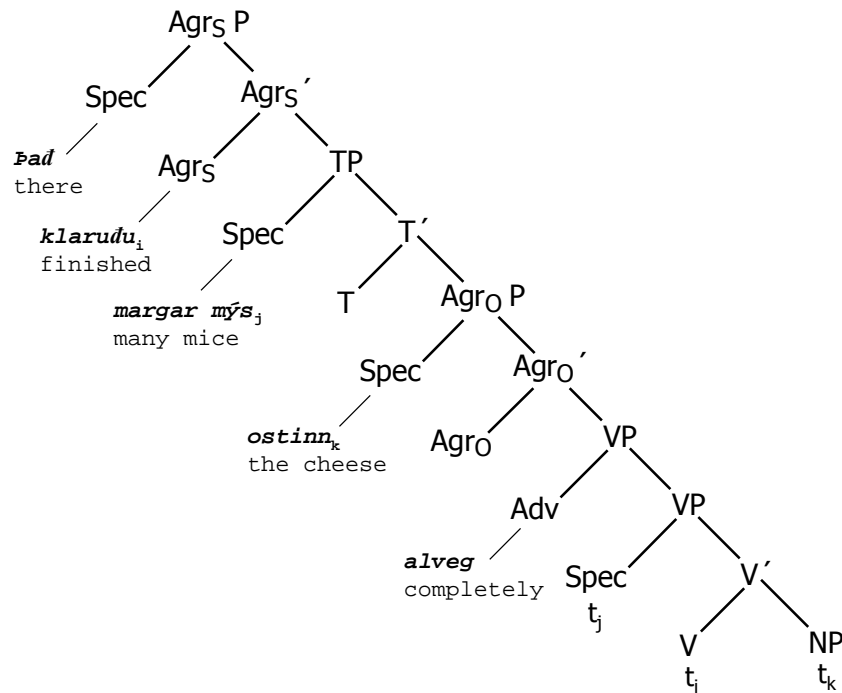


Fig. 1: Syntactic structure of an Icelandic sentence with the English reading “Many mice completely finished the cheese”, displaying both object shift, leaving trace t_k in NP, and a transitive expletive expression that accounts for a trace of the subject t_j in Spec-VP and a trace of the verb t_i in V. Agr_SP and Agr_OP denote agreement phrases of the subject and object, respectively (after Bobaljik and Jonas, 1996).

²⁹ W. Shakespeare (Troilus and Cressida, Scene 2 in 5th Act)

³⁰ John Lyly, 1580

5.3. V2 word order and *pro*-drop: a negative conspiracy in Western Romance

During the medieval ages, Western Romance languages (most importantly, French, Spanish, Portuguese) as well as northern Italian dialects used to display both V2 word order and the null-subject parameter turning them into *pro*-drop grammatical systems that permitted the pronominal subject to be left out (*e.g.*, Vanelli *et al.*, 1986; Adams, 1987; Yang, 2000: 243). Importantly, a cluster of conspired syntactic properties is reported for V2 word order and the null-subject parameter each: At least in Germanic, several traits were described to go along with V2, as mentioned in section 3.3., and as for null-subject languages, clitic raising, infinitive-clitic word order, free inversion of subject and verb, and the lack of complementizer-trace effects are often cited as accompanying phenomena (*cf.* Roberts, 2007: 34-35).

In the history of French, diachronic development with respect to the configuration of V2 and the null-subject parameter is particularly well-studied (*e.g.*, Adams, 1987; Roberts, 1993): Subsequent examples (Roberts, 1993: 124; Clark and Roberts, 1993: 332) illustrate that V2 (verb underlined) and omission of pronominal subjects were perfectly grammatical in Old French, the period from the documents known as Oaths of Strasbourg in 842 up to 1300 (Cerquiglini, 2007; Adams, 1987: 26), and Middle French, following up to approximately 1500 (Adams, 1987: 26):

*Tresqu' en la mer cunquist la tere altaigne.*³¹

Until - the - sea - conquered (3rd pers. sing.) - the - land - high

'He conquered the land all the way to the sea.'

Or ai eu plusieurs fois grant imagination.

Now - have (1st pers. sing.) - had - several - times - great - imagination

'Now I have had a great imagination several times.'

In conjunction, free inversion of verb and subject (the latter underlined in the following example, appearing as last word in the sentence) occurred as further property associated with the *pro*-drop parameter during Old and Middle French (Vance, 1997: 77):

Tant fu de bone hore nez li chevaliers!

So much - was - of - good - hour - born - the - knight

'The knight was born at such a propitious hour!'

³¹ Roland, 1. 3

During Middle French, however, a rise of expressions incompatible with a V2 grammar can be observed, like in the subsequent sentence where the verb (underlined) is in a V>2 position (Yang, 2000: 243):

Après disner le chevalier me dist [...]

After - dinner - the - knight - to me - said

‘After dinner the knight said to me [...]’

For the transition from Middle French to early modern French, Roberts (1993: 199) furthermore reported a summary of corpus-based percentages of sentences where the verb followed the subject (SV pattern, compatible with both an SVO and a V2 grammar) or the verb was followed by the subject (VS pattern including XVSO and OVS, compatible with a V2 grammar, but not with the SVO type), and the percentage of sentences where the subject pronoun could be dropped: Whereas in sample sentences from the 15th century, the SV pattern occurred in 48% of cases, being already more dominant than the VS pattern in 10% of instances, with the subject pronoun omitted in 42% of sample sentences, in corpora from 16th century French sentences showing the SV pattern increased to 77% while VS type sentences and drop of the subject pronoun declined concomitantly to 3% and 15%, respectively. For early modern French, information offered by contemporary grammarian Maupas at the beginning of the 17th century (cited in Roberts, 1993: 215ff.) portends the absence of subject pronouns still in certain contexts, notably after conjunctions like *et* (“and”) and *si* (“so”, “thus”), when the subject pronoun is 1st or 2nd person plural, and in non-referential expressions, the latter of which became the only null-subjects in fixed expressions during the 17th century. As mentioned before, whereas French is devoid of both V2 expressions and the *pro*-drop parameter, other Western Romance languages like Spanish and Portuguese dispensed with only V2 word order while keeping the option to drop pronominal subjects, entailing also other conspired properties like free inversion of verb and subject. Two example sentences from modern Spanish, the first demonstrating *pro*-drop and the second free inversion (verb in bold, subject underlined), might illustrate the situation of to-date:

*Cuando se hizo experto en el uso y manejo de sus instrumentos, [...]*²²

When - himself - made - expert - in - the - usage - and - handling - of - his - instruments

‘When he became expert in the usage and handling of his instruments, [...]’

*José Arcadio Buendía no creyó que **fuera** tan rígida la voluntad de su mujer.*²²

José - Arcadio - Buendía - not - believed - that - was - so - rigid - the - intention - of - his - wife

‘José Arcadio Buendía did not believe that his wife’s intention was so rigid.’

As furthermore expected, modern French that is devoid of null-subjects, also lacks free inversion. However, a phenomenon labeled “Stylistic Inversion” can be analyzed as null-subject-related rudimentary trace in modern French (Kayne and Pollock, 2001) that makes its appearance in but restricted contexts like interrogative sentences and relative clauses, subject underlined in the example sentences below (Adams, 1987: 24):

D’où est venue cette idée?

From - where - is - come - this - idea

‘From where has this idea come?’

Le roman qu’avait écrit Margot

The - novel - that - had - written - Margot

‘The novel that Margot had written’

Syntactic explanations for the disappearance of (at least) the V2 word order parameter in the presence of the *pro*-drop property in the history of Western Romance languages include those put forth by Roberts (1993) and Yang (2000): Roberts (1993) construed the decline of both V2 and the null-subject in the history of French as a multi-step process: In Middle French, the analysis of the verb pertaining to an SVO syntactic sequence in the CP was optional, since it could as well localize in the agreement phrase (Roberts, 1993: 153). During early modern French in the early 16th century, the eventual elimination of V2 clauses was due to the underlying change in the nominative case assignment parameter of the verb in C⁰ that licensed or case-marked the subject. The result was loss of the agreement feature of C⁰ from where the verb governed *pro* in the specifier position of the agreement phrase, thus precluding the verb from being raised to CP. Subsequently, grammatical contexts decreased in frequency that licensed null subjects in the agreement phrase under government (agreement had been ruled out as an option for licensing *pro* already during Old French when the extensive repertoire of morphological agreement markers diminished), leading to the erosion of null-subjects in the course of the 17th century (Roberts, 1993: 204 ff.).

By contrast, an alternative explanation to account for the loss of V2 at the expense of the SVO word order type was suggested by Yang (2000: 241 ff.): The existence of V2 along with the dropping of subject pronouns created contexts where XVS word order patterns could better be parsed by an SVO type grammar, resulting in first a diminished advantage and then in a disadvantage of a V2 type relative to an SVO type grammar. In French, the option to drop subject pronouns was lost after the demise of V2, whereas Spanish and Portuguese both preserved the null-subject parameter once V2 had disappeared. Hence, under this analysis it is

plausible to posit that “the combination of *pro*-drop and V2 is intrinsically unstable” (Yang, 2000: 243) and that a V2 word order type and the null-subject parameter represent a (tentatively universal) bilateral negative conspiracy.

6. Discussion and conclusions

6.1. Classification of conspiracies

In section 3.3., the basic concept of conspiracies was introduced as an association between or implication of linguistic traits, also to be termed correlation if found statistically significant, that in addition harbors the aspect of common diachronic development. This idea will be elaborated into a systematic classification in what follows: Unless explicitly confined to the synchronic dimension in which case rendering conspiracies synonymous with (synchronic) implications, it is suggested herein that a conspiracy in its general meaning cover the correlated development of linguistic traits on the diachronic axis as well, since the concept incorporated in the term “conspiracy” bears the intrinsic property of mutual attraction of the elements that constitute it. This notion of “attraction” implies a process, thus including (diachronic) development over time, whose stages or results can be observed at certain time points, *i.e.* synchronically. By contrast, the notion of implications is a merely descriptive concept devoid of the processual component inherent in what is expressed by the concept of “conspiracy”. In this regard, the qualitative duality of diachrony *vs.* synchrony is akin to Coseriu’s distinction between static and dynamic universals of language from the perspective of “activity” (Coseriu, 1974: 56), a dichotomy which is reconciled in the notion of conspiracy. To be sure, both diachronic and synchronic aspects are equally relevant, since on the one hand, a longitudinal diachronic study can reveal correlative tendencies in a scenario of an ongoing linguistic change that, as weak associations at best, might fall short of statistical significance in a cross-sectional synchronic study at solely one time point amidst the process of change. Furthermore, an implication can so be traced to one initial change in the grammatical system that entailed one or several ensuing other associated changes. On the other hand, reliable data for synchronic (usually contemporary) linguistic analyses are in most cases available, whereas comprehensive written historical records, in the best case representative of more than one epoch, or reconstructed linguistic data, might be not. A large body of valuable data amenable to testable hypotheses can hence be generated from synchronic corpora.

For further classification, a number of feasible specifications can be assigned to the term conspiracy: Aside from the obvious attributes referring to linguistic domains (*e.g.*, syntactic, morphological, phonological, and so forth) and next to the intrinsic duality of both synchronic (equaling the notion of implication in the usual sense) and diachronic denotation, conspiracies are observed intra-, inter-, or cross-linguistically: While intra-linguistically, conspiracies are related to the diachronic development of correlated phenomena or the

synchronic distribution of certain parameter values in different syntactic contexts in only one language, the exploration of inter- and cross-linguistic conspiracies (like syntactic properties correlated with the relative position of the object to the verb, *cf.* section 5.1., or the mutual exclusion of V2 and *pro*-drop in the history Western Romance, *cf.* section 5.3.) aims at the identification of principles underlying the design of a group of languages, *e.g.* a language family, or the design of human language as a whole, as are goals of typology and universalism. In this vein, conspiracies can be universal, comprising subcategories appearing as universal statistical and universal absolute (or exceptionless) conspiracies; if confined to the synchronic dimension, in synonymy with implications of these types. Similarly, for a conspiracy to be labeled bilateral or reciprocal, each element of a pair of correlated linguistic traits thus constituting a bipartite conspiracy must be able, if subject to change, to effect the same outcome over time in mutual dependency, regardless whether this development takes place simultaneously in the two elements or not, and irrespective of which element first touched off the change. Analogously, the notion of a unilateral conspiracy is justified if out of two correlated elements, only one adheres to the other but not vice versa in synchronic studies, backed up by historical proof of that one-sided dependency. For example, notwithstanding that transitive expletive constructions and object shift require movement of the verb up to the TP in Germanic (*cf.* section 5.2.), these phenomena in turn certainly do not occur in all languages that raise the verb up to TP, as examples from and outside Germanic languages demonstrate. In extension of this, if more than two elements like certain parameter values are unravelled to cluster together displaying correlation over time, tripartite, quadripartite, quinquartite, and so forth, in a nutshell multipartite, conspiracies can be distinguished. As examples of multipartite conspiracies, properties that go along with the dropping of pronouns, mentioned in section 5.3., can be adduced, as well as several features associated with the V2 word order type in Germanic languages, as touched upon in section 3.3. Also in cases of a multipartite conspiracy, disparate implicational properties can conceivably be sorted out from each other, giving rise to unilateral, bilateral, trilateral, and so forth, or multilateral relations. For instance, a tripartite conspiracy can be bilateral if one element of the three, although correlated with the other two, occurs also in contexts devoid of either of the other two elements. Moreover, contingent upon the exact formulation, a conspiracy can be either positive (“if *x*, then also *y* and *z*”) or negative (or inverse) (“if *x*, then not *y* and *z*”). In summary, Fig. 2 illustrates the multidimensional classification entailed with the term conspiracy.

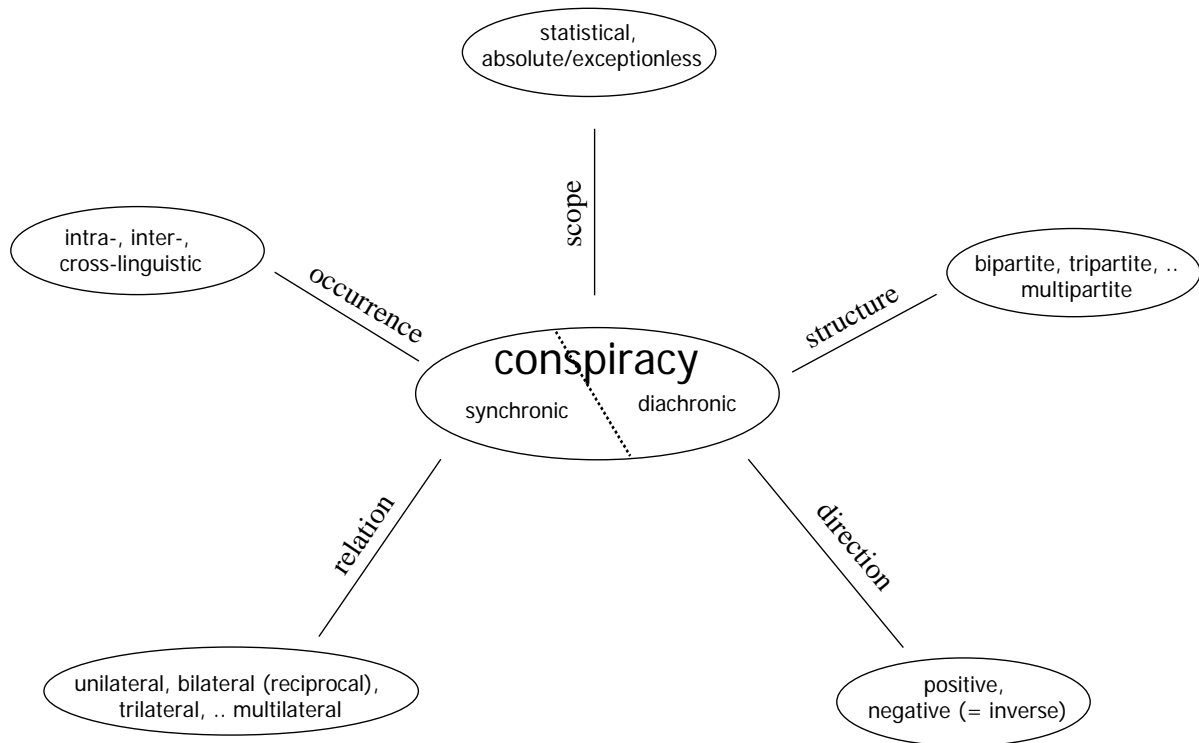


Fig. 2: Schematic representation of the concept conveyed by the term “conspiracy” and its specifications: The duality of synchrony and diachrony implied by this one term is central. Axes from the center project to the periphery along the dimensions of occurrence, scope, structure, direction, and relation, thus connecting to corresponding attributes.

6.2. Factors explaining (syntactic) conspiracies

In this subsection, possible factors underlying emergence and existence of language universals and universal implications will be summarized and will be asserted to bear relevance for conspiracies likewise. Owing to the emphasis on general phenomena of universal validity, focus will be on the cross-linguistic perspective, however, without ruling out that the same factors are underlying conspiracies that are intra-linguistically restricted.

Several explanations that can account for the existence of language universals appear in and are discussed by Comrie (1981: 22ff.) that by extension tentatively also apply to conspiracies. Comrie’s reservations to abundantly invoking innateness as plausible cause for the occurrence of universals notwithstanding, aspects of cognitive psychology that ensure ease of cognitive processing, for instance salience hierarchies in perception that are paralleled by corresponding syntactic phrase hierarchies, are rated properly plausible. Noteworthy further explanations encompass functional and pragmatic factors that serve for ease of communication like the universal presence of deictic systems in human languages. Whereas Comrie dismisses the idea that a common monogenetic origin of human language might explain universal linguistic properties and recurrent patterns of grammar that are retained in

nowaday's languages as highly speculative, in case of inter-linguistic conspiracies restricted to a group of related languages, like phenomena associated with V2 word order or with raising movement of V to T in Germanic, inheritance of correlated traits is definitely one viable explanation. This view is endorsed by a recent cross-linguistic analysis by Dunn *et al.* (2011) who, conversely, underscore the lineage-specificity of correlated linguistic traits, concomitantly denying general cognitive mechanisms as explanations for commonalities throughout languages. In addition, also language contact was suggested to play a role in the occurrence of universal linguistic features, as advocated by Moravcsik (1978).

In explaining implicational word order universals, plenty of ideas were put forth by Hawkins's account on word order universals (1983). He therein proposed that noun modifiers are placed to the right or to the left of their heads in principled ways (Hawkins, 1983: 88ff.). Such principles are based on "heaviness" and "mobility" of modifiers like relative clause, genitive, adjective, demonstrative, and numeral, with different modes of operation in prepositional and postpositional languages following the PrNMH and the PoNMH (*cf.* section 3.2.). Defining "heaviness" of noun modifiers as composite property embracing their length of morphemes and quantity of both words and morphemes, as well as the number of branching nodes and dominated constituents within the noun modifier phrase, Hawkins established a hierarchy ranging from relative clauses as the most heavy noun modifiers over genitives and adjectives, to both demonstratives and numerals as equally least heavy elements. This ordering gives rise to the Heaviness Serialization Principle (HSP), according to which heavier noun modifiers exhibit a cross-linguistic propensity to occur to the right of the noun, devoid of exceptions in the set of prepositional languages examined by Hawkins (1983). A second rule, the Mobility Principle (MP), is hence introduced to account for counter-instances to the HSP amongst postpositional languages, according to which noun modifiers are categorized by "mobility", *i.e.* their ability to move around their heads in departing from the serialization dictated by the adposition phrase as entailed by the PrNMH and the PoNMH, whereby adjectives, demonstratives, and numerals alike are equally or more mobile than relative clauses and genitives. Combination of the HSP and the MP leads to the MHIP (Mobility and Heaviness Interaction Principle) which, by assigning more weight to the MP in cases with divergent classifications of the two principles, achieves to make all predictions correct.

A further regularity pinpointed by Hawkins (1983: 133) refers to the observation that the ratio of all preposed operators to all postposed ones relative to the operand within one phrasal category remains constant for most languages over all phrasal categories. This cross-

linguistic generalization that is attributed by Hawkins to an underlying principle of Cross-Category Harmony (CCH), is demonstrated to make more reliable predictions than Vennemann's Natural Serialization Principle (NSP; *cf.* section 3.2.) that claims that the bulk of languages serialize all operators either to the right or to the left of their operands (Vennemann, 1974a,b). Departure from the cross-category harmony is accompanied by a proportional decline in language frequencies - the more violations to CHH, the fewer languages to be found with this profile. Accordingly, languages with instantiations of a given implicational universal statement show a characteristic numerical distribution pattern across all attested instantiations, with unequal occurrence frequencies. For instance, the unilateral implication stating that in postpositional languages of SOV word order type the genitive (G) precedes the noun (N) if the adjective (Adj) does likewise, allows for three possible word order instantiations of which Adj-N & G-N is most frequent, followed by N-Adj & G-N and N-Adj & N-G as the least frequent type (*cf.* Hawkins, 1983: 135).

In providing causal explanations underlying the three observed patterns of HSP, MP, and CCH, Hawkins (1983), as for the HSP, adduced psycholinguistic evidence involving language processing. Citing work by Moore (1972) who found that relations between the verb and its arguments are processed before argument-internal relations, Hawkins conjectured from these findings that likewise modifiers of the head of a noun phrase are processed later than the head nouns themselves. He went on drawing conclusions on processing difficulties associated with late appearance of the head within the noun phrase as well as with numerous and "heavy" (in the sense of the HSP) modifiers preceding the noun. By contrast, Hawkins attributed the MP to syntactic and language historical factors. In brief, non-branching constituents including adjective, demonstrative, and numeral that do not dominate noun phrases and involve little complexity are syntactically more mobile than their "more settled" noun modifying counterparts genitive and relative clause. In phases of ongoing language change, less complex noun modifiers undergo reordering before the more complex ones. In combination, mobility is explained as a grammar-based phenomenon interacting with gradualness of linguistic change (Hawkins, 1983: 109). Finally, CCH is explained as interrelatedly tri-factorial principle (Hawkins, 1983: 179-183): Syntactic-semantic relations like that of the modifier to its head but notably not the function-argument principle are able to account for decreasing occurrence frequencies of languages with less similar cross-categorical localization of the head. Next, analogy is claimed to play a role in generalization from one phrasal category to another in order to achieve a cross-categorical harmonic balance of preposed and postposed operators. And lastly, grammatical, more specifically, syntactic

complexity is invoked as third factor, since Hawkins contends that there exists an inverse relation between cross-category harmony and grammatical complexity, whereby the fewer less harmonic languages exhibit more complex grammatical structures. It is at this point noteworthy that CCH is not a principle that explains implications or conspiracies themselves, but their distribution, being nevertheless a universal principle just like the HSP and the MP.

In section 2, language acquisition, usually first language learning, was introduced and discussed as influential factor for language change. Moreover, section 4 provided evidence for linking universally correlated traits like conspired implications with language acquisition: In particular, the development of creole and sign languages were claimed to mirror commonalities and regularities of first language acquisition. But also the psycholinguistic factors restraining language processing advocated by Hawkins (1983) to underlie the HSP, bear relevance as constraints for language acquisition that contribute towards shaping possible grammars. In this vein, intriguingly, not only is first language acquisition affected by such constraints, even adult language learning is skewed towards rule-based syntactic generalizations upon complex input, as recent investigations have shown (*cf.* Hudson Kam and Newport, 2009; Reali and Griffiths, 2009).

One of such studies provided experimental evidence for the propensity of acquiring syntactic structure consistent with one of the Greenbergian implicational universals, contingent upon varying syntactic learning input conditions: By applying an artificial language learning paradigm, Culbertson *et al.* (2012) demonstrate that adults exposed to several simultaneous grammatical systems, mimicking a state of ongoing diachronic change, develop a word order preference described by Greenberg's universal 18, *i.e.* to place the demonstrative and the numeral before the noun if the adjective likewise precedes the noun. In their study, 65 native speakers of English who in part were competent also in other natural languages, were taught vocabulary for several nouns, adjectives, and numerals in an artificial language. For acquiring syntax, participants were subsequently split into five groups of learning conditions that differed by the major input pattern of possible word orders of nouns relative to adjectives and nouns in relation to numerals, allowing for one control condition. Whereas each of the four possible word ordering instances was presented to the participants with equal frequency in the control condition, in the other four groups the dominant instances of the adjective position and the numeral position relative to the noun were presented to the participants with a majority frequency of 70% each. These four experimental conditions with unequal word order distribution corresponded on the one hand to the two most common cross-linguistic types according to Greenberg (1966) and Dryer and Haspelmath (2011) thus labeled

“harmonic” conditions, *i.e.* noun preceding or following both adjective and numeral, and on the other corresponded to two “non-harmonic” word order types, *i.e.* a less common “unmarked” (numeral preceding noun and followed by adjective) as well as a rare “marked” type (adjective preceding noun and followed by numeral). When the study subjects were required to describe pictures by using the artificial language, it turned out that they preferred to use the syntactic variant they had been predominantly (*i.e.*, by 70%) exposed to during the training phase only if they had not been part of the learning condition representing the marked type. In detail, results showed that the study participants used expressions following the word order they had been trained on significantly more often than in 70% of cases in the two harmonic groups and the unmarked group. Participants trained in the control and marked conditions, by contrast, displayed no significant deviation from the input frequencies, even though there was a trend for less use in the group that had been exposed to the marked word order type. Further comparisons revealed that the two harmonic groups did not differ significantly in (elevated) use of the predominant word order of the training phase. Although in the unmarked group use frequency increased likewise compared with the training set, this increment was significantly less than that in the two harmonic groups, due to less frequent use of numerals in the “correct” word order. As for the marked condition, statistical significance was observed relative to all other conditions, whereby only the less frequently “correct” use of numerals accounted for the difference to the unmarked condition. For a deeper insight into the learners’ behavior, a biplot was constructed showing each participant’s proportions (or probabilities) for numeral and adjective positions relative to the noun, revealing varying shifting tendencies contingent upon learning conditions (Tab. 2):

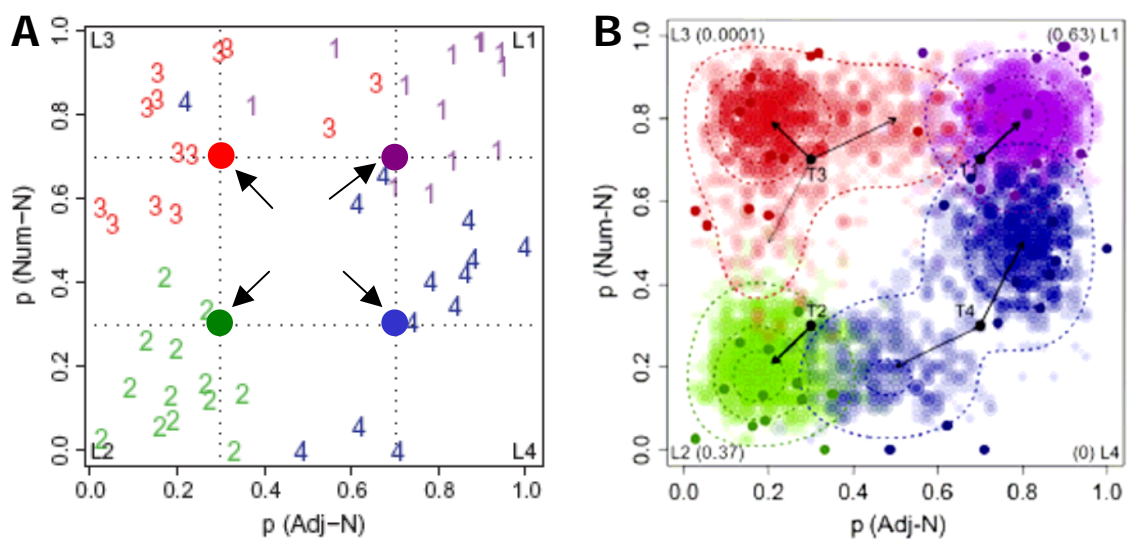


Fig. 3: Biplots showing the distribution of syntactic word order proportions in the utterances of each participant following any of the four training conditions (conditions L1, L2, L3, and L4 correspond to

violet, green, red, and blue color, respectively). Input training conditions are indicated by colored dots pointed at by black arrows in (A). Each learner is marked by a colored number indicative of his/her training condition, and position in the biplot corresponds to language production output. Results of a Bayesian model that assumes learning biases as discussed in the text are represented in (B). Numbers in brackets assign weight to the learning condition next to them in each corner. T1-T4 refer to initial training conditions, with arrows indicating the directions of learners' biases (after Culbertson *et al.*, 2012).

As Fig. 3 shows, starting from the initial training conditions that exposed the participants to input training utterances at a ratio of 7:3 for each noun modifier position relative to the noun, pointed at by arrows, learners' own utterances exhibited different patterns across learning groups. Whereas participants acquiring the two harmonic conditions L1 (both modifiers, adjective and numeral, precede the noun, corresponding to English) and L2 (both modifiers follow the noun) tended to increase the word order ratio in their own utterances, reflected by accumulation of data points of the corresponding learning condition in the upper right and lower left corner, respectively, of Figs. 3A and 3B, learners having acquired the non-harmonic unmarked word order labeled L3 (*i.e.*, numeral preceding the noun, adjective following it), although represented fairly firmly in the upper left square, displayed a tendency for moving away towards one or the other harmonic word order. In stark contrast, none of the learners in the marked L4 group was found to increase the word order predominantly offered by the training condition (*i.e.*, adjective preceding the noun, numeral following it), witnessed by the virtual emptiness of the lower right square. Participants of this learning group in their utterances tended towards one of the harmonic word order patterns, in particular the English type L1. Importantly, experimental results are borne out by a Bayesian model accounting for the learners' biases (Fig. 3B). One important possible limitation of the study is that all participants had English as first language, which might account for the observed preferences. However, this potential concern is dispelled by the authors resorting to arguments like the absence of statistically significant difference between preferences for the two harmonic conditions, one of which followed the English type, and the observation that there was no more than chance use of English type word order in the control condition. However, as admitted by the authors, that participants in the two non-harmonic groups preferentially shifted towards the harmonic English word order type is unexpected in view of cross-linguistic preference for the other harmonic type and might be attributed to an effect of native language.

In summary, this experimental study suggests that next to a regularization bias for heterogeneous linguistic input, there exists also a harmonic learning bias in adults leading to generalizable preferences for certain implicational syntactic structures over others, claimed to

be mirrored by typological patterns and their cross-linguistic frequencies. Findings therein are therefore direct evidence that cognitive underpinnings pertaining to acquisition of grammatical structure account for the principle of cross-category harmony (CCH) as described by Hawkins (1983), rather than implying as factors the purely intra-grammatical triad of syntacto-semantic relations, analogy, and grammatical complexity (see above). Experimental evidence thus advocates biases in, notably also adult, language learning, reflecting constraints imposed by universal grammar (UG), whose possible results are probabilistically distributed and become manifest in typological patterns whose distribution is inclined towards the more harmonic types. To be sure, drawing a relation between UG, language learning, and typology seems warranted in light of one possible interpretation of UG as prior learning bias resulting in constraints on cross-linguistic variation (Kirby *et al.*, 2004).

6.3. The role of language acquisition and conspiracies in diachrony

Experimental learning experiments that demonstrate the rise of regular linguistic structure from conditions without or less grammatical regularities, including the conspired emergence of syntactic combinations that are favored over other such combinations, crucially all imply the notion of language change that heads towards a certain direction (*e.g.*, Reali and Griffiths, 2009; Culbertson *et al.*, 2012). This idea of directed diachronic change is invariably linked with Edward Sapir's concept of linguistic drift that underlies the continual emergence of divergent dialectic variations resulting in the decline and rise of languages observed in the course of time (Sapir, 1921: 160 ff.). In this view, Vennemann (1975: 286 ff.) notes a close relation between Sapir's diachronic concept of drift and tendencies in the world's languages that are correlated, manifested by implicational linguistic universals described by Greenberg (1966), by the structural word order-related principle on the basic arrangement of verb and object identified by Lehmann (1973), and by the Natural Serialization Principle (NSP) (Vennemann, 1974a,b), all of which capture first and foremost the synchronic situation. Vennemann (1975) hence implicitly bridges the two dimensions, *i.e.* synchronic and diachronic, claimed herein to be harbored by the notion of conspiracies. Along the same lines, it is certainly justified to adduce as further explanation for the notion of drift the principle of CCH, epitomizing a general cross-linguistic tendency for harmony and typological consistency, presumably guided by cognitive principles manifest in language learning. Or, as Roberts (2007: 197) puts it for word order change in the history of English, the emergence of a new typological pattern such as VO from OV word order involves the interaction of several parameters that tend to act harmonically. This process of interaction might furthermore be

equated with a propensity for independent parameters to conspire (hence, to form a conspiracy) to give rise to a certain grammar type (Roberts, 2007: 101-102), during the period of acquiring a language.

If there is a general tendency, mediated by (first) language learning, that all grammars strive for harmonic ordering reflected by universal conspiracies, it is plausible to ask what mechanisms create diversity and why linguistic systems, having reached a state of sufficient harmony, change at all. Several points might serve as viable answers to such a question: First of all, the principle of CCH does allow of disharmonic orderings, it thereby predicts that these are less likely to occur. Likewise, models of artificial learning (*e.g.*, Culbertson *et al.*, 2012) account for this distributional principle insofar as their outcomes are probabilistic. Then, it should be borne in mind that, as touched upon in section 2.1. as well as implied by the model of Yang (2000), it is also through mechanisms other than language acquisition that languages undergo change, *e.g.* language contact. In this regard, stable grammatical systems like those reached by transmission over many generations in iterated learning models of artificial languages (*cf.* Kirby *et al.*, 2004) might lose equilibrium triggered by the effects of language contact, inducing yet another process of acquisition-mediated development towards a stable outcome quite different from the previous one, characterized by distinct grammatical structure and other conspired grammatical traits. As discussed in section 3.3., several ideas have been put forward to parallel such developmental routes towards a stable state, often termed transitional, with typologies inconsistent in their correlated phenomena (*e.g.*, Vennemann, 1974b; Croft, 1990; Greenberg, 1995). Conceptually, however, an alternative explanation is conceivable if it is assumed that confounding factors possibly impede or reduce the “attraction” of elements for each other to appear as conspired traits, conferring stability to such a grammatical system with concomitantly no need to claim it in a transitional state. Such elements that block the conspiring of certain syntactic traits would not necessarily have to pertain to the realm of syntax, *e.g.* certain morphological properties or elements that could prevent the grammatical system from a “harmonic” arrangement of modifiers relative to the noun phrase. This, being tentatively a source of linguistic diversity, would explain the occurrence of languages of inconsistent outliers with respect to typological classifications. To corroborate such assumptions, diachronic investigations that take advantage of multivariate statistical techniques would be expected to yield insightful results.

If the process of language acquisition, as argued herein, constitutes a driving force for diachronic change, it is worthwhile to ask whether individual mismatch learning, evidenced as imperfections for instance in abductive learning (*cf.* Andersen, 1973), is sufficient for

diachronic changes to become manifest, as suggested by some approaches to linguistic change involving learning cited in section 2.1., *e.g.* Kiparsky (1974) and Parker (1976). Such a mechanism is also implied by the models of Clark and Roberts (1993) as well as Niyogi and Berwick (1995, 1997, 1998) (*cf.* section 2.2.), but denied by Yang (2000). Reconciling disparate stances on this issue, already Sapir, in his explanation of drift as principle of directed diachronic change, acknowledged the role of both the individual speaker and the group in the process of dialectic split (Sapir, 1921; chapter 7): While individual variation, whatever its cause, is not conceded sufficient power for the emergence of dialects, “unconscious selection” by the speakers of these variations, when accumulating as part of a directed movement, do underly the forces that drive diachronic development. Notably, this “unconscious selection”, even though not explicitly stated by Sapir, might be regarded as a metaphor for processes involved in language learning.³² In this vein, drawing on computational machine learning models of language evolution, the inadequacy of mere individual mislearning is captured by the principle of collective dynamics as a pre-requisite for language evolution and diachronic change to take place, supplementing an equally important inductive bias on part of the individual learners (Zuidema and Westermann, 2001). On an explanatory level, what makes a deviation from the preceding generation’s grammar on individual level become a collective phenomenon seen as diachronic alteration and another become leveled out by the majority grammar to eventually vanish is still poorly understood, being an issue deemed due to multiple factors (*cf.* Beckner *et al.*, 2009: 14ff.): For example, frequency effects, speech production economy, and social influences, to name a few of such factors, might be operative in learners to varying degrees and not at the same time, with early stages of language change certainly differing from later stages in these respects. Computational models, arguably in particular such that incorporate multivariate statistical methodology, as well as computer simulations might be especially well-suited for uncovering factors that enable acquisition-mediated linguistic inconsistencies with respect to the previous generation to successfully effect a sustainable change in the grammatical system (for review, *cf.* Christiansen and Chater, 2008).

6.4. Concluding remarks

The present work has attempted, with due focus on syntax, to provide a synthesis of

³² “Now dialects arise not because of the mere fact of individual variation but because two or more groups of individuals have become sufficiently disconnected to drift apart, or independently, instead of together. [...] The drift of language is constituted by the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction.” (Sapir, 1921; chapter 7)

several intensively investigated linguistic fields encompassing diachronic change, language acquisition, linguistic universals, typology and grammatical structure, universal grammar, and their possible influences on correlated phenomena in the guise of syntactic conspiracies. The relations between these notions have been demonstrated to manifest themselves as follows: Emergence of conspired linguistic traits, in particular if observed recurrently in cross-linguistic contexts, has been argued to be furthered by mechanisms of language learning, most notably the setting of grammatical parameters, which are constrained by an inborn endowment of universal grammar common to all human beings. Given such a constraint on language acquisition, the process of learning is a key factor in the emergence of not only unrestricted but also implicational linguistic universals as well as of grammatical structure of distinct typologies reflected by conspiracies. In the course of development or adaptation of linguistic structure, repeated learning over generations is witnessed as diachronic change arguably following mechanisms captured by models outlined in section 2.2. Even though being not the only factor eliciting linguistic change, language learning is considered to make an essential contribution to the forces by which diachronic development is driven. Besides regularities of language acquisition, also other factors are definitely involved in the emergence of linguistic structure or serve to explain why expected regularities do not form, including language contact, inheritance of common traits amongst related languages, and historical factors (*cf.* section 5.2.). In this regard, there exists a tight relationship between the concepts of language acquisition, diachronic change, and conspiracies, for they all harbor the intrinsic aspect of dynamic development.

7. References

- Abels, K., 2011. Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. *Groninger Arbeiten zur Germanistischen Linguistik* 53.2, 1-28.
- Adams, M., 1987. From Old French to the Theory of Pro-Drop. *Nat. Lang. Linguist. Theory* 5, 1-32.
- Aitchison, J., 1989. Spaghetti junctions and recurrent routes: some preferred pathways in language evolution. *Lingua* 77, 151-171.
- Aitchison, J., 1996. *The Seeds of Speech. Language Origin and Evolution*. Cambridge University Press, Cambridge, U.K.
- Andersen, H., 1973. Abductive and deductive change. *Language* 49, 765-793.
- Bailey, C.-J.N., 1973. *Variation and linguistic theory*. Center for Applied Linguistics, Arlington, Virginia.
- Barnes, M.P., 2008 [1999]. *A New Introduction to Old Norse, part I: Grammar*. 3rd ed., Viking Society for Northern Research, University College London, London.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M.H., Croft, W., Ellis, N.C., Holland, J., Ke, J., Larsen-Freeman, D., Schoenemann, T., 2009. Language Is a Complex Adaptive System: Position Paper. *Lang. Learn.* 59 suppl. 1, 1-26.
- Bennett, P.A., 1979. On Universal 23. *Linguist. Inq.* 10, 510-511.
- Berwick, R.C., 1985. *The Acquisition of Syntactic Knowledge*. MIT Press, Cambridge, Massachusetts.
- Bickerton, D., 1984. The language bioprogram hypothesis. *Behav. Brain Sci.* 7, 173-203.
- Bickerton, D., 1999. How to acquire language without positive evidence: What acquisitionists can learn from Creoles. In: M. DeGraff (ed.), *Language Creation and Language Change*. MIT Press, Cambridge, Massachusetts, pp. 49-74.
- Bloomfield, L., 1933. *Language*. Henry Holt, New York.
- Bobaljik, J.D., Jonas, D., 1996. Subject positions and the roles of TP. *Ling. Inq.* 27, 195-236.
- Bures, A., 1992. (Re)-cycling expletives and other sentences. General paper, MIT, Cambridge, Massachusetts.
- Bush, R., Mosteller, F., 1958. *Stochastic models for learning*. Wiley, New York.
- Campbell, L., 2004 [1998]. *Historical Linguistics. An Introduction*. 2nd ed., Edinburgh University Press, Edinburgh.
- Cerquiglini, B., 2007 [1991]. *La Naissance du français*. 3rd ed., Presses Universitaires de France, Paris.
- Chomsky, N., 1959. A review of B. F. Skinner's *Verbal Behavior*. *Language* 35, 26-58.
- Chomsky, N., 1965. *Aspects of the theory of syntax*. MIT Press, Cambridge, Massachusetts.
- Chomsky, N., 1981. *Lectures on government and binding*. Foris Publications Holland, Dordrecht.
- Chomsky, N., 1986a. *Barriers*. MIT Press, Cambridge, Massachusetts.
- Chomsky, N., 1986b. *Knowledge of Language: its Nature, Origin and Use*. Praeger Publishers, New York.
- Chomsky, N., 1995. *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.
- Chomsky, N., Lasnik, H., 1993. The Theory of Principles and Parameters. In: J. Jacobs, von Stechow, A., Sternefeld, W., Vennemann, T. (eds.), *Syntax: An International Handbook of Contemporary Research*. Mouton de Gruyter, Berlin, pp. 506-569.
- Christiansen, M.H., Chater, N., 2008. Language as shaped by the brain. *Behav. Brain Sci.* 31, 489-558.

- Cinque, G., 1996. The antisymmetric programme: Theoretical and typological implications. *J. Linguistics* 32, 447–464.
- Cinque, G., 2000. On Greenberg's Universal 20 and the Semitic DP. In: L. Brugè (ed.), *University of Venice working papers in linguistics*, vol. 10(2). University of Venice, Dipartimento di Scienze del Linguaggio, Venice, pp. 45-61.
- Cinque, G., 2005. Deriving Greenberg's universal 20 and its exceptions. *Ling. Inq.* 36, 315-332.
- Cinque, G., 2009. Greenberg's Universal 23 and SVO languages. In: L. Brugè (ed.), *University of Venice working papers in linguistics*, vol. 19. University of Venice, Dipartimento di Scienze del Linguaggio, Venice, pp. 29-34.
- Clark, R., Roberts, I., 1993. A computational model of language learnability and language change. *Linguist. Inq.* 24, 299-345.
- Comrie, B., 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. University of Chicago Press, Chicago & Oxford.
- Coseriu, E., 1974. Les universaux linguistiques (et les autres). In: L. Heilmann (ed.), *Proceedings of the Eleventh International Congress of Linguists, I, Bologna-Florence, Aug. 28 – Sept. 2, 1972*, pp. 47-73.
- Croft, W., 1990. *Typology and Universals*. Cambridge University Press, Cambridge.
- Culbertson, J., Smolensky, P., Legendre, G., 2012. Learning biases predict a word order universal. *Cognition* 122, 306-329.
- Darwin, C., 1859. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.
- Darwin, C., 1871. *The descent of man and selection in relation to sex*. John Murray, London, U.K.
- Darwin, C., Wallace, A.R., 1858. On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *J. Proc. Linn. Soc. Zool.* 3, 45-62.
- de Saussure, F., 1916. *Cours de linguistique générale* (C. Bally, A. Séchehaye, eds.). Payot, Lausanne/Paris.
- Dressler, W.U., 1977. Elements of a polycentristic theory of word-formation. *Wiener Linguistische Gazette* 15, 13-32.
- Dryer, M.S., 1992. The Greenbergian word order correlations. *Language* 68, 81-138.
- Dryer, M.S., 2009. The order of demonstrative, numeral, adjective, and noun: an alternative to Cinque. In: *Conference on Theoretical Approaches to Disharmonic Word Orders, May 30th - June 1st 2009*, Newcastle.
- Dryer, M.S., Haspelmath, M. (eds.), 2011. *The World Atlas of Language Structures Online*. Max Planck Digital Library, Munich.
- Dunn, M., Greenhill, S.J., Levinson, S.C., Gray, R.D., 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79-82.
- Evans, N., Levinson, S.C., 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behav. Brain Sci.* 32, 429-492.
- Finck, F.N., 1899. *Der deutsche Sprachbau als Ausdruck deutscher Weltanschauung: acht Vorträge*. N.G. Elwert, Marburg.
- Fischer, O., Los, B., van Kemenade, A., van der Wurff, W., 2000. *The Syntax of Early English*. Cambridge University Press, Cambridge.
- Fitch, W.T., 2011. Unity and diversity in human language. *Phil. Trans. R. Soc. B* 366, 376-388.
- Gibson, E., Wexler, K., 1994. Triggers. *Linguist. Inq.* 25, 407-454.

- Graebner, F., 1911. *Methode der Ethnologie*. Carl Winter, Heidelberg.
- Greenberg, J.H., 1966 [1963]. Some universals of grammar with particular reference to the order of meaningful elements. In: J.H. Greenberg (ed.), *Universals of language*. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961, 2nd ed. MIT Press, Cambridge, Massachusetts, pp. 73-113.
- Greenberg, J.H., 1995. The Diachronic Typological Approach to Language. In: M. Shibatani, Bynon, T. (eds.), *Approaches to Language Typology*. Oxford University Press, Oxford, pp. 145-166.
- Griffiths, T.L., Kalish, M.L., 2007. Language evolution by iterated learning with bayesian agents. *Cogn. Sci.* 31, 441-480.
- Harris, M., 1978. *The Evolution of French Syntax: A Comparative Approach*. Longman, London.
- Haspelmath, M., 1997. *Indefinite Pronouns*. Oxford University Press, Oxford.
- Haspelmath, M., 2008. Parametric versus functional explanations of syntactic universals. In: T. Biberauer (ed.), *The limits of syntactic variation*. Benjamins, Amsterdam, pp. 75-107.
- Haspelmath, M., Dryer, M.S., Gil, D., Comrie, B., 2005. *The World Atlas of Language Structures*. Oxford University Press, Oxford.
- Hawkins, J.A., 1983. *Word Order Universals*. Academic Press, New York.
- Hockett, C.F., 1960. The origin of speech. *Sci. Am.* 203, 89-96.
- Hockett, C.F., 1966 [1963]. The problem of universals in language. In: J.H. Greenberg (ed.), *Universals of language*. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961, 2nd ed. MIT Press, Cambridge, Massachusetts, pp. 1-29.
- Holm, J.A., 2000. *An introduction to pidgin and creoles*. Cambridge University Press, Cambridge, U.K.
- Holmberg, A., 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53, 1-39.
- Hróarsdóttir, T., 2000. *Word Order Change in Icelandic: From OV to VO*. John Benjamins Publishing, Amsterdam.
- Hudson Kam, C.L., Newport, E.L., 2009. Getting it right by getting it wrong: When learners change languages. *Cogn. Psychol.* 59, 30-66.
- Humboldt, W. von, 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. F. Dümmler, Berlin.
- Husserl, E., 1913. *Logische Untersuchungen: Band 2, Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*, 1. Teil. Max Niemeyer, Halle/Saale.
- Jackendoff, R., 2002. *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford University Press Inc., New York.
- Jakobson, R., 1972 [1968]. *Child Language, Aphasia and Phonological Universals*. 2nd ed., Mouton, The Hague (translation by A.R. Keiler of monograph from 1941: *Kindersprache, Aphasie und allgemeine Lautgesetze*, Almqvist & Wiksell, Uppsala).
- Jonas, D., 1996. *Clause Structure and Verb Syntax in Scandinavian and English*. PhD Dissertation, Harvard University.
- Kalish, M.L., Griffiths, T.L., Lewandowsky, S., 2007. Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonom. Bull. Rev.* 14, 288-294.
- Kayne, R., Pollock, J.-Y., 2001. New Thoughts on Stylistic Inversion. In: A. Hulk and Pollock, J.-Y. (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*. Oxford University Press, Oxford, pp. 107-162.
- Kiparsky, P., 1974. Remarks on analogical change. In: J.M. Anderson and Jones, C. (eds.), *Historical Linguistics II*, North-Holland Linguistic Series 12b, Amsterdam, pp. 257-276.

- Kirby, S., Hurford, J.R., 2002. The emergence of linguistic structure: an overview of the iterated learning model. In: A. Cangelosi, Parisi, D. (eds.), *Simulating the evolution of language*. Springer, Berlin, pp. 121–148.
- Kirby, S., Smith, K., Brighton, H., 2004. From UG to Universals. *Linguistic adaptation through iterated learning*. *Stud. Lang.* 28, 587–607.
- Kroch, A., 1989. Reflexes of grammar in patterns of language change. *Lang. Var. Change* 1, 199–244.
- Kroeber, A.L., 1939. *Cultural and Natural Areas of Native North America*. University of California Press, Berkeley.
- Labov, W., 1994. *Principles of Linguistic Change: Internal Factors*. Blackwell, Oxford.
- Lehmann, W., 1973. A structural principle of language and its implications. *Language* 49, 47–66.
- Lehmann, W., 1993. *Theoretical Bases of Indo-European Linguistics*. Routledge, London.
- Lightfoot, D., 1979. *Principles of Diachronic Syntax*. Cambridge University Press, Cambridge.
- Lightfoot, D., 1991. *How to set parameters: Arguments from language change*. MIT Press, Cambridge, Massachusetts.
- MacSwan, J., 2000. The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Lang. Cogn.* 3, 37–54.
- Mallinson G., Blake, B.J., 1981. *Language Typology: Cross-linguistic Studies in Syntax*. North-Holland Publishing Company, Amsterdam.
- Moore, T.E., 1972. Speeded recognition of ungrammaticality. *J. Verbal Learning Verbal Behav.* 11, 550–560.
- Moravcsik, E.A., 1978. Language contact. In: J.H. Greenberg, Ferguson, C.A., Moravcsik, E.A. (eds.), *Universals of human language*, vol. 1 (Method and Theory). Stanford University Press, Stanford, California, pp. 93–122.
- Morgan, L.H., 1877. *Ancient Society*. C.H. Kerr, Chicago.
- Muysken, P., 1988. Are creoles a special type of language? In: F. Newmayer (ed.), *Linguistics: The Cambridge Survey*, vol. 2. Cambridge University Press, Cambridge, pp. 285–301.
- Niyogi, P., Berwick, R.C., 1995. The logical problem of language change. MIT Artificial Intelligence Laboratory Memo No. 1516, Cambridge, Massachusetts.
- Niyogi, P., Berwick, R.C., 1997. Evolutionary consequences of language learning. *Linguist. Philos.* 20, 697–719.
- Niyogi, P., Berwick, R.C., 1998. The logical problem of language change: a case study of European Portuguese. *Syntax* 1, 192–205.
- Nowak, M.A., Komarova, N.L., Niyogi, P., 2002. Computational and evolutionary aspects of language. *Nature* 417, 611–617.
- Osgood, M., 1966 [1963]. Language universals and psycholinguistics. In: J.H. Greenberg (ed.), *Universals of language*. Report of a conference held at Dobbs Ferry, New York, April 13–15, 1961, 2nd ed. MIT Press, Cambridge, Massachusetts, pp. 299–322.
- Parker, F., 1976. Language change and the passive voice. *Language* 52, 449–460.
- Patel, A.D., 2003. Language, music, syntax, and the brain. *Nat. Neurosci.* 6, 674–681.
- Paul, H., 1886 [1880]. *Prinzipien der Sprachgeschichte*. 2nd ed., Max Niemeyer, Halle/Saale.
- Pintzuk, S., 1991. *Phrase Structure in Competition: Variation and Change in Old English Word Order*. PhD Dissertation, University of Pennsylvania.
- Pintzuk, S., 1999. *Phrase Structure in Competition. Variation and Change in Old English Word Order*. Garland Publishing, New York.

- Pintzuk, S., 2002. Verb-object Order in Old English: Variation as Grammatical Competition. In: D.W. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford University Press, Oxford, pp. 276–300.
- Platzack, C., 1990. Fred Weerman: The V2 Conspiracy. A Synchronic and a Diachronic Analysis. Dordrecht: Foris Publications. 1989. *Nordic Journal of Linguistics* 13, 77-88.
- Real, F., Griffiths, T.L., 2009. The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition* 111, 317-328.
- Real, F., Griffiths, T.L., 2010. Words as alleles: connecting language evolution with Bayesian learners to models of genetic drift. *Proc. R. Soc. B* 277, 429-436.
- Reilly, F.E., 1970. Charles Peirce's theory of scientific method. Fordham University Press, New York.
- Rizzi, L., 1982. *Issues in Italian Syntax*. Foris Publications, Dordrecht.
- Roberts, I.G., 1993. *Verbs and diachronic syntax: A comparative history of English and French*. Kluwer Academic Publishers, Dordrecht.
- Roberts, I., 1995. Object Movement and Verb Movement in Early Modern English. In: H. Haider, Olsen, S., Vikner, S. (eds.), *Studies in Comparative Germanic Syntax*. Kluwer, Dordrecht, pp. 269-284.
- Roberts, I., 1999. Verb Movement and Markedness. In: M. DeGraff (ed.), *Language Creation and Language Change*. MIT Press, Cambridge, Massachusetts, pp. 287-328.
- Roberts, I., 2007. *Diachronic Syntax*. Oxford University Press, Oxford.
- Sandler, W., Meir, I., Padden, C., Aronoff, M., 2005. The emergence of grammar: Systematic structure in a new language. *Proc. Natl. Acad. Sci. USA* 102, 2661-2665.
- Sankoff, G., Laberge, S., 1973. On the acquisition of speakers by a native language. *Kivung* 6, 32-47.
- Sapir, E., 1921. *Language. An Introduction to the Study of Speech*. Harcourt, Brace & Company, New York.
- Sapir, E., 1949. *Selected Writings of Edward Sapir in Language, Culture, and Personality* (D.G. Mandelbaum, ed.). University of California Press, Berkeley.
- Seiler, H., 1995. Cognitive-Conceptual Structure and Linguistic Encoding: Language Universals and Typology in the UNITYP Framework. In: M. Shibatani, Bynon, T. (eds.), *Approaches to Language Typology*. Oxford University Press, Oxford, pp. 273-325.
- Senghas, A., Kita, S., Özyürek, A., 2004. Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305, 1779-1782.
- Song, J.J., 2001. *Linguistic Typology: Morphology and Syntax*. Longman Linguistic Library, London.
- Stocking, G.W., 1974. *A Franz Boas reader: The shaping of American anthropology, 1883-1911*. University of Chicago Press, Chicago.
- Traugott, E.C., 1972. *A History of English Syntax*. Holt, Rinehart & Winston, New York.
- van Kemenade, A., 2007. Formal syntax and language change. *Developments and outlook*. *Diachronica* 24, 155-169.
- Vance, B.S., 1997. *Syntactic Change in Medieval French: Verb Second and Null Subjects*. Kluwer Academic Publishers, Dordrecht.
- Vanelli, L., Renzi, L., Benincà, P., 1986. *Typologie des pronoms sujets dans les langues romanes*. Actes du XIIe Congrès de Linguistique et Philologie Romanes, Aix-en-Provence.
- Vennemann, T., 1975. An Explanation of Drift. In: C.N. Li (ed.), *Word Order and Word Order Change*. University of Texas Press, Austin, pp. 269-305.
- Vennemann, T., 1974a. Theoretical word order studies: results and problems. *Papier zur Linguistik* 7, 5-25.

- Vennemann, T., 1974b. Topics, subjects, and word order: from SXV to SVX via TVX. In: J. Anderson and Jones, C. (eds.), *Historical Linguistics: Proceedings of the First International Congress of Historical Linguistics*, Edinburgh, September 1973, vol. II. North-Holland, Amsterdam, pp. 339-376.
- Vincent, N., 1988. Latin. In: M. Harris, Vincent, N. (eds.), *The Romance Languages*. Routledge, London, pp. 26–78.
- Weerman, F., 1989. *The V2 Conspiracy. A Synchronic and a Diachronic Analysis*. Foris Publications, Dordrecht.
- Whorf, B.L., 1956. *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J.B. Carroll, ed.). MIT Press, Cambridge, Massachusetts.
- Yang, C.D., 2000. Internal and external forces in language change. *Lang. Var. Change* 12, 231-250.
- Zaenen, A., Maling, J., Thráinsson, H., 1990. Case and grammatical functions: the Icelandic passive. In: J. Maling, Zaenen, A. (eds.), *Modern Icelandic Syntax, Syntax and Semantics* vol. 24. Academic Press, New York.
- Zuidema, W.H., Westermann, G., 2001. *On the Relevance of Language Evolution Models for Cognitive Science*. Technical report, AI-Lab, VUB Brussels.

Appendix I: Deutsche Zusammenfassung (German abstract)

Empirische, zumeist synchrone linguistische Arbeiten haben immer wieder über ein in den Sprachen der Welt gemeinsames oder sich gegenseitig ausschließendes Auftreten bestimmter sprachlicher Elemente, auch im Bereich der Syntax, berichtet. Die Beschreibung von überwiegend als Implikationen formulierten Sprachuniversalien durch Joseph H. Greenberg in den 1960er Jahren übte beträchtlichen Einfluss auf die Sprachtypologie aus und gilt als eine der einflussreichsten Publikationen dieser Art für die Syntax. Die vorliegende Arbeit betrachtet als eine hauptsächliche Erklärung für das wiederkehrende gemeinsame Auftreten syntaktischer Eigenschaften Regelmäßigkeiten im Spracherwerb, der den diachronen Sprachwandel formt. In dieser Hinsicht werden die wichtigsten quantitativen statistischen Computermodelle vorgestellt und diskutiert, die sowohl Spracherwerb als auch Sprachwandel integrieren. Des weiteren, gemäß der großen Bedeutung der Untersuchung diachronen Wandels, wird eine Definition und Klassifikation von „Konspirationen“ als Assoziationen linguistischer Eigenschaften nicht nur durch Demonstration ihrer Affinität (oder Unvereinbarkeit) mit anderen solchen Eigenschaften auf synchroner Ebene, sondern auch in ihrer diachronen Entwicklung vorgestellt. Einige synchrone und diachrone Beispiele, hauptsächlich aus Indogermanischen Sprachen, sollen das Konzept des Konspirations-Begriffes veranschaulichen, darunter syntaktische Korrelationen mit entweder VO oder OV Wortstellung, transitive expletive Konstruktionen und Objekt-Verschiebung, und die Unvereinbarkeit von V2 Wortstellung und dem Auftreten von Null-Subjekten als Fall einer negativen Konspiration. Dabei unterstützen eigene statistische Analysen die dargestellten Ergebnisse. Des weiteren unterstreicht eine detaillierte Diskussion wahrscheinlicher und denkbarer kausaler Ursachen für das Entstehen von Konspirationen die Rolle kognitions-psychologischer Faktoren, die während des Spracherwerbs wirksam sind. Eine experimentelle Studie über Spracherwerb einer künstlichen Sprache bei Erwachsenen, deren Ergebnisse Vorhersagen einer von Greenberg formulierten universellen Implikation als auch einer prinzipiellen Tendenz zur Harmonie syntaktischer Kategorien (welche dem Konzept der „Drift“, das von Edward Sapir in den 1920er Jahren beschrieben wurde, als zugrundeliegend diskutiert wird) darstellen, verdeutlicht die Wichtigkeit kognitions-psychologischer Faktoren beim Spracherwerb für die Entstehung syntaktischer Struktur. Aber auch andere Ursachen für diachronen Sprachwandel und die Entstehung oder Abwesenheit (erwarteter) Konspirationen werden diskutiert, darunter Sprachkontakt, Tradierung gemeinsamer Eigenschaften bei Sprachverwandtschaft, und sozio-kulturelle Faktoren; ebenso wird die Wichtigkeit kollektiver Dynamik für den Sprachwandel erläutert.

Appendix II: Scientific CV

***Curriculum vitae* Mag. Dr. Wolfgang Brozek (MSc, PhD)**

Date and place of birth: December 23, 1975, Vienna, Austria

Nationality: Austrian

Post-graduate professional experience:

July 2011 - present: postdoctoral scientist at the “Ludwig Boltzmann Institute of Osteology” (Vienna, Austria), headed by Prof. Klaus Klaushofer, MD, group of Eleftherios P. Paschalis, PhD.

Aug. 2008 - June 2011: research and development for private company Synmed Research GmbH (Vienna, Austria)

Feb. 2006 - June 2008: research associate at the “Ludwig Boltzmann Institut für klinisch-experimentelle Erforschung der Prostataerkrankungen, Cluster Urologie” (Vienna, Austria) headed by Prof. Gero Kramer, MD.

Studies:

Oct. 1994 - Feb. 1996: Food Technology and Biotechnology at the University of Natural Resources and Life Sciences, Vienna (Universität für Bodenkultur, Wien) (unfinished)

March 1996 - May 2008: Biology at Vienna University

Oct. 1998 - April 2002: specialization in Ecology; Diploma Thesis: “Secondary metabolite profiling of the genus *Aglaia* (Meliaceae)” at the Institute of Botany, Department of Comparative and Ecological Phytochemistry, under the supervision of Prof. Harald Greger, PhD.

April 2002: Master’s Degree (MSc / Mag. rer. nat.) in Biology with distinction

July 2002 - May 2008: specialization in Microbiology/Genetics, Doctoral Thesis: “Roles of interleukin-6 and vitamin D₃ in modulation of cell growth and immunity in human colorectal cancer (CRC) and inflammatory bowel disease (IBD)” at the Medical University of Vienna, Department of Pathophysiology, under the supervision of Prof. Meinrad Peterlik, PhD, MD.

May 2008: Doctoral Degree (PhD / Dr. rer. nat.) in Biology with distinction

Oct. 2004 - : Linguistics at Vienna University

June 2006 - : specialization in General Linguistics, core area: syntax and grammar theory; Diploma Thesis: “Syntactic Conspiracies in a Diachronic Perspective” supervised by Prof. Hans Martin Prinzhorn, MA, PhD.

Further studies (unfinished): Translation Sciences (English, Spanish), Cultural and Social Anthropology (both Vienna University)

Education:

1982 - 1994: Elementary school and high school: Albertus Magnus Schule der Marianisten, Vienna, Austria; school leaving examination with distinction (June 20, 1994)

1993 and 1994: Qualification for and participation in the nationwide competition of the Austrian Chemistry Olympiad

Selected summer jobs and special courses:

Arche Noah (Austrian seed savers association), Schiltern (Lower Austria):

Summer job, Aug. 1995

Raiffeisen Ware Austria (RWA), Vienna, Austria:

Summer job in the accounts department, Sept. 1995

Novartis Research Institute, Vienna, Austria:

Summer job, Aug. 1997

Ludwig Boltzmann Institut für Ökologischen Landbau und Angewandte Ökologie (LBI for Ecological Agriculture and Applied Ecology), Vienna, Austria:

Summer job, July – Sept. 1998

Supervision of the alpine botanical garden, Rax (Lower Austria):

Week-end job, July and Aug. 2002

Austrian Society for Cytometry: practical course in flow cytometry, September 2002.

Vienna University: advanced statistics course "Multivariate analysis in anthropology and systematics", visiting Professor J. F. Rohlf, May and June 2004.

Languages:

Native language: German

Fluency in English

Very good knowledge of Spanish (D.E.L.E.-básico [now "-intermedio"] in 1999, various courses and stays in Spain, México, and Ecuador; advanced intensive course at the Instituto Latinoamericano in Vienna, Austria, Sept. 2005; advanced conversation course at i-diom Language Institute Vienna, Austria (March 2007 – June 2008))

Good knowledge of French (4 years at school), Portuguese (2 years' course 1997-1999), and Latin (6 years at school)

Basics in Italian (self-educated), Chinese (2 years' course 1996-1998) and Greek (self-educated)

Proficiency in molecular techniques:

Chromatography (TLC, MPLC, HPLC, gel filtration, affinity chromatography, ion exchange chromatography, GC-MS), FT-IR and UV spectroscopy

PCR (semi-quantitative, real-time), DNA sequence analysis, RNA and DNA extraction, Dig-labeled *in situ* hybridization

SDS-PAGE, Western Blot, various ELISAs, immuno-histochemistry, immune fluorescence

Diverse colorimetric, fluorimetric and radioactive labeling assays ($[^3\text{H}]$ thymidine incorporation, apoptosis, cytotoxicity, total protein, cellular differentiation, *et c.*)

Flow cytometry and MACS

Cell culture, in particular malignant human cell lines and primary cultures

Digital image analysis

Proficiency in statistical analysis:

Broad range of standard methods, sound knowledge of multivariate techniques including principal components and principal coordinates analysis, singular value decomposition, factor analysis, cluster analysis, non-metric multidimensional scaling, discriminant analysis, correspondence analysis, Cox regression analysis, *et c.*

Excellent skills in SPSS, SYSTAT, S-PLUS, NTSYS-pc; good and basic programming skills in Mathematica and R, respectively

Teaching experience:

March 2005 – Feb. 2007: students' tutorial "Functional Pathology" at the Department of Pathophysiology, Medical University of Vienna

Regular private tutoring in chemistry for students

Reviewer for:

Anthropologischer Anzeiger (Journal of Biological and Clinical Anthropology)

Calcified Tissue International

Clinical Chemistry and Laboratory Medicine

Cytokine

Scientific memberships:

Österreichische Gesellschaft für Endokrinologie und Stoffwechsel (ÖGES) (Austrian Society for Endocrinology and Metabolism)

Verein zur Erforschung der Flora von Österreich (Society for the Exploration of the Flora of Austria)

○ Original Publications

Full papers

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2005. Differentiation-dependent expression and mitogenic action of interleukin-6 in human colon carcinoma cells: relevance for tumour progression. *Eur. J. Cancer* 41(15), 2347-2354.

<http://www.ncbi.nlm.nih.gov/pubmed/16176872>

Wutzi, A., **Brozek, W.**, Lernbass, I., Rauner, M., Hofbauer, G., Schopper C., Watzinger, F., Peterlik, M., Pietschmann P., 2006. Bone morphogenetic proteins 5 and 6 stimulate osteoclast generation. *J. Biomed. Mater. Res.* 77A(1), 75-83.

<http://www.ncbi.nlm.nih.gov/pubmed/16355411>

Kramer, G., Mitteregger, D., Maj-Hes, A., Sevchenko, S., **Brozek, W.**, 2007. Chronische Inflammation als Promotor- und Behandlungstarget bei benigner Prostatahyperplasie (BPH) und beim Prostatakarzinom. *Urologe A* 46, 1095-1096.

<http://www.ncbi.nlm.nih.gov/pubmed/17671777>

Brozek, W., Bises, G., Fabjani, G., Cross, H.S., Peterlik, M., 2008. Clone-specific expression, transcriptional regulation, and action of interleukin-6 in human colon carcinoma cells. *BMC Cancer* 8, 13.

<http://www.ncbi.nlm.nih.gov/pubmed/18205904>

Brozek, W., Kriwanek, S., Bonner, E., Peterlik, M., Cross, H.S., 2009. Mutual associations between malignancy, age, gender, and subsite incidence of colorectal cancer. *Anticancer Res.* 29(9), 3721-3726.

<http://www.ncbi.nlm.nih.gov/pubmed/19667170>

Wutzi, A., Sinko, K., Shengelia, N., **Brozek, W.**, Watzinger, F., Schicho, K., Ewers, R., 2009. Examination of dental casts in newborns with bilateral complete cleft lip and palate. *Int. J. Oral Maxillofac. Surg.* 38(10), 1025-1029.

<http://www.ncbi.nlm.nih.gov/pubmed/19481423>

Wutzi, A., Gruber, R., **Brozek, W.**, Hofbauer, G., Lernbass, I., Brosch, S., Pietschmann, P., 2010. Mechanisms involved in the inhibition of osteoclast generation by the benzothioephene SERM LY117018. *Wien. Klin. Wochenschr.* 122(21-22), 626-632.

<http://www.ncbi.nlm.nih.gov/pubmed/20938747>

Mitteregger, D., Aberle, S.W., Makristathis, A., Walochnik, J., **Brozek, W.**, Marberger, M., Kramer, G., 2012. High detection rate of *Trichomonas vaginalis* in benign hyperplastic prostatic tissue. *Med. Microbiol. Immunol.* 201(1), 113-116.

<http://www.ncbi.nlm.nih.gov/pubmed/21660495>

Brozek, W., Manhardt, T., Kállay, E., Peterlik, M., Cross, H.S., 2012. Relative expression of vitamin D hydroxylases, *CYP27B1* and *CYP24A1*, and of cyclooxygenase-2 and heterogeneity of human colorectal cancer in relation to age, gender, tumor location, and malignancy: results from factor and cluster analysis. *Cancers* 4(3), 763-776.

<http://www.mdpi.com/2072-6694/4/3/763>

Abstracts

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2004. Expression of interleukin-6 (IL-6), IL-6 receptor- α , and gp130 in human colorectal cancer cells: implications for tumor growth. 18th Congress of the European Association for Cancer Research (EACR), Innsbruck, Austria, 3 – 6 July 2004; Book of Abstracts, 200.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2004. $1\alpha,25$ -Dihydroxyvitamin D₃ up-regulates interleukin-6 mRNA and protein expression in human colon carcinoma cells: implications for vitamin D anti-cancer therapy. *Wien. Klin. Wochenschr.* 116, A20-21.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2004. Interaction of $1\alpha,25$ -dihydroxyvitamin D₃ and interleukin-6 on control of human colon carcinoma cell growth. *Anticancer Res.* 24(5D), 3445.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2004. Role of interleukin-6 in modulation of human colon carcinoma cell growth. *Anticancer Res.* 24(5D), 3598-3599.

Wutzi, A., **Brozek, W.**, Hofbauer, G., Lernbass, I., Peterlik, M., Pietschmann, P., 2004. Der Einfluss von Bone morphogenic protein-5 auf Osteoblasten und Osteoklasten. *Stomatologie* 101(5), 132.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2005. $1\alpha,25$ -dihydroxyvitamin D₃ and interleukin-6: possible intersections of signaling pathways on control of human colon carcinoma cell growth. *Anticancer Res.* 25, 2280-2281.

Wutzi, A., Schopper, C., **Brozek, W.**, Rauner, M., Watzinger, F., Pietschmann, P., 2005. Die Regulation von Knochenzellen durch Bone morphogenetic protein 5 in vitro. 9. Kongreß der Österreichischen Gesellschaft für Mund-, Kiefer- und Gesichtschirurgie, Bad Hofgastein, Austria, 1 - 5 February 2005; Book of Abstracts, 110.

Mitteregger, D., **Brozek, W.**, Kramer, G., Schwarz, S.M., Maj-Hes, A., Mazal, P., Stancik, I., Albrecht, C., Marberger, M., 2007. Novel identification of interleukin (IL)-23 and IL-23 receptor (R) in the prostate: impact on growth regulation as revealed by inhibition of PC3 cell proliferation. *J. Urol.* 177, S177.

Cross, H.S., Nittke, T., **Brozek, W.**, Kállay, E., 2008. Regulation of vitamin D hydroxylase expression and $1,25$ -dihydroxyvitamin D₃ synthesis in colon mucosa. *Anticancer Res.* 28(3A), 1614.

Brozek, W., Nittke, T., Kriwanek, S., Bonner, E., Peterlik, M., Cross, H.S., 2008. Mutual associations between age, gender, anatomical location and malignancy of colorectal cancers – relationship to $1,25$ -dihydroxyvitamin D₃ tissue concentrations. *Anticancer Res.* 28(5C), 3223.

Brozek, W., Manhardt, T., Kriwanek, S., Bonner, E., Kállay, E., Cross, H.S., 2008. Expression of the vitamin D system and COX-2 in human colorectal cancer tissue. *Anticancer Res.* 28(5C), 3222-3223.

Mitteregger, D., **Brozek, W.**, Susani, M., Stancik, I., Marberger, M., Kramer, G., 2008. Increased cytoplasmic expression of the receptor (R) for interleukin (IL)-23 is a marker for epithelial proliferation, prostatic intraepithelial neoplasia (PIN), and prostate cancer (CaP). *Anticancer Res.* 28(5C), 3221-3222.

Mitteregger, D., Aberle, S.W., Makristathis, A., Walochnik, J., **Brozek, W.**, Marberger, M., Kramer, G., 2009. High detection rate of intraprostatic trichomonads correlates with immune response in obstructive benign prostatic hyperplasia. *Eur. Urol. Suppl.* 8(4), 207.

Hassler, N., **Brozek, W.**, Varga, F., Klaushofer, K., Paschalis, E.P., 2012. Influence of homocysteine and ibandronate on extracellular matrix production of osteoblast-like MC3T3-E1 cells investigated by FTIR attenuated total reflection (ATR) spectroscopy. Bone 51, S5.

Gamsjaeger, S., **Brozek, W.**, Hofstetter, B., Zwettler, E., Klaushofer, K., Paschalis, E.P., 2012. Feasibility of Raman spectroscopic analysis of skin to discriminate between healthy and diseased organic matrix: A case study. Bone 51, S8.

Brozek, W., Hassler, N., Varga, F., Klaushofer, K., Paschalis, E.P., 2012. Effect of bisphosphonates on gene expression of fibroblasts cultured in the presence of homocysteine. Bone 51, S8.

Brozek, W., Hassler, N., Varga, F., Klaushofer, K., Paschalis, E.P., 2012. Effect of homocysteine on fibroblast gene expression. Bone 51, S17.

- Theses

Brozek, W., 2002. Secondary metabolite profiling of the genus *Aglaia* (Meliaceae). Diploma Thesis, University of Vienna, Austria.

Brozek, W., 2007. Roles of interleukin-6 and vitamin D₃ in modulation of cell growth and immunity in human colorectal cancer (CRC) and inflammatory bowel disease (IBD). Doctoral Thesis, University of Vienna, Austria.

<http://media.obvsg.at/AC05036120>

- Oral and poster presentations

- Presenting author (* talk)

- * Brozek, W., Bises, G., Cross, H.S., Kaiser, H.E., Peterlik, M. Expression of interleukin-6 (IL-6), IL-6 receptor- α , and gp130 in human colorectal cancer cells: implications for tumor growth. Wintertagung der Gesellschaft der Ärzte in Wien, Jan. 28, 2004, Vienna, Austria.

- * Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M. 1 α ,25-Dihydroxyvitamin D₃ up-regulates interleukin-6 mRNA and protein expression in human colon carcinoma cells: implications for vitamin D anti-cancer therapy. 9. Jahrestagung der Österreichischen Gesellschaft für Endokrinologie und Stoffwechsel und 12. Österreichisches Osteoporoseforum, May 6 – 8, 2004, St. Wolfgang, Austria.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M. Expression of interleukin-6 (IL-6), IL-6 receptor- α , and gp130 in human colorectal cancer cells: implications for tumor growth. 18th Congress of the European Association for Cancer Research (EACR), 3 - 6 July 2004, Innsbruck, Austria.

- * Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M. Interaction of 1 α ,25-dihydroxyvitamin D₃ and interleukin-6 on control of human colon carcinoma cell growth. 7th International Conference of Anticancer Research, Oct. 25 - 30, 2004, Corfu, Greece.

- * Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M. Interaktion von 1,25-Dihydroxyvitamin D₃ und Interleukin-6 bei der Wachstumskontrolle humaner Kolonkarzinomzellen. 4. Seminar: Grundlagenforschung in der Onkologie, Dec. 3, 2004, Vienna, Austria.

Mitteregger, D., Aberle, S.W., Makristathis, A., Walochnik, J., Brozek, W., Marberger, M., Kramer, G. Infectious microorganisms in the prostate: unexpectedly high prevalence of the protozoan *Trichomonas vaginalis* in surgery derived benign prostatic hyperplasia (BPH) specimens. YSA PhD Symposium of the Medical University of Vienna, 28 – 29 May 2008, Vienna, Austria.

* Brozek, W., Nittke, T., Kriwanek, S., Bonner, E., Peterlik, M., Cross, H.S. Mutual associations between age, gender, anatomical location and malignancy of colorectal cancers – relationship to 1,25-dihydroxyvitamin D₃ tissue concentrations. 8th International Conference of Anticancer Research, Oct. 17 - 22, 2008, Kos, Greece.

* Brozek, W., Manhardt, T., Kriwanek, S., Bonner, E., Kállay, E., Cross, H.S. Expression of the vitamin D system and COX-2 in human colorectal cancer tissue. 8th International Conference of Anticancer Research, Oct. 17 - 22, 2008, Kos, Greece.

* Mitteregger, D., Brozek, W., Susani, M., Stancik, I., Marberger, M., Kramer, G. Increased cytoplasmic expression of the receptor (R) for interleukin (IL)-23 is a marker for epithelial proliferation, prostatic intraepithelial neoplasia (PIN), and prostate cancer (CaP). 8th International Conference of Anticancer Research, Oct. 17 - 22, 2008, Kos, Greece.

* Brozek, W., Hassler, N., Varga, F., Klaushofer, K., Paschalis, E.P. Effect of bisphosphonates on gene expression of fibroblasts cultured in the presence of homocysteine. International Conference on Progress in Bone and Mineral Research, Nov. 29 - Dec. 1, 2012, Vienna, Austria.

Brozek, W., Hassler, N., Varga, F., Klaushofer, K., Paschalis, E.P. Effect of homocysteine on fibroblast gene expression. International Conference on Progress in Bone and Mineral Research, Nov. 29 - Dec. 1, 2012, Vienna, Austria.

Brozek, W., Reichardt, B., Kimberger, O., Kritsch, D., Klaushofer, K., Zwettler, E. Mortality after osteoporosis-related hip fractures in Austria 2008-2010. 40th Annual Congress of the European Calcified Tissue Society, May 18 - 21, 2013, Lisbon, Portugal.

Brozek, W., Reichardt, B., Kimberger, O., Kritsch, D., Klaushofer, K., Zwettler, E. Anti-osteoporosis treatment amongst Austrian hip fracture patients: *Status quo*, and effects on mortality and subsequent fracture risk. 40th Annual Congress of the European Calcified Tissue Society, May 18 - 21, 2013, Lisbon, Portugal.

Not presenting (selection) [§]

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2004. Role of interleukin-6 in modulation of human colon carcinoma cell growth. 7th International Conference of Anticancer Research, Oct. 25 – 30, 2004, Corfu, Greece.

Brozek, W., Bises, G., Girsch, T., Cross, H.S., Kaiser, H.E., Peterlik, M., 2005. 1 α ,25-dihydroxyvitamin D₃ and interleukin-6: possible intersections of signaling pathways on control of human colon carcinoma cell growth. 2nd International Symposium on Vitamin D Analogs in Cancer Prevention and Therapy, May 7 – 8, 2005, Lübeck, Germany.

Mitteregger, D., **Brozek, W.**, Kramer, G., Schwarz, S.M., Maj-Hes, A., Mazal, P., Stancik, I., Albrecht, C., Marberger, M., 2007. Novel identification of interleukin (IL)-23 and IL-23 receptor (R) in the prostate: impact on growth regulation as revealed by inhibition of PC3 cell

[§] presenting author underlined

proliferation. Annual meeting of the American Urology Association (AUA), 19 – 24 May 2007, Anaheim, CA.

Selig, S., Nittke, T., Kállay, E., **Brozek, W.**, Holt, P., Cross, H.S., 2007. Regulation of COX-2, a marker for IBD and colorectal tumours, and of the synthesizing vitamin D hydroxylase CYP27B1 by 17 β -estradiol. Vfwf-Universitätsvorlesung 2007, June 22, 2007, Vienna, Austria.

Mitteregger, D., **Brozek, W.**, Schwarz, S.M., Stancik, I., Mazal, P., Albrecht, C., Marberger, M., Kramer, G., 2007. A novel pathway of tissue homeostasis in the prostate hallmarked by identification of interleukin (IL)-23 and IL-23 receptor (R) in prostatic cells and inhibition of epithelial cell proliferation in response to IL-23R signaling. 1st European multidisciplinary meeting on prostate and kidney cancer, 2 – 3 November 2007, Barcelona, Spain.

Mitteregger, D., **Brozek, W.**, Susani, M., Stancik, I., Marberger, M., Kramer, G., 2008. Increased cytoplasmic expression of the receptor (R) for interleukin (IL)-23 is a marker for epithelial proliferation, prostatic intraepithelial neoplasia (PIN), and prostate cancer (CaP). YSA PhD Symposium of the Medical University of Vienna, 28 – 29 May 2008, Vienna, Austria.

Hassler, N., **Brozek, W.**, Varga, F., Klaushofer, K., Paschalis, E.P., 2012. Influence of homocysteine and ibandronate on extracellular matrix production of osteoblast-like MC3T3-E1 cells investigated by FTIR attenuated total reflection (ATR) spectroscopy. International Conference on Progress in Bone and Mineral Research, Nov. 29 – Dec. 1, 2012, Vienna, Austria.

Gamsjaeger, S., **Brozek, W.**, Hofstetter, B., Zwettler, E., Klaushofer, K., Paschalis, E.P., 2012. Feasibility of Raman spectroscopic analysis of skin to discriminate between healthy and diseased organic matrix: A case study. International Conference on Progress in Bone and Mineral Research, Nov. 29 – Dec. 1, 2012, Vienna, Austria.

○ Patent

Brozek, W. Method for characterizing, in particular quantifying, molecular markers that are intracellularly absorbed from tissues by blood macrophages that are recirculated from the tissues into the circulatory system. European Patent, filed in 2009, application number EP09781481.

http://v3.espacenet.com/publicationDetails/biblio?CC=WO&NR=2010015633&KC=&locale=en_WO&FT=E