

## Extent and limits of linguistic diversity as the remit of typology – but through constraints on WHAT is diversity limited?

FRANS PLANK

### 1. Finishing unfinished business

1.1. The first true linguistic typologist – recognisable as a predecessor by the likes of ourselves, now publishing in journals like *Linguistic Typology* – was Giovanni Domenico Campanella (1568–1639). A Dominican friar from Calabria (Fra Tommaso by monastic name) who was to rise to lasting fame as a utopian social thinker with his *La città del sole*, Campanella was a philosopher in the wide and colourful Renaissance sense. As was the standard repertoire, his *Philosophia realis* (1623) covered *Physiologia*, *Ethica*, *Politica*, and *Oeconomica*, and his *Philosophia rationalis* (1638) continued with *Dialectica*, *Grammatica*, *Rhetorica*, *Poetica*, and *Historiographia*; but, with labour not yet divided between scholar and wizard, Campanella also distinguished himself as a seeker of arcaner truths.

However far he may have strayed from his early rationalism, the grammar part of his *Philosophia rationalis* did show Campanella (now in exile in Paris, after almost thirty years in prison for religious and political reasons) the sober champion of sense experience that he had been since his very first, anti-Aristotelian beginnings: he was able, or indeed bound, to advance linguistic theory because he let himself be guided by observation rather than doctrine. This was in stark contrast to current practice in the ever popular genre of Universal Grammar (UG), where conventional accounts of Latin and Greek, and whatever could be accommodated (however awkwardly) of the Romance or Germanic mother tongues of the respective Universal Grammarians, were felt to be all it needed to keep speculation going, if mostly in circles.

Campanella was aware of what was then known about these particular languages: the contemporary Romance vernaculars (Italian, French, and Spanish, the latter the language of Southern Italy's rulers), Latin, Ancient Greek, Biblical Hebrew (and Chaldean), (Classical) Arabic, Turkish, and Chinese and Viet-

name (lingua *Coconchinesium* or *Concincinorum*). His information about the languages he drew on, a modest sample but significantly including two from further afield,<sup>1</sup> was partial and partly deficient, but it was enough to convince him that linguistic theory of his day was wrong about the two major parts of speech. Nouns were NOT (to be defined as) words inflecting for case and number: number was the only category they inflected for in the Romance vernaculars, Hebrew, and (or so believed Campanella) Arabic, and they did not inflect at all in Chinese and Vietnamese; nor could nouns be said to be words which would either inflect for case or be accompanied by (*p*)*articuli*:<sup>2</sup> they could be in the simultaneous company of both kinds of markers, as in Ancient Greek. Verbs were NOT (to be defined as) words inflecting for tense, person-number, and possibly further categories: they did not inflect at all in Chinese and Vietnamese (or so missionaries had recently been reporting, not perhaps wholly accurately).

Thus, there was more variability to be observed for parts of speech than had been assumed by UG on too narrow a factual basis; yet variation was not unlimited because relevant variables, though not connected logically, were not varying independently. Campanella's typological discoveries consisted in identifying these dependencies. The first was that the inflectional behaviours of nouns and verbs appeared not to be independent of one another; the second was that two inflectional categories of nouns, case and number, appeared not to be independent of one another either. Stated implicationally, Campanella in effect discovered (i) that if nouns inflect for case, then verbs inflect too, for whatever category (tense, person-number, . . .), but not vice versa; and (ii) that if nouns inflect for case, then they also inflect for number, but not vice versa. And at this Campanella would leave it: no explanations, however arcane, were offered for these inflectional asymmetries among parts of speech and among the categories they inflect for.

The next typological discovery, now in the structural domain of word order, was made by François (de) Mesgnien (1620/23?–98). Born in Lorraine, but widely travelled, as a student (at Rome) and after, Mesgnien published *grammaticae civiles* of French, Italian, and Polish, which suggests that he must have been teaching modern languages for a living. Long based in Constantinople as a diplomat in Polish (hence the name form Meninski) and Austrian service, he then made a name for himself as an Oriental scholar with his monumental *Thesaurus linguarum orientalium* (1680), produced at his own printing

---

1. A little to the east of the island of Taprobana where Campanella had located his utopian republic, apparently with Ceylon as the geographical model.

2. A confused category, not only in Campanella's system, alternately referring to definite articles, prepositions, or fusions of prepositions and definite articles (like Italian *nel* = *in* + *il* 'in the').

press at Vienna. The three volumes of this *Thesaurus* were accompanied by a contrastive grammar, *Linguarum orientalium turcicae, arabicae, persicae institutiones seu Grammatica turcica* (1680),<sup>3</sup> demonstrating Mesgnien's familiarity not only with Turkish, Arabic, and Persian, but also with French, Italian, German, Polish, Hungarian, Greek, Latin, and several further languages mentioned in passing.

Mesgnien accordingly could not but find fault with one tenet of UG after another, insofar as many things supposedly invariant could be missing or be different: definite articles (not found in Turkish and Persian); prepositions (postposed in Turkish); personal and possessive pronouns (the former only optional in some languages, the latter suffixed to nouns in Turkish); genders (no such contrast made in Turkish and Persian); most cases (ranging from really only a single one in Persian, contrasting with a multi-purpose form in *ra*, to three in Arabic and six in Turkish, with the nominative here coinciding with a special *accusativus indeterminativus*); numbers (with a separate dual in Arabic, and with the nominal plural identical to the verbal plural in Turkish); negation (forming part of verbal inflection in Turkish); inflection (with categories expressed separately rather than cumulated in Turkish).

On the positive side, in Part 6 of the *Grammatica*, *De syntaxi*, Section 2, *De ordine constructionis* (1680: 146–148), Mesgnien could report a discovery: on his evidence, the linear order of constituents did not vary independently across a whole range of constructions, especially ones instantiating the general relation of government – subject, object, and VERB; nominal attribute and HEAD NOUN; adjectival attribute and HEAD NOUN; noun phrase and ADPOSITION; clause and CLAUSAL PARTICLES such as conjunctions and interrogative words (with the respective governors in small capitals). Mesgnien's basic generalisation was that “regens debet semper postponi suo recto, seu casui quem regit, ideoque Verbum, quòd omnia regere videatur, ultimum orationis locum obtinet”, as in one group of languages prominently including Turkish but also Hungarian and German, or “horum ferè omnium contrarium evenit”, as in the complementary group including Arabic and less strictly Persian.<sup>4</sup> No one particular instance of a governor-governed construction was considered by Mesgnien to be more decisive than the others: what was seen as being expressed through linear order, either one way or the other, was the relationship of government as such. Simplicity of the overall particular grammar was the

---

3. The bulk of the first edition was destroyed by fire at the siege of Vienna in 1683; the re-editions of 1756 and 1780–82 came too late for the next generation of typologists, centred around the *Encyclopédie* at Paris, to benefit from.

4. To translate: The governor must always be placed after its governee, or the case which it governs, and likewise the verb, which seems to govern everything, comes last in its clause [in Turkish etc.]. [...] But the opposite of nearly all this obtains [in Arabic etc.].

obvious rationale of such cross-categorical harmony in arrangements; but learnability, ease of processing and/or production, clarity, or iconicity took another century to gain prominence as explanatory notions – too much prominence sometimes, eclipsing facts inconsistent with supposed explanations, such as rigid OV and AN ordering, cold-shouldered by word order typologists of the eighteenth century for whom verb/action HAD to come before object/effect and noun/substance before adjective/property, because this was the linguistic order that was iconically mirroring the “natural” order of thought.

1.2. And so, for more than three centuries, typological discovery was to follow upon typological discovery. Individually they were perhaps not very spectacular, but they were adding up, with meanwhile several thousand universals on record (as documented in THE UNIVERSALS ARCHIVE, at <http://typo.uni-konstanz.de/archive>, if incompletely). Progress was not always cumulative: old discoveries would sometimes be forgotten and perhaps be independently rediscovered (like, famously, Mesgnien's word-order harmony). Sometimes, progress was eliminative: what had seemed a discovery would eventually be shown to have been a figment of a system seeker's imagination, due to inaccurate or insufficient knowledge of languages or faulty generalisation (e.g., nouns MAY inflect for case without verbs inflecting too, or so a counterclaim to Campanella's would have it, 327 years later: Capell 1965; adjective–noun does NOT pattern with object–verb, with Mesgnien standing corrected after only 308 years: Dryer 1988); sometimes, rather than being abandoned, universals found inadequate would be rescued by weakening or otherwise modifying them. Alternatively, superior factual evidence would sometimes lend support to universals previously not firmly grounded. All too rarely were universals confirmed or disconfirmed upon a replication of their first proponent's investigation.<sup>5</sup>

1.3. Although much has been accomplished by way of adding to, deleting from, and otherwise revising the fund of universals, much remains to be done for today's and tomorrow's typologists in pursuit of what has remained essentially the same research programme for over three and a half centuries. In particular:

First, knowledge about more and more particular languages needs to be acquired (as spoken or signed, perhaps whistled or drummed, too; as developing from infancy to old age of the users of a language), to be made generally accessible, and to be conscientiously utilised by typologists. The typologist's special responsibility here is to see to it that descriptions are in such formats as to facilitate, or indeed allow, comparisons across languages, which minimally

---

5. See Plank (2001, 2007) for a fuller story to go with this historiographical précis.

means that those describing particular languages need to be (made) aware of the potentials for variation across languages, and maximally that a master plan is made available for the description of all languages. (Typologists will specially regret that so many languages have been, and are being, discontinued undescribed or undocumented: after all, their universals aim to be true of all languages at all times, rather than only for those left for them to compare. Should mass decimations of languages, or of individual distinctive traits, have occurred in the 100,000 or 200,000 years of linguistic history, our conception of what is or is not a possible language is in danger of taking for necessity what is mere chance.)

Second, knowledge about particular languages needs to be more comprehensive, more in-depth (informed by “thick” rather than “thin” description), and more reliable: typology will only be as good as the language-particular descriptions it can draw on. The typologist’s role here is, not only to be a beneficiary, but also to give guidance, and perhaps lend a hand in describing undescribed or poorly described languages.

Third, and this is the genuinely typological remit, more and more individual variables – parameters along which languages can vary – need to be identified, and it then needs to be ascertained whether or not these variables co-vary; where co-variation is found, it will want an explanation. The occupational hazard here is that it is the easier to persuade oneself that something, or a connection between some two things, is invariant the less one knows about actual variation. There are sceptics who doubt whether anything will remain which could truly claim universality once our awareness of diversity is sufficiently wide; if they are right (which is an empirical question), the typological programme in its classical formulation can be discontinued in favour of historical research, in particular (macro-)areal linguistics and population history (in the manner of Nichols 1992), with the remit of tracing to joint inheritance and borrowing what relevant languages (non-universally) share. This would, after all, still be the same general enterprise of seeking to answer the big question of predictability: What is predictable about a language, any language, on the basis of what? – with possible predictors including human languagehood as such (assuming there are universals), other structural properties of the language concerned, its internal and external history, the current and previous locations of its speech community, and current and previous contacts with other speech communities. What typologists may usefully ponder, until this question of what are valid predictors is decided, is whether anything should be EXPECTED to co-vary with anything else, unless logically connected. Surely, interpreted non-trivially, the old dictum that languages are systems, organisms, or mechanisms where *TOUT se tient* is a gross exaggeration. When common heritage (which in turn raises a question: Why should two independent traits be JOINTLY retained?), common contact possession (Why JOINTLY borrowed?), and chance are discarded, what

is left that could nourish the expectation of co-variation for any two variables? Possible answers are: very general and abstract organising principles of (sub-) systems of linguistic articulation of thought; or developmental scenarios with stages and cycles, of particular durations, which happen to occur in sync.

Fourth, to be able to convincingly confirm or disconfirm co-variation, the methods of inductive generalisation, of drawing valid inferences about universals from limited crosslinguistic evidence, need cultivating. The questions of how to sample and what are appropriate statistics (addressed a lot in *LT*'s pages) will therefore have to remain on the methodological agenda. As things are, the representativeness of samples is defined genealogically and/or areally; and much has been achieved here in guarding against the mistaking of common heritage from protolanguages and/or borrowings among (macro-)areal neighbours for universals. Obviously, what typological samples ought to be representative of is the full range of structural diversity, as brought about by individual innovations diversifying speech communities upon being adopted: without us knowing how, and how fast, languages can and cannot change through successful innovations, on their own or in contact, genealogical and areal balancing does not perforce guarantee structural balance.

Fifth, further specialisation and professionalisation will be inevitable for typological research in future, dictated by the growing workload (more and more languages and more and more variables to be taken into account) and the due refinements in methodology (with microscopic scan and calculation replacing macroscopic panorama and impression). Whereas in the old days (in fact, until rather recently) *literati*, philosophers, diplomats and other travellers, jurists, or enterprising philologists could make typological discoveries, typology now is likelier to be done by trained typologists who, in mutual awareness and often in collaboration, pursue the quest for unity in crosslinguistic diversity as an end in itself, not as a pastime or as a by-product of other concerns. However, given the possibility of constraints on diversity reflecting constraints on change, or of universals mutating into inheritance or borrowing histories, the requisite sort of specialisation had better include an expertise in (crosslinguistic) historical linguistics. And doing typology professionally is unlikely to ever mean not also to be doing linguistics as such: just as you cannot hope to contribute meaningfully to phonological, morphological, syntactic, or lexical theory without an awareness of how languages can and cannot differ, phonological, morphological, syntactic, lexical, or indeed whole-system typology cannot be advanced in a theoretical vacuum or on the pretence of a-theoretical analyses. There are no typological "facts" that would speak for themselves, however sophisticated one's fact-finding and fact-processing methods.

1.4. In an ostensibly innocuous manner of speaking, also adopted here up to now, the business of typology can be said to be the mapping of the extent

to which languages do differ and the discovery of constraints on how they may differ. In the remainder of this paper I would like to suggest that this is perhaps not the most perspicuous remit and that, seeking to finish unfinished typological business, we ought to be clearer about what it is that is supposed to be constrained – the human linguistic faculty in its rigidity or its flexibility (Section 2); concrete or abstract representations of grammatical and lexical forms and structures (Section 3).

## **2. Constraints on analysis or reanalysis: Timeless laws or laws of change?**

Disregarding the many ways in which this cherished concept is problematic, “a language”, differing or not differing from “another language” in one or another particular, is something individual and social. Difference comes about when individuals innovate and other members of their community follow their example: this is where constraints must be effective.

The linguistic know-how of an individual is the grammar & lexicon represented in that individual's mind (well, brain). The know-how to express and communicate thought can be shared among individuals, or else it distinguishes one group of individuals from another: speech communities, with no differences in linguistic know-how among their individual members (an idealisation); families/genera/phyla and *Sprachbünde*/diffusion zones/macro-areas, with the members of each (sets of speech communities) sharing what is “inherited” or “borrowed” respectively; the whole subspecies of *homo sapiens*<sub>1</sub> *sapiens*<sub>2</sub>, with at least that in common which earns it the attribute *sapiens*<sub>2</sub> and its near-synonym *loquens* (openness or indeed recursion, double articulation, ...), and with further, non-definitional universals to be discovered through typological inquiry.

Grammars & lexicons are ACQUIRED: the linguistic histories of individuals and of populations are the histories of the acquisition of linguistic know-how by individuals within speech communities (with the early life span of individuals probably the most crucial), possibly enriched by input (early or late in an individual's life) across speech communities or from other sections of heterogeneous speech communities. Constraints on crosslinguistic diversity, then, can only be constraints on acquisitions and on what acquirers can make of them and what gets accepted by speech communities as the general norm.

Now, when SOME grammars & lexicons, but not all, are found in some respect to be the same (and what is being compared are mental constructs), this will naturally be attributed to the particular linguistic experiences of the acquirers concerned having been (relevantly) the same – notwithstanding the possibilities that identical experiences can also lead to different grammars & lexicons and different experiences to identical grammars & lexicons. If ALL grammars & lexicons can plausibly be assumed to be the same, in one particular or more

likely in the way two or more variables co-vary, and this identity is not a design feature “language” is defined through, there are two ways to make sense of such lack of diversity: (i) through constraints on acquisition (be they genetic or “functional”, specifically linguistic or more generally cognitive) which always and everywhere invariably enforce the same grammatical & lexical representations irrespective of the particular linguistic experience of acquirers within their speech communities; or (ii) through constraints which in light of particular linguistic experiences encourage or even force acquirers (including borrowers) to set up grammatical & lexical representations which are different, though uniformly different, from those of their linguistic models.

Typological constraints can accordingly be conceived of A-chronically or DIA-chronically, as timeless laws (i) or as laws of change (ii):

- (i) Universals, seen as timeless laws, impose (absolute or conditional) limits on variation across grammars & lexicons at any and all times, regardless of particular linguistic experiences;  
they thereby constrain change insofar as a grammar & lexicon must not change (not be restructured) so as to violate such a universal, or at any rate not without subsequent changes swiftly redressing the balance.  
(Possibly: There are no laws of change itself; anything can be reanalysed as anything else independently of anything else, as long as no timeless law is violated.)
- (ii) Particular targets (forms, categories, constructions, rules, constraints) can only result by particular mechanisms of change (reanalysis) operating on particular sources (forms, etc.) which an individual experiences at a particular time in a particular speech community;  
such laws of change thereby impose limits on how grammars & lexicons can differ: they can only be what they could become, under the contingent historical circumstances.  
(Possibly: There are no timeless universals. Or: Co-variation is due to co-evolution, with concomitance or consecutiveness of historical reanalyses being superintended by timeless laws.)

In one or another guise, the view that has diachrony in charge of the constraining (ii) is the more traditional one, with types conceived of as developmental stages inexorably succeeding one another and perhaps recurring in cycles. For a while superseded by position (i), it has lately been regaining popularity, inspired by Greenberg’s programme to “dynamicise” typology (with several important papers collected in Greenberg 1990). In the case of implicational constraints, the most straightforward way of dynamicisation has been to read “implies” ( $\supset$ , material implication) as “derives from” ( $\leftarrow$ , is reanalysed from). To briefly illustrate with three examples from syntax, morphology, and phonology.



Prep NP  $\supset/\leftarrow$  N Genitive, NP Postp  $\supset/\leftarrow$  Genitive N:

When adpositions are “grammaticalised” from head nouns in attributive constructions, which is one of their sources and manners of origin (that is, when such constructions are analysed by acquirers differently from their models’ analysis, along preordained lines leading from more lexical to more grammatical forms-in-constructions), then they will remain in the same position relative to the NPs which they continue to be in some sort of similar construction with, automatically harmonising in head-dependent ordering or in branchingness direction with the source construction, provided the grammar & lexicon of the new generation within that speech community remains otherwise unaltered (thus, e.g., English *beside the house*  $\supset/\leftarrow$  *by the side of the house*; Turkish *Bariş-a inat* Bariş-DAT despite ‘despite Bariş’  $\supset/\leftarrow$  *Bariş-in inat-ı* Bariş-GEN persistence-3SG.POSS ‘Bariş’s persistence’; see further Aristar 1991, in a Greenbergian sense dynamicising seemingly achronic word order universals of Greenberg’s 1963).

Infixes  $\supset/\leftarrow$  adfixes:

The (almost) only way infixes can come about is from adfixes being metathesised inside stems, around initial or final consonants or other phonological constituents, serving to improve syllable or other prosodic structures; in the company of stems where prosodic structures are unobjectionable anyhow, adfixes will remain external. (Gabelentz 1891: 330; more on in- and adfixation below, Section 3.5.)

Nasal vowel  $\supset/\leftarrow$  corresponding oral vowel:

The (almost) only way distinctive nasal vowels can come about is from oral vowels being nasalised through the influence of a following nasal consonant, itself consequently dropped from lexical representations. (For a recent all-out attempt to dynamicise phonological typology see Blevins 2004.)

Though seemingly straightforward in selected cases, this sort of “ $\supset = \leftarrow$ ” dynamicisation is too facile, and cases are easily found where it is nonsense – e.g., dual  $\supset$  plural, an instance of marked  $\supset$  unmarked: duals are hardly ever due to reanalyses of plurals, \*dual  $\leftarrow$  plural, while, the other way round, dual forms frequently acquire plural meanings, dual  $\supset$  plural. The full diachronic stories behind implications are far more complex,<sup>6</sup> and to be meaningfully told,

6. As, from my own point of view, adumbrated for a few morphological cases in Plank (1999, 2001, 2003a/b) and Plank & Schellinger (2000). There are numerous papers in the first ten volumes of *LT* where crosslinguistic patterns are accounted for diachronically, often under the rubric of “grammaticalisation”; but what is supposed to be the essence and force of diachronic constraints would merit livelier discussion. Hundreds of entries in *THE UNIVERSALS*

they need to be embedded in the histories of acquisitions of grammars & lexicons through individuals within their speech communities and with possible input also across speech communities.

As grammars & lexicons are being shaped, to some extent predictably, by the particular experiences acquirers happen to make, there would still seem a role to play, as per the achronic view of universals (i), for experience-independent laws prescribing or proscribing certain grammatical & lexical representations and superintending the concomitance or consecutiveness of reanalyses. But the question remains, and should be high on the typological agenda, precisely how (diachronic) constraints on reanalyses would relate to (achronic) constraints on analyses in limiting differences between mental grammars & lexicons which individuals can come up with, at any and all times as well as in particular historical circumstances.

And, to be able to govern the structuring or restructuring of mental representations, universal constraints themselves need to be (recognised as) part of the mental equipment of individual linguistic agents. But then, the most abstract universals on record perhaps excepted, which are arguably owed to human genetics, the mental status of universals, and especially of implicational universals, has rarely been an issue, however solid their “functional” grounding (facilitating the expression of thought; facilitating storage, access, production, processing; subserving inertia or extravagance; giving pleasure; etc.): it decidedly ought to be one in future.

### 3. Typology modularised: Constraints on concrete or abstract representations?

3.1. Typological discovery has always been data-driven: Campanella and Mesgnien had been alerted to patterns of uniformity in diversity by knowing (of) languages which deviated in inflectional paradigms and surface-syntactic arrangements from those languages which then current linguistic theory had taken its inspiration from. In turn, theory – Advanced Linguistic Theory (ALT) – has been driven by typological discovery: with Campanella and Mesgnien, linguistic theory of old, chapters *partes orationis* and *constructio*, became an anachronism. In the meantime the quantity and quality of data potentially relevant to typology has been greatly expanding and improving: linguistic structures, as analysed today, much inspired by a growing awareness of diversity, are more complex than they ever used to be perceived as; their acquisition and loss, beginning to be seriously investigated across languages and cultures (with Slobin 1985–97 a milestone for acquisition), have been observed to be orderly

---

ARCHIVE, from phonetics/phonology to semantics/lexicon, admit of achronic as well as diachronic interpretations, with the ambivalence often unresolved.

rather than random; and direct access is being gained to their mental/neural representations, with English no longer the only language studied in experimental psycho- and neurolinguistics. Driven by ever richer data, typological discovery will continue to drive theory – ALT, at any rate – also in future.

One issue, with many facets, over which contemporary theorising and experimenting has been divided is the concreteness or abstractness of linguistic representations. The point I would like to make here is that abstract representations derive strong support from typology, insofar as constraints on diversity are plausibly conceived of as differentially constraining abstract and concrete representations. Looking at it the other way round, it is only by reckoning with abstract as well as concrete representations that diversity can be seen to be subject to constraints and variables and invariants can be correctly identified.

3.2. In barest outline, concrete representations typically have the following ingredients, adding up to a rich and complex picture of linguistic structure:

- forms-in-constructions, at complexity levels of segments, syllables, feet, morphemes, (phonological/syntactic) words, (phon/syn) phrases, (phon/syn) clauses, (phon/syn) sentences, paragraphs, texts;
- some sort of strength or salience of particular forms (exemplars) in particular constructions, owing to their frequency of occurrence;
- (sub-)classes of forms, (sub-)classes of constructions:
  - form classes, distribution/position classes, meaning classes;
- relation of (immediate) precedence in constructions (i.e., linear order);
- relation of (immediate) constituency in constructions (i.e., part-whole);
- syntagmatic relationships between manifest constituents of constructions, not necessarily specific to particular classes of constructions:<sup>7</sup>
  - subjunction, or dependency (head – dependent),
    - with subtypes: argument, circumstance; modification, determination, complementation; attribution, predication, apposition, . . . ; quantification, classification;
  - conjunction,
  - adjunction,
  - . . . ;
- several kinds of influence (preferably local, possibly also at a distance) among manifest constituents of constructions, not specific to particular classes of constructions (and it is in how to express such influences that analytic frameworks differ greatly, making translation from one to another

---

7. “Not specific” means “general”, which is to be distinguished from “abstract”. Ditto for “form/construction classes” (general) as opposed to “exemplars” (particular).

so difficult, badly hindering communication among subcommunities of linguists):

- exclusion of co-presence,
- requirement of co-presence,
- license of co-presence,
- government, command,
- agreement (directional or other),
- referential binding,
- scope-taking,
- ...;
- paradigmatic relationships;
- kinds of constructional marking:
  - relationship-identifying,
  - relatedness-indicating,
  - linking, separating,
  - speech-event anchoring and context anchoring;
- kinds of grammatical meanings/functions:
  - reference, reference-modification, predication, predicate-modification;
  - information-structuring: topic, comment, focus, ...;
  - assertion, question, command, exclamation, ...

Such structural concepts are what “descriptive” or “reference” grammars are usually based on, however variable their formats. For typical concrete descriptions of particular languages which uniformly follow one format, specified through the questionnaire of Comrie & Smith (1977), see the (ex-) *Lingua Descriptive Series*. For typological constraints as typically cast in concrete terms see THE UNIVERSALS ARCHIVE (at <http://typo.uni-konstanz.de/archive>).<sup>8</sup> And in the ten volumes so far of *LT* concreteness has been the predominant mood, too:<sup>9</sup> in circumscribing crosslinguistic patterns essential reference to non-concrete representations is made in barely a dozen papers.

Again very generally speaking, linguistic representations can be abstract, first, by virtue of excluding all sorts of things present in the speech signal, but considered irrelevant for grammar; second, by virtue of including some sorts of things not present in the speech signal, but considered relevant for grammar; third, by virtue of structural relations between forms-in-constructions differing from those obtaining in manifest structures. Sometimes, it is only by assuming constraints on such abstract representations – where structurally NOTHING

---

8. Or, for that matter, the constraints of Optimality Theory (not to put too fine a point on their typological credentials).

9. Or, for that matter, in the Rutgers Optimality Archive (not to put too fine a point on typological credentials).

corresponds to manifestly SOMETHING; where structurally SOMETHING corresponds to manifestly NOTHING; where structurally relevant PRECEDENCE of parts of wholes differs from manifest precedence (movement/displacedness); or where structurally relevant constituency differs from manifest CONSTITUENCY (restructuring) – that invariants can be factored out from variation. Abstract representations, therefore, should not be anathema in language-particular description, and should receive meticulous theoretical attention informed by current knowledge about crosslinguistic diversity.

3.3. To first illustrate, from segmental phonology/phonetics, how manifestly something can correspond to nothing in an abstract representation, here is a thumbnail sketch of the “Coronal Syndrome” – the asymmetry between sounds pronounced with the blade of the tongue raised from its neutral position (dental, alveolar, and probably palatal consonants, front vowels), distinguished through the (privative) feature [CORONAL], and ones with the tongue blade in neutral position (labial, velar, uvular, pharyngeal, glottal consonants, central and back vowels):

First, coronals are the most frequent articulator choice on a number of counts. In the UG phonetic alphabet, the Coronal articulator supports a larger number of dependent (consonantal) features than Labial, Dorsal, and Pharyngeal. In the phonemic systems of individual languages, coronals typically outnumber the other Place categories as well. Finally, they have been documented as among the most frequent consonants in speech corpora of English and Spanish. Second, coronal is the normal outcome of rules and constraints that neutralize Place contrasts [...]. Third [...] phonologists have the impression that coronal is the most commonly chosen epenthetic or otherwise dummy oral consonant [...]. Fourth, coronals more freely combine with each other as well as with other consonants, eluding phonotactic restrictions that are enforced on labials and velars. [...] Fifth, coronals are more susceptible to Place assimilation than noncoronals [...].<sup>10</sup> Finally, coronals are more likely to be transparent to transconsonantal vowel-echo rules than labials or velars. (Kenstowicz 1994: 516–517)

For an in-depth examination of the coronal syndrome see Paradis & Prunet (1991); amidst some theoretical controversy (with pros and cons fairly summarised in Lahiri & Reetz 2002, 2007), much crosslinguistic work since has confirmed the basic factual diagnosis, which has also been substantiated through acquisition studies and psycho- and neurolinguistic experimentation for a range of languages (see again Lahiri & Reetz 2002, 2007).

---

10. That is to say, non-coronals may assimilate in place as well as coronals; but, for any language, if non-coronals assimilate in place, then coronals assimilate in place too (cf. Mohanan 1993).

Thus, there is a whole family of patterns, rules, or constraints where coronal and non-coronal segments behave asymmetrically, with the patterns of inequality always the same: in any language, either only coronal segments can do something which non-coronal segments can't, or whatever non-coronals do, corresponding coronals do, too. Rather than separately positing some implicational law to take care of this asymmetry, it is, in the most simple and natural manner, accounted for through lexical phonological representations which are more abstract than phonetic representations (and perhaps than non-lexical phonological representations), lacking something inevitably present in instructions for pronunciation and in the speech signal: the full specification of all segments for all features. The requisite abstract representations are underspecified, with coronal segments remaining universally unspecified for place of articulation;<sup>11</sup> it would only be for purposes of pronunciation and perception that the feature [CORONAL] is specified, or perhaps already for non-lexical representations in the case of those languages where this feature [CORONAL] is phonologically active (e.g., when coronals assimilate in place too, like non-coronals – if the assimilation of non-coronals indeed is a process of the same kind as that of coronals: but it probably isn't).

Assuming specifications of features only where necessary to distinguish lexical items of a language and where a feature is phonologically active in that language, thus, provides an angle – more advantageous than full-listing or exemplar approaches which eschew abstract phonological representations with predictable and redundant featural information missing – on how segmental systems and processes can and cannot differ across languages. What is seen to be constrained categorically and unconditionally – and timelessly, not as a consequence of any law of change – are lexical representations, featurally underspecified with respect to [CORONAL]; pronunciation and perception forms and perhaps non-lexical phonological representations admit of diversity. As is yet to be examined in greater typological detail, this diversity is patterned rather than random, too, equally showing a coronal/non-coronal asymmetry.

Underspecification of morphological, syntactic, and semantic representations may prove equally attractive for typological theorising. In phonology again, whole prosodic constituents (moras and syllables) can be phonetically realised, at the edges of prosodic domains, but yet be absent from representations relevant for metrical structure: and such extrametricality appears to have a role to play in the typology of word stress.

3.4. The opposite relationship of structurally something corresponding to manifestly nothing can again be exemplified from phonology. In what is known

---

11. Which in addition permits the identification of a set of consonants which are often found to form a natural class: those which are (abstractly) placeless (Ghini 2003: 447).

as catalexis,<sup>12</sup> prosodic constituents (moras and syllables) are included in metrical structures which remain segmentally empty, confined to edges of prosodic domains like extrametrical constituents are: subject to the peripherality constraint, a manifestly light syllable can count as abstractly heavy with a catalectic mora added, and a syllabic foot of manifestly only one light syllable can count as abstractly binary with a catalectic syllable added. Perhaps, though without direct phonetic realisation, catalexis can make itself felt through segmental lengthening of a syllable in the same foot when phrase-final (Flores 2004), or through the way it interferes with syllabification (as when overtly word-final consonants behave like onsets, followed by a catalectic mora (Ghini 2003)).

Catalexis has become a popular analytic tool in metrical phonology, but one that continues to be controversial, and for just about any individual language where it has been invoked, alternative analyses have been considered that are supposedly less abstract, doing without such inaudible moras or syllables. Nonetheless, it is worth pointing out that the potential benefits of catalexis, itself representing a parameter where acquirers have a choice of values (yes or no, apparently with word-minimality effects as the trigger experience), are primarily typological.

First, bona fide universals which would otherwise have to be sacrificed (or to be admitted to be violable) can be rescued if catalexis is accepted as part of abstract phonological representations: the universal foot inventory can remain limited, excluding degenerate feet, insofar as seeming degenerates (monomoraic or monosyllabic, depending on the presence or absence of a weight contrast) are redeemed through catalexis; a universal constraint on word minimality (minimally one foot, bimoraic or disyllabic, depending on the presence or absence of a weight contrast) can be upheld, insofar as seemingly subminimal words are redeemed through catalexis. Second, catalexis itself promises to be implicated in universal correlations: catalexis in monosyllables appears to imply catalexis in polysyllables (Kager 1995); word-minimality effects appear to correlate with stress patterns, insofar as, for example, rightward-trochaic stress systems have or lack secondary stresses on final odd-numbered syllables depending on whether or not they permit manifestly subminimal, abstractly catalectic words (Kager 1995, building on Kiparsky 1991); as to wider rhythmic typology, a favourable climate for catalexis appears to be stress-timing (or mixed stress/syllable-timing), but not syllable-timing (Flores 2004).

---

12. Better known perhaps in poetics than in linguistics; but through work of Paul Kiparsky's (1991, apparently unpublished) catalexis has been introduced to phonologists. Catalexis and related matters are examined across a wide range of languages in Kager (1995) – 44 rightward trochaic languages, 57 penultimate-stress languages: not bad at all by *LT* standards.

Outside phonology, traces and empty categories, or possibly also ellipsis, are examples from syntax where manifestly nothing arguably corresponds to structurally something: it remains to be seen whether their recognition receives special support from, and sheds extra light on, the study of variation and invariance.

3.5. Infixation provides an example where abstract (morphological) order is profitably recognised as being different from concrete (phonological) order.

On the face of it, affix order is evidently variable: there are suffixes, most common crosslinguistically; prefixes, less common; circumfixes (probably also transfixes, unlike circumfixes morphologically not bipartite); and infixes, least common crosslinguistically. Variation is reined in by this implicational constraint: If there are infixes, there will also be adfixes (= suffixes and/or prefixes). This constraint can be seen, achronically, as a particular case, pertaining to stems, of a more general constraint against discontinuous constituents, dispreferred as being more difficult to store, access, and process than continuous constituents. It can also be seen as following from a diachronic law, with adfixes being the only productive source of infixes and with phonological reordering as the only productive mechanism of getting adfixes inside stems (in the right phonological circumstances, with adfixes remaining external elsewhere).<sup>13</sup>

An even stricter constraint would be not to permit infixation at all, anywhere and at any time, rather than making it (achronically or diachronically) contingent on adfixation. Though it is typologically desirable to push constraints as far as possible, this move would seem glaringly at odds with crosslinguistic reality. However, when word representations are separated into an abstract morphological one and a concrete one that is to be pronounced, and the NO INFIXES! constraint is only imposed on the former, the theory fits the facts, and in addition sheds light on why the facts are as they are.

The price to pay are morphological representations of relevant words in relevant languages where what are manifestly “infixes” are (abstractly) adfixes – e.g., present stems of verbs of the relevant conjugation class in Latin (Italic, Indo-European) such as *fu-N-d-* ‘shed’, *vi-N-k-* ‘conquer’, *ru-N-p-* ‘break’ (with the nasal assimilating in place to the following consonant); perfective verbs in Tagalog (Malayo-Polynesian, Austronesian) such as *k-um-ain* ‘ate’, *p-um-asok* ‘entered’; causative verb stems in Tsiene (Bantu, Niger-Congo) such as *l-as-ab-* ‘cause to walk’; construct state forms of nouns in Ulwa (Sumu, Misumalpan) such as *suu-ka-lu* ‘(his) dog’, *siwa-ka-nak* ‘(his) root’, *karas-ka-mak* ‘(his) knee’, with the “infixes” underlined. This enables the NO INFIXES! universal to be upheld, though valid only for morphological representations, where

---

13. Another mechanism on record is “entrapment”, with an outer affix very occasionally reanalysed as part of the stem. See Moravcsik (1977, 2000) for general surveys of infixation.



the order of word parts may be different from the manifest ordering of these parts.<sup>14</sup>

With the origin of affixes as a rule external, owed to univerbation, morphology prefers order to be rigid: rarely allowing for reorderings of affixes among each other and relative to stems, it never condones partial interlacings of stems and external addenda – and why should it! Pronunciation, however, may find morphological order ill-suited to its own purposes, and may have a licence to improve matters through metathesis or other reorderings. The positioning of “infixes” – adfixes in morphological representations – is, thus, always and ever accomplished by PHONOLOGICAL means, whose aim is to optimise prosodic structures – namely those (i) of syllables (as syllabified at the stem or perhaps also the word level, aiming at CV patterns, at permissible clusters, or at sonority sequencing; as in Tagalog and Latin<sup>15</sup>); (ii) of syllable groupings, i.e., metrical feet (aiming at the foot type preferred in the language; as in Ulwa, or also in English expletive and *-ma-* “infixation”, not otherwise very typical: *abso-bloomin-lutely*, *secre-ma-tary*; Green 1999: §4.3.2, McCarthy 1982, Yu 2004); or also (iii) of stem templates (as in Tiene, where derived verb stems are of the shape  $C_1VC_2VC_3$  and need to form a “prosodic trough” with  $C_2$  coronal and  $C_3$  non-coronal; Hyman 2006).

Whether they have been phonologically reordered or (very rarely) entrapped, “infixes” have always been adfixes in their previous lives (that is, in the grammars of previous generations of speakers). This is reflected, synchronically speaking, by “infixes” always being EDGE-BOUND: they are never found further inside stems than after/before the initial/final constituents of the relevant prosodic unit – before plosive syllable-coda in Latin; after any syllable-onset in Tagalog; after the first iambic foot in Ulwa; at left or right edges of final/initial

---

14. One might further hypothesise that, if there are prefixes, there will also be suffixes – were it not for individual languages or families which are predominantly prefixing, however massively outnumbered by exclusively or predominantly suffixing languages and families. For Athapaskan, where prefixation prevails in spite of other typological properties which one would expect to favour suffixation, a case has been made (Rice 2000) for analysing manifest prefixes as abstract suffixes, with morphological representations reordered for semantic or phonological reasons (scope, prosody). Radicalising the implicational constraint to, simply, NO PREFIXES (in morphology)! might be going too far, although there undeniably is a general suffixing preference which needs to be accounted for, and it is the interplay of morphology with (cliticising and further fusing) phonology that presumably does the accounting, just as with infixation.

15. In Latin, it is only at the STEM level that the nasal stem formative in suffixal position would yield an impermissible coda cluster, with a more sonorous followed by a less sonorous consonant (*fud-N-*, *vik-N-*, *rup-N-*); as all inflectional suffixes begin with a vowel, WORD FORMS would be syllabified acceptably without metathesis (*fud.nō*, *vik.nō*, *rup.nō*, etc.). So, phonological infixation has to be seen as either applying at stem level, motivated by a coda cluster constraint, or at word-form level, in which case the motivation would be a cross-syllable dis-preference of non-sonorous coda followed by sonorous onset.

trochees in English; before the final non-coronal consonant of the template in Tiene. And they are, in the relevant languages, always also realised as adfixes with stems which do not require prosodic optimisation through phonological reordering – cf. Latin stems of the same conjugation class without a stem-final plosive such as *si-N-* ‘leave’, *ker-N-* ‘separate’, *(con-)tem-N-* ‘despise’, *pell-* (< *pel-N-*) ‘expell’; Tagalog vowel-initial stems such as *um-awit* PERF of ‘sing’; Tiene verb stems with a final coronal, forming a “prosodic trough” when suffixes with a non-coronal consonant remain external, such as *mat-is-* (→ *maas-*) ‘cause to go away’; Ulwa iambic stems of two moras, contributed by one heavy syllable or two light ones, or of three moras, contributed by a light syllable followed by a heavy one, such as *kii-ka* ‘(his) stone’, *sana-ka* ‘(his) bee’, *sapaa-ka* ‘(his) forehead’.<sup>16</sup> Edge-boundedness, with “edge” defined prosodically, and external occurrence in the case of some edges strongly support the analysis of “infixes” as created by phonological reordering from morphological adfixes.

As to the accessibility of abstract morphological representations for grammar acquirers, given that there are always instances of adfixes in both a stem-external and a stem-internal position depending on the prosodic shape of stems, learners will be able to infer adfixal positioning for abstract representations, undoing optimising phonology in those cases where concrete pronunciation forms do not match them.

Reasserting their morphological adfix status, as continually encoded in abstract representations, “infixes” can again be RE-EXTERNALISED, even from the very stems they used to surface inside of. Thus, the adfix *um-*, which is widespread in Malayo-Polynesian and whose manifest position is either before stems when stems are vowel-initial or after the first consonant of consonant-initial stems, as illustrated above from Tagalog, is increasingly found in prefix position also with consonant-initial stems, with its own final consonant assimilating in place (/m/ is not a coronal, but coronals assimilate, too): e.g., *ud-dátu*, now alternating with earlier *d-um-átu* ‘wiser’, *ug-gógo*, now alternating with earlier *g-um-ógo* ‘stronger’, *ul-lógo*, now alternating with earlier *l-um-ógo* ‘drier’ (Crowhurst 1998).<sup>17</sup> Such re-externalisations will not happen once “infixes” have become lexicalised, part of internally unanalysed stem allomorphs, which they probably were as Latin turned into the Romance vernaculars; but then, being lexicalised, they are not constructive parts of morphological representations, either.

16. Atypically, American English (more particularly, Homeric, that is, Homer Simpsonic) *-ma-* does not occur externally, and never has, and when (noun-)external, expletives (*fuckin’ Chicago* etc.) are words (adjectives), not adfixes.

17. With stems beginning with a labial or nasal consonant, *um-* had always remained external in Toba Batak, with /m/ assimilating in place: e.g., *up-pásak* ‘has beaten’, *uN-Náli* ‘colder’.

*Extent and limits of linguistic diversity as the remit of typology* 61

What can be assumed to be crosslinguistically invariant, then, is order in morphological representations, always and ever obeying the constraint **NO INFIXES**! What is variable is how, at any given time, particular speech communities rank phonological (in particular, prosodic) optimality relative to morphological faithfulness in pronouncing complex words. If complex words are not required to be phonologically optimal, or if stems and affixes have phonological shapes which, when combined, yield preferred prosodies anyhow (syllables, feet, templates), then adfixes will be realised as adfixes; otherwise they will be internalised around edges, with phonological (prosodic) constituents in an order which sounds better than arrangements faithful to the morphology. The task for seekers of system in diversity (= typologists), yet to be seriously undertaken, is to identify circumstances which license or indeed require, or also which proscribe, phonological improvements of morphology as individuals are acquiring a grammar, and as members of speech communities may come up with different results. Given the historical mutability of “infixation”, these conducive or adverse circumstances, probably to do with prosodic complexity, would also be expected to be variable, and the diachronic job would be to examine co-variation as co-evolution.

In sum, while it is true to say, achronically as well as diachronically, that infixes imply adfixes, this implication as such has no status in mental grammars (and, as such, is of little theoretical interest); it is the (prosodic) phonology, acting on infix-less morphological representations and obeying constraints of its own, that masterminds overt variation (and demands in-depth typological-cum-diachronic examination).

3.6. In the case of infixation the ordering discrepancy reflects the impact of phonology (concrete representations) on morphology (abstract), each with rationales of their own. In the final example – stacked attributive adjectives preceding or following a noun – both abstract and concrete order are of the same kind, namely syntactic. Nonetheless, variability of concrete order and invariance of abstract order can arguably be traced to different, and not necessarily harmonic, influences on syntax, referential semantics and information structure.

When adjectives of different semantic classes are to be combined with a noun in attributive constructions, two decisions are to be made: first, to put the adjectives (all or some) before or after the noun; second, how to order the adjectives among each other. (Adjectives are property-concept words with a grammar of their own, distinct from the grammars of nouns and verbs. Where adjective is not a distinct word class, property-concept words of a nominal or verbal nature should show similar positional proclivities.) In languages where the ordering is relatively rigid at phrase level, the first decision is usually clear-cut; while the second tends to be less categorical, there usually are clear pref-

62 *Frans Plank*

erences, as illustrated for only three semantic adjective classes in the following examples:

- |     |     |                                     |                      |
|-----|-----|-------------------------------------|----------------------|
| (1) | a.  | English et al.                      |                      |
|     |     | <i>a beautiful big red ball</i>     | VALUE SIZE COLOUR N  |
|     | b.  | Bahasa Indonesia et al.             |                      |
|     |     | <i>bola merah besar tjantik</i>     | N COLOUR SIZE VALUE  |
| (2) | a.  | Maltese et al.                      |                      |
|     |     | <i>ballun sabiħ kbir aħmar</i>      | N VALUE SIZE COLOUR  |
|     | a.' | Italian et al.                      |                      |
|     |     | <i>una bella grande palla rossa</i> | VALUE SIZE N COLOUR  |
|     | b.  | unattested (or so it seems)         |                      |
|     |     | <i>*redd bigg beautifull ball</i>   | *COLOUR SIZE VALUE N |

In establishing such unmarked orderings among adjectives, all kinds of considerations have to be borne in mind: the language has to permit the stacking of adjectives to begin with (some/many don't, or only do so very reluctantly); the several adjectives should not be (asyndetically) coordinated (which is the only option in non-stacking languages); adjectives should be modifying the noun rather than each other (*a beautifully red ball*); there should be no intonation breaks between the adjectives suggesting non-cohesiveness of the construction; sometimes inherently emphatic adjectives (especially 'big') push forward; there may be further factors relevant for ordering, such as the length or other phonological properties of adjectives; further semantic classes of property concepts should be taken into consideration, such as weight, age, material, provenance, and purpose, tending to be increasingly closer to the noun (*a beautiful big heavy new red woollen Swiss medicine ball*).

With these provisos, this descriptive generalisation, richly supported in the relevant literature (most substantially in Hetzron 1978 and Sproat & Shih 1990), can be made over (1a/b), which are by far the most common orderings across languages:

- (i) The relative distance (position class) of semantic classes of adjectives from the noun is the same, whether the noun comes after or before the adjectives. (That is, the two most common orderings are mirror images of each other.)

A higher-level descriptive generalisation over (1a/b) is as follows, drawing on conceptual differences of words subsumed under one word class and of conceptual similarities across word classes:

- (ii) The nounier a modifier, the closer to the noun.

The nouniness ranking COLOUR > SIZE > VALUE is independently motivated, on language-particular and general grounds. Well-known relevant evidence includes: the nounier modifier words actually are themselves nouns or are derived from nouns (e.g., *wooll-en* MATERIAL, *medicine* PURPOSE); they can enter a morphological relation with head nouns (compounding: e.g., *snow-ball* MATERIAL, *black-ball* COLOUR); their range of applicability to nouns of different semantic classes is narrower (e.g., *beautiful*/\**red idea* VALUE/\*COLOUR in construction with abstract noun). The less nouny modifier words are verbal forms (participles) or are derived from verbs (e.g., *ugly* VALUE, originally deverbal, borrowed from Old Norse *ugga* ‘to dread’); they do not compound; their range of application is wider.

The nouniness feature mentioned last suggests a generalisation over the pattern at issue at an even higher level, where description metamorphoses into explanation. As modifiers are stacked, their natural stacking-order, hierarchical rather than linear, reflects the conceptual closeness or distance of modifiers from ultimate heads: less nouny property-concept modifiers prefer to take scope over nounier concepts. Thus, unmarked linear order is iconically motivated by preferences of scope construal:

- (iii) Linear closeness – VALUE SIZE COLOUR N / N COLOUR SIZE VALUE – mirrors scoping hierarchy – (VALUE (SIZE (COLOUR (N)))) – as itself determined by conceptual distance.

This is the sort of fundamental principle that one would like to be able to invoke as a general constraint on the construction of wholes from meaningful parts, and in particular their arrangement. And iconicity is an undoubted major force in universally governing linear order in a wide range of syntactic domains.<sup>18</sup>

It follows from this account that the anti-iconic ordering in (2b) should not occur, and it apparently does not. But neither should the equally anti-iconic ordering (2a), which does, even if not so frequently. In view of the existence of (2a), the obvious question is: Why is there no mirror image of (2a), i.e., (2b)? And more alarmingly, the question is whether a prized universal, stated at whatever level of generality, as in (i), (ii), or (iii), is invalidated by the overtly anti-iconic ordering in (2a).

The universal is rescued, and (2a)’s lack of a mirror image is explained, if Maltese et al. (with Semitic and Celtic languages as *alii* on record), instantiating the surface ordering in (2a), are analysed as being like English et al.: as having NPs where N is in final position. This similarity, evidently, can only

---

18. It remains to be seen whether the several other semantic categories claimed by Cinque (2006) to be also involved in mirror-image orderings around nouns can be subsumed under this same generalisation.

hold at an abstract level of syntactic representation – at a level where linear order is dictated by scope construal determined by conceptual proximity, only concerned with rendering referential meaning and unencumbered by any other expressive responsibilities. Thus, as to the relative ordering among multiple adjectives, iconicity rules OK everywhere, and timelessly, and the only variable here is whether modifiers come before nouns (1a, 2a/a') or after (1b, 2a').

The price to pay for such an abstract analysis is a syntactic rule of N-fronting (2a), or half-way fronting as in (2a'), exemplified by Romance, tampering with abstract order.<sup>19</sup> And the question, and typological research project, that comes with it is why only a few languages front or half-way front N, while many languages leave N where it is. When such discrepancies between concrete and abstract syntactic order are accepted, yet another question needs to be addressed, namely, why there are no abstract representations in line with iconicity which end up with a counter-iconic concrete order through N-BACKING – that is, with manifest (2b) derived from abstract (1b). Some explanatory mileage might be gotten out of the particular directional asymmetry in this respect where grammars are variable – displacing or not displacing N; but if displacement, then only by fronting, never by backing. Ordering under the iconic supervision of referential semantics can be interfered with as the information to be presented in context is being structured, with the requirements of topic-comment or focus-background articulation potentially advising that abstract scope-iconic order better be deviated from. Perhaps, as modifiers will naturally be comments, the displacement of N in NPs is to be understood as some sort of topicalisation – and, universally, topics tend to be overtly fronted, not backed. As to the question of why languages differ in letting or not letting nouns be displaced to begin with, another lead that remains to be pursued typologically is that languages with N-fronting are ones where V is initial, too (as in Semitic and Celtic), at least (arguably) in abstract representations.<sup>20</sup> Should there really turn out to be such a connection between noun phrases and verb phrases/clauses, making theoretical sense of it might require representations where noun phrases and clauses are seen to share constructional principles that are perhaps impossible to express in concrete representations.

Abstract N-finality, at variance with concrete N-initiality, may seem hard for acquirers of grammars to infer from the concrete forms they get to hear in their speech communities. Maltese is an instructive case where relevant experience seems to be altogether lacking. In Maltese, adjectives are as a rule postnominal.

---

19. Following Cinque (1994) and Longobardi (1994), who took their inspiration from Romance, N-fronting has been much discussed in generative syntax, with somewhat unequal attention paid to the technical implementation and the typological milieu of such a rule.

20. Inflectional differences in the marking for gender and number have also been claimed to be implicated; but this seems factually more dubious.

A few adjectives do precede their nouns, but only in highly marked formal and stereotyped constructions: e.g., *il-kbir Alla* DEF-big god 'God Almighty'; *Għażiż Pawlu* 'Dear Paul', the address formula in letters; *l-gharef hu-k* DEF-wise brother-2SG 'your clever brother' (only ironic). Synthetic comparatives, themselves postnominal, are placed ahead of their noun to express the superlative grade (with one definiteness marker suppressed): e.g., *il-belt il-qawwi-ja* DEF-city DEF-beautiful-FEM 'the beautiful city', *belt aqwa* city beautiful.COMP, 'a more beautiful city', *l-aqwa belt* DEF-beautiful.COMP city 'the most beautiful city'. Analytic comparatives/superlatives as such stay behind; only the comparative marker on its own may be in front of the noun (again, with one definiteness marker dropped): e.g., *il-ktieb l-aktar sabiħ* DEF-book DEF-more beautiful 'the more/most beautiful book', *l-aktar ktieb sabiħ* DEF-more book beautiful 'the most beautiful book'. And, being in some ways similar to adjectival modifiers, numerals other than 'one' also precede nouns (which are in the singular with numerals above 'ten'): e.g., *għoxrin suldat* twenty soldier (SG) 'twenty soldiers'. Whatever the right "thick description" (concrete rather than abstract in Borg & Azzopardi-Alexander 1997, in *LDS* format), it seems doubtful that the manifest ordering of nouns in very marginal modifier constructions could suffice to induce acquirers to set up abstract N-final representations at odds with manifest ordering in ordinary constructions encountered much more frequently.<sup>21</sup> Nonetheless, N-finality appears to be forced on acquirers of Maltese, by achronic law, and the option of N-fronting appears to be embraced with it, for when attributive adjectives are to be stacked (assuming they can be stacked in simple NPs, with (2a) above not a juxtaposition of three or four NPs: 'a ball, a beautiful [one], a big [one], a red [one]'), they unerringly follow the order which needs the noun to be final, at the (abstract) level where representation is iconic.

In sum, the price for abstract syntactic representations seems well worth paying in the case of adjective stacking, especially when your business is to tease apart, and explain, what is variable and what invariable across languages and to ascertain how variables co-vary.

*Received: 23 February 2007*

*Universität Konstanz*

*Revised: 5 March 2007*

---

21. In an analogous case, it is perhaps easier to infer for acquirers of German that abstract representations should be uniformly verb-final, thus also for main clauses where finite verbs are overtly in second position: there would seem sufficient evidence for verb-finality being recognisably unmarked, with verb-final clauses lacking all sorts of things which make a clause "main".

*Correspondence address:* Sprachwissenschaft, Universität Konstanz, 78457 Konstanz, Germany;  
e-mail: frans.plank@uni-konstanz.de

*Acknowledgements:* As usual, what is being written here has been said before, on more occasion than one, in private and in public; and on some points I may even be repeating myself writing (to correct, augment, generally improve). For listening and arguing and for asking and answering questions, I'm grateful to private interlocutors (especially Albert Borg, Guglielmo Cinque, Hans den Besten, Larry Hyman, Barış Kabak, Aditi Lahiri, and Edith Moravcsik) and public audiences (especially at the LSA Typology Aftershop, Oakland, Cal., January 2005; at an SFB Colloquium, Konstanz, December 2005; at the Bologna Universals Conference, January 2007; and at the University of Malta, February 2007). In part my work over the past years has been supported by the Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 471 (Projects A11 and A20).

## References

- Aristar, Anthony R. (1991). On diachronic sources and synchronic patterns: An investigation into the origin of linguistic universals. *Language* 67: 1–33.
- Blevins, Juliette (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Borg, Albert & Marie Azzopardi-Alexander (1997). *Maltese*. (Descriptive Grammars.) London: Routledge.
- Campanella, Tommaso (1638). *Philosophiæ rationalis partes quinque. Videlicet: Grammatica, dialectica, rhetorica, poetica, historiographia, iuxta propria principia. Suorum operum tomus I*. Paris: Ioannes du Bray. [Included, without *Dialectica*, in *Tutte le opere, I: Scritti letterari*, ed. by L. Firpo. Milano: Mondadori, 1954.]
- Capell, Arthur (1965). A typology of concept domination. *Lingua* 15: 451–462.
- Cinque, Guglielmo (1994). Evidence for partial N-movement in the Romance DP. In Guglielmo Cinque et al. (eds.), *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne*, 85–110. Washington, D.C.: Georgetown University Press.
- (2006). Mapping nominal functional structure: Some DP/CP (non-)parallelisms. Paper at the conference *DP-internal Information Structure*, Utrecht, 17–18 November.
- Comrie, Bernard & Norval Smith (1977). *Lingua Descriptive Studies: Questionnaire*. *Lingua* 42: 1–72.
- Crowhurst, Megan J. (1998). *Um* infixation and prefixation in Toba Batak. *Language* 74: 590–604.
- Dryer, Matthew S. (1988). Object–verb order and adjective–noun order: Dispelling a myth. *Lingua* 74: 185–217.
- Flores, Maria-Bonita (2004). The role of catalexis in Spanish rhythm structure: A phonological and phonetic study of catalexis in Spanish syllable-timed poetry. Magister thesis, Lunds universitet. [http://www.ling.lu.se/education/essays/MariaBonitaFlores\\_D.pdf](http://www.ling.lu.se/education/essays/MariaBonitaFlores_D.pdf)
- Gabelentz, Georg von der (1891). *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel.
- Ghini, Mirco (2003). From prosody to place: The development of prosodic contrasts into place of articulation contrast in the history of Miogliola. In Paula Fikkert & Haike Jacobs (eds.), *Development in Prosodic Systems*, 419–455. Berlin: Mouton de Gruyter.
- Green, Thomas Michael (1999). A lexicographic study of Ulwa. Doctoral dissertation, Massachusetts Institute of Technology. <http://www.slaxicon.org/papers/thesis.pdf>
- Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful constituents. In Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, Mass.: MIT Press.
- (1990). *On Language: Selected Writings*. Edited by Keith Denning & Suzanne Kemmer. Stanford: Stanford University Press.
- Hetzron, Robert (1978). On the relative order of adjectives. In Hansjakob Seiler (ed.), *Language Universals*, 165–184. Tübingen: Narr.



*Extent and limits of linguistic diversity as the remit of typology* 67

- Hyman, Larry M. (2006). Affixation by place of articulation: Rare and mysterious. Paper at the conference on *Rara*, Leipzig, 29 March – 1 April.
- Kager, René (1995). Consequences of catalexis. In Harry van der Hulst & Jeroen van de Weijer (eds.), *Leiden in Last: HIL Phonology Papers 1*, 269–298. Den Haag: Holland Academic Graphics.
- Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.
- Kiparsky, Paul (1991). Catalexis. Unpublished manuscript, Stanford University.
- Lahiri, Aditi & Henning Reetz (2002). Underspecified recognition. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 637–674. Berlin: Mouton de Gruyter.
- (2007). Distinctive features: Signal to representation. In G. N. Clements (ed.), *Phonetic Bases for Distinctive Features*, to appear.
- Longobardi, Giuseppe (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25: 609–665.
- McCarthy, John J. (1982). Prosodic structure and expletive infixation. *Language* 58: 574–590.
- Mesgnien, François de (1680). *Thesaurus linguarum orientalium turcicae, arabicae, persicae, praecipuas earum opes à Turcis peculiariter usurpatas [...]*, with a separate fourth volume *Linguarum orientalium turcicae, arabicae, persicae institutiones seu Grammatica turcica [...]*. Windobonae: Operâ, typis, & sumptibus Francisci à Mesgnien Meninski. [The *Grammatica* re-edited in two volumes by Adam Franz Kollar as *Institutiones linguae turcicae, cum rudimentis parallelis linguarum arabicae & persicae [...]*. Windobonae: Ex typographico orientali Schilgiano, 1756. The whole *Thesaurus* re-edited by Bernhard von Jenisch, Franz von Klezl, & Jacob von Wallenberg in four volumes as *Lexicon arabico-persico-turcicum [...]*. Wien: Kurzböck, 1780–1802. Original edition reproduced in facsimile, edited by Mehmet Olmez and with an introduction by Stanisław Stachowski, 6 volumes, Istanbul: Simurg, 2000.]
- Mohanan, K. P. (1993). Fields of attraction in phonology. In John Goldsmith (ed.), *The Last Phonological Rule*, 61–116. Chicago: University of Chicago Press.
- Moravcsik, Edith A. (1977). On rules of infixing. Bloomington: Indiana University Linguistics Club.
- (2000). Infixation. In Geert Booij, Christian Lehmann, & Joachim Mugdan (eds.), *Morphology: An International Handbook on Inflection and Word-Formation*, Volume 1, 545–552. Berlin: Walter de Gruyter.
- Nichols, Johanna (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Paradis, Carole & Jean-François Prunet (eds.) (1991). *The Special Status of Coronals: Internal and External Evidence*. San Diego: Academic Press.
- Plank, Frans (1999). Split morphology: How agglutination and flexion mix. *Linguistic Typology* 3: 279–340.
- (2001). Typology by the end of the 18th century. In Sylvain Auroux et al. (eds.), *History of the Language Sciences: An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present*, Volume 2, 1399–1414. Berlin: Walter de Gruyter.
- (2003a). Noun phrase structure: *An und für sich*, in time, and in space. In Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, 3–33. Berlin: Mouton de Gruyter.
- (2003b). There's more than one way to make sense of one-way implications – and sense they need to be made of. *Linguistic Typology* 7: 128–139.
- (2007). *Early Typology*. To appear.
- Plank, Frans, Elena Filimonova, & Thomas Mayer (eds.) (2007). THE UNIVERSALS ARCHIVE. <http://typo.uni-konstanz.de/archive>
- Plank, Frans & Wolfgang Schellinger (2000). Dual laws in (no) time. *Sprachtypologie und Universalienforschung* 53: 46–52.
- Rice, Keren (2000). *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*. Cambridge: Cambridge University Press.

68 *Frans Plank*

- Slobin, Dan I. (ed.) (1985–97). *The Crosslinguistic Study of Language Acquisition*. 5 volumes. Hillsdale, then Mahwah, N.J.: Erlbaum.
- Sproat, Richard & Chilin Shih (1990). The cross-linguistic distribution of adjective ordering restrictions. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary Approaches to Language: Essays in Honor of Sige-Yuki Kuroda*, 565–593. Dordrecht: Kluwer.
- Yu, Alan C. L. (2004). Reduplication in English Homeric infixation. *North-Eastern Linguistic Society* 34: 619–633.