Abstract

This paper explores the concept of typological explanation, with particular reference to iconicity. It is suggested, first, that iconicity should be clearly distinguished from the (language-internal) principle of ‘one meaning – one form’ (also known as ‘isomorphism’) and, second, that iconicity with no clear ontological support should be de-emphasized. With these caveats, iconicity is shown to possess genuine explanatory force.

1. The 1M1F Principle vs. Iconicity

Grammatical morphemes tend to exhibit what Anttila (1989 [1972]) calls the ‘principle of one meaning – one form’ (to be henceforth abbreviated as the ‘1M1F principle’). Iconicity, in turn, is generally understood as structural similarity — a ‘picture-like’ relationship — between extralinguistic reality and language. These two notions should be kept apart.

1.1. The 1M1F Principle

The deviations from the 1M1F principle may be summarized as follows:

<table>
<thead>
<tr>
<th>Syntagmatic</th>
<th>Paradigmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) 2M1F = portmanteau morpheme</td>
<td>c) 2M1F = grammatical ambiguity</td>
</tr>
<tr>
<td>b) 1M2F = discontinuous morpheme</td>
<td>d) 1M2F = allomorphy</td>
</tr>
<tr>
<td>non-1M1F</td>
<td></td>
</tr>
</tbody>
</table>

Type a) represents the case where one form has (simultaneously) more than one meaning; e.g. the suffix in the Latin *glob-us* (‘ball’, ‘group’) expresses the three meanings ‘masculine’, ‘singular’, and ‘nominative’. Type a) also includes the case where a zero morpheme has at least one meaning (often exemplified by the first member in such pairs as ‘singular vs. plural’, ‘nominative vs. accusative’, ‘absolutive vs. ergative’). It follows that, as an instance of the (paradigmatic) type c), the zero may in a given language be massively ambiguous. Type a) is also known under such (slightly misleading) designations as ‘fusion’ and ‘cumulation’. Type b) represents the case where one meaning is (simultaneously) expressed by more than one form; e.g. in the Swedish *den röd-a stol-en* (‘the red chair’) the meaning ‘definite’ is expressed by the three underlined units, as opposed to *en röd stol* (‘a red chair’), where the meaning ‘indefinite’ is expressed in the 1M1F fashion.

While syntagmatic relations are instances of ‘both – and’, paradigmatic relations are instances of ‘either – or’. Type c) is exemplified by the Latin *glob-is*, where the suffix
expresses either the meaning ‘dative (plural)’ or the meaning ‘ablative (plural)’, but not of course both simultaneously. Wherever there are several declensions or conjugations, type d) is exemplified; for instance, ‘nominative singular’ is expressed in Latin not only by the allomorph -us, as in glob-us, but also by the allomorph -a, as in puell-a (‘girl’).

In the domain of grammatical morphemes, 1M1F represents the ideal situation, and it is traditionally used to explain analogical changes that occur within inflectional paradigms, i.e. changes where either 2M1F or 1M2F becomes 1M1F (cf. Paul 1975 [1880]: 106–120, 190–199; Anttila 1989 [1972]: Ch. 5; Itkonen 1982: 104–112, 1983: 208–211). In the domain of lexical morphemes, by contrast, it is not reasonable to consider 1M1F as an ideal to be sought after. Here, types c) and d) are known as polysemy (~ homonymy) and synonymy, respectively. Polysemy or 2M1F, rather than monosemy or 1M2F (let alone synonymy or 1M2F) clearly represents the typical case.

It is significant that while type a) is well-established in the grammatical domain, it has no natural counterpart in the lexical domain. The grammatical morpheme -us, when uttered as part of the Latin word globus, must always have the three meanings ‘masculine’, ‘nominative’, and ‘singular’. But the lexical morpheme glob-, when uttered as part of the word globus, has only one meaning at a time (for instance, either ‘ball’ or ‘group’, but not both). The exceptions to this rule can easily be recognized as non-typical, as when uttering He didn’t grasp her the speaker intends to convey both the concrete or literal meaning and the abstract or metaphorical meaning. The lexical counterpart of type b) is exemplified by expressions like John looked her up. In the lexical domain, however, there seems to be a tendency for the parts of the discontinuous form to be asymmetric, whereas in the grammatical domain they are symmetric.

In the preceding paragraph, an argument was given for the traditional grammatical vs. lexical distinction, whose usefulness is denied e.g. in today’s so-called cognitive linguistics. An additional argument is provided by the fact that in all languages lexical meanings like ‘child’ are expressed in the same way, namely by the simple form – meaning correlation, whereas grammatical meanings like ‘plural’ are expressed in typologically dissimilar ways, namely by suffixation, prefixation, reduplication, internal change (Ablaut), or zero.

Up to now, grammatical meanings have been portrayed as being expressed by single morphemic forms. However, many grammatical meanings (e.g. ‘passive’) are typically expressed by (multi-partite) constructions which go beyond single words. This illustrates the syntactic counterpart of type b). In syntax, type c) is illustrated e.g. by the formal similarity, in some languages, between topic constructions and conditional clauses, or between questions and conditional clauses. Type d) — called ‘allotaxy’ — is uncommon, but may be exemplified, in English, by such more or less equivalent cases of sentence negation as I didn’t see anybody and I saw nobody.

1.2. Iconicity

Iconicity is based on comparing two independent entities A and B, and on deciding that there exists enough of structural similarity between A and B for A to be called a ‘picture’ of B (or vice versa). Suppose that on a specific day in a specific region there was first rain and then sunshine. This chain of events may be expressed either by After it rained, the sun shined or Before the sun shined, it rained. It is easy to see that the former sentence is an iconic representation of the extralinguistic state of affairs spoken about. Why? Because its constituent clauses exemplify the same (temporal) order as the events spoken about (= first rain, then sunshine). For the corresponding
reason, the latter sentence is a non-iconic representation of what happened. It is important to realize that these judgments can be made quite objectively. That is, there is an intersubjective agreement on the truthfulness (or otherwise) of such judgments. And there is no logical circle involved in reaching such judgments.

The 1M1F principle does not qualify as a subtype of iconicity, for the simple reason that, in a given language, the meanings of words (or morphemes) do not exist independently of the corresponding forms. Rather, meanings and forms are interdependent. This fact was accurately captured by Saussure’s notion of ‘linguistic sign’. Notice that in the example of the previous paragraph we were not comparing forms with meanings. We were, rather, comparing sentences (= form-and-meaning entities) with extralinguistic states of affairs, with the qualification that (as is always the case in iconic comparison) the emphasis was on the sentence-forms, not on sentence-meanings.

The concern with iconicity is ultimately justified by the wish to explain linguistic structure. Such explanations can be conceived of as constituting the following type of taxonomy:

“Ontological explanations refer to the way in which the structure of extralinguistic reality is reflected in linguistic structure, producing a relation of structural iconicity (or isomorphism) between the two. It goes without saying that there can be no ‘pure’ ontology; rather, each ontology is already a result of conceptualization.

Cognitive explanations refer to the way in which a human being relates himself to what is ontologically definable. Precisely because there is no pure ontology, the difference between ontology and cognition, though real, can only be an approximative one. In a situation like this, it is important first of all to establish the clear cases. That one event temporally precedes another, is an ontological fact, although both ‘event’ and ‘temporal precedence’ are certainly results of human concept-formation. By contrast, either denying or inferring the occurrence of an event is a cognitive fact (more precisely, a cognitive operation), because, instead of being part of the extramental reality, it is directed toward this reality (more precisely, applies to mental representations of this reality).

Finally, social explanations refer to the interaction between human beings relating themselves to what is ontologically definable” (Itkonen 1994: 38).

As was noted in this quotation, prototypical iconicity is based on comparing ontology and language. As far as the distinction between ontology and cognition is concerned, the following rule of thumb is quite useful. Whatever can be pictured or photographed, is ontologically given, and whatever cannot, is of non-ontological character. Consider the sentence Demi Moore is combing her hair. Wittgenstein has taught us that pictures are not self-explanatory but need rules of interpretation. With this caveat in mind, however, it is surely possible to make a picture that can be plausibly claimed to describe the state of affairs that is truthfully described by the above-mentioned sentence.

Now consider the sentence Demi Moore is not combing her hair. Clearly, it is not possible to draw any single picture or take any single photograph that could with any plausibility be claimed to describe the state of affairs that makes this sentence true. Indeed, according to the extreme ‘logicist’ view, any state of affairs, except the one consisting in Demi Moore combing her hair, would make this sentence true. This example shows that, instead of referring to some ontological entity that is part of the external world, the negation-sign expresses a cognitive operation performed by the
speaker. To put it succinctly, negation does not exist in the world.

Next, consider the sentence *Demi Moore is my favourite actress*. Again, it is not possible to give a picture of the state of affairs that makes this sentence true, and the reason is the same as above. By means of the copula *is*, this sentence expresses an act or operation performed by the speaker, namely an act of identification. And this kind of act, just like the act of negation, does not exist in the world, in the sense of being an entity that can be pictured or photographed. To summarize, negation and identification are cognitive acts or processes with no clear ontological correlates.

From what precedes, it follows, first, that ontology and cognition must be kept apart and, second, that iconicity should be based on ontology. The situation is more complicated than that, however. Remember that iconicity was stated to be based on a comparison between two independent entities A and B. Now, to the extent that cognition can be shown to be independent of language, it *is* meaningful to speak of iconicity between language and cognition, after all. The study of infant cognition seems to indicate that much of the conceptual apparatus, including such cognitive operations as negation and identification, is in place already at the age of 12 months or so, i.e. before the emergence of language. This issue remains a controversial one, however, and therefore it is wise to stick to the ontological variety of iconicity as much as possible (cf. below).

The social variety of iconicity seems even more doubtful than the cognitive one. For instance, no iconic principle seems to govern the way that expressions of old vs. new information are distributed in discourse. If metaphorical use of iconicity is allowed, however, the situation may turn out be different (cf. below).

Ontological iconicity may be subclassified in different ways. The relevant dimensions include at least temporal order, cohesion/distance, and quantity. The first of these was briefly illustrated above. The relevance of *cohesion* is summarized in the slogan *(the expression of) what belongs together goes together*: ontological wholes and/or hierarchies are reflected in linguistic structure. Suppose I utter a true sentence *There is a blond woman in the park walking a black dog*. The only reason why *blond* occurs close to *woman* and *black* close to *dog*, and not vice versa, is that in the language-external reality as I perceive it the qualities *’blond’* and *’woman’*, on the one hand, and *’black’* and *’dog’*, on the other, occur together. When Jackendoff (1987) argues that both the extralinguistic reality and language exemplify the same type of structure, or what he calls *’headed hierarchy’*, he is just reasserting the existence of cohesion-based iconicity.

The relevance of ontological *distance* becomes evident e.g. in the domain of causation (where the expression becomes less compressed the more indirect the causation is) or in the expression of inalienable possession (= short) vs. alienable possession (= long). At the metaphorical level, an increasing degree of social distance between the speakers produces an increasing length of expression (= so-called politeness forms).

In the simplest case, ontological *quantity* is linguistically reflected in the difference between singular (= ontologically and linguistically less) vs. plural (= ontologically and linguistically more). The distinction between existence and non-existence may be viewed as a subtype of ontological quantity. An ontological explanation involving this distinction will be discussed in Sect. 2.

In all these cases, we are entitled to speak of ontological iconicity in the *strong* sense insofar as the notions of temporal order, cohesion/distance, and quantity are the *same* in ontology and in language: just by looking at linguistic expressions we can tell which ontological alternatives they describe (i.e. prior vs. posterior, close vs. distant,
less vs. more). We can, however, also speak of ontological iconicity in the weak sense. For instance, the near-universal noun vs. verb distinction is clearly based on a corresponding ontological difference (cf. Itkonen 2001). But the qualitative difference between nouns and verbs is not supported by the same ontological quality: it is in general not possible, just by looking at two words, to tell which one is a noun and which one is a verb. Interestingly, this is not true of sign languages. In the Finnish Sign Language, for instance, the same sign (= digging motion) expresses the two meanings ‘(a/the) spade’ and ‘to dig’, but in the former case the motion is more restricted than in the latter.

Let it be repeated that, in what precedes, the terms ‘ontology’ and ‘cognition’ are convenient shorthands for ‘ontology-with-cognition’ and ‘cognition-without-ontology’. If taken literally, it does not make sense to speak of a distinction between ‘ontology’ and ‘cognition’.

Finally, it is good to point out that ‘ontological explanation’ is necessarily a shorthand for a more complex notion. Taken as such, ontology lies inertly there and does not explain anything. It becomes an explanatory notion only when it is embedded in the human need and/or wish to communicate. Apparently, structural similarity between what is to be communicated and the means of communicating it is conductive to clarity and ultimately facilitates communication.

1.3. Discussion: Croft (2003) on ‘Iconicity’

Haiman (1985) distinguishes between two subtypes of ‘iconicity’, namely ‘isomorphism’ and ‘(diagrammatic) motivation’. Of these two, the former is identical with our 1M1F principle whereas the latter contains what has been characterized here as iconicity proper. Croft (2003) continues the same terminological usage, with the qualification that he divides ‘iconicity’ into ‘isomorphism’ and ‘diagrammatic iconicity’. This inevitably leads to some confusion.

Typological linguistics deals with the variation in how particular languages have formally expressed the conceptual space common to all languages. Croft (2003: 134) opposes the conceptual space to ‘semantic maps’ each of which represents the choice made by one particular language. This distinction is paralleled by the distinction between universally valid ‘concepts’ and language-specific ‘meanings’. Consequently enough, Croft speaks of ‘meanings’ (and not of ‘concepts’), when discussing ‘isomorphism’ (alias 1M1F principle) (pp. 102–110). Afterwards, however, he claims (p. 202) that ‘isomorphism’ deals with ‘concepts’ (and not ‘meanings’), which is a mistake. He correctly divides ‘isomorphism’ into a syntagmatic variety (pp. 102–105) and a paradigmatic variety (pp. 105–110). In the paradigmatic context he inexplicably confines his discussion to the meanings (and forms) of words: “The second type of isomorphism is the correspondence between form and meaning in the inventory of words stored in the mind; this is called paradigmatic isomorphism” (p. 105). As a consequence, he refers to our type c), i.e. paradigmatic 1M2F, ‘synonymy’, but this term must be taken to subsume also allomorphy and allotaxy.

Croft correctly points out that an explanatory relationship can only exist between two independent entities: “Thus, the types of universals that one can identify through cross-linguistic comparison are universals of the relationship between linguistic form (morphosyntactic or phonological) and external function or reality. This is the sense in which the typological approach to grammar is functionalist. [...] The explanations described in the preceding chapters lies [sic] in the relationship between linguistic form
and external function” (p. 194; the second emphasis in the original). We have already noted, however, that if X is a form-meaning entity (like a morpheme, word, or sentence), the relation of X’s meaning to X’s form is too intimate to be genuinely explanatory. Ontological structures can explain corresponding linguistic structures, but meanings cannot explain corresponding forms. Therefore, to repeat, it is wrong to view the 1M1F principle as a subtype of iconicity. It is a different matter that, as noted above, the 1M1F principle can be causally effective in analogical changes, insofar as a particular instantiation of this principle constitutes the goal which speakers are unconsciously entertaining and which they try, unconsciously, to achieve by changing their language accordingly.

Iconicity and the 1M1F principle are not just distinct but, occasionally, also contradictory notions. This is graphically shown by what Croft considers the ‘iconic motivation’ of the singular vs. plural marking. Substituting the 1M1F principle for iconicity, he assumes that the SG vs. PL distinction is ‘iconically motivated’ in those languages where either both categories are expressed by non-zero morphemes (= Zulu) or neither is (= Minor Mlabri): “In the Zulu example, for instance, each conceptual category, both singular and plural, is overtly encoded in the word form. Iconicity motivates symmetry in the grammatical expression [...] The Minor Mlabri case is also symmetric in its own way, in that it expresses neither value for number [...]” (p. 102). This amounts to denying the existence of iconicity based on ontological quantity, which is generally thought to be a prototypical instance of iconicity.

‘Diagrammatic iconicity’ (alias iconicity proper) is discussed by Croft on pp. 201–219. Considering the central role that iconicity seems to play in Croft’s overall view of typological linguistic, the conclusion that he reaches is surprisingly pessimistic:

“The primary difficulty in evaluating arguments in favor of iconicity is that the structure of what is signified, experience as we put it in §4.2, is not well established. [...] At this point, one faces [...] the shortage of firm evidence for the structure of experience outside of language itself. [...] However, this argument can be turned on its head. The iconicity of human language is a hypothesis that can be used to propose hypotheses of cognitive structure” (p. 203).

Is Croft’s pessimism justified? Of course not. Why not? Because, failing to distinguish between ontological iconicity and cognitive iconicity, he sees all of iconicity as tainted by those doubts that rightly attach to the notion of cognitive iconicity. However, to use our favourite example once again, it would be irrational to doubt that we can objectively compare the temporal order of extralinguistic events and that of uttered (i.e. spoken or signed) sentences.

Croft’s over-extensive notion of iconicity makes it possible for him to subsume not just cognitive explanations, but also “pragmatic” (p. 65), i.e. (ultimately) social, explanations under ‘iconicity’. Consider the following three (“pragmatic”) principles, enumerated on p. 66 and again on p. 202, that are claimed to explain word order:

1. What is old information comes first, what is new information comes later, in an utterance.
2. Ideas that are closely connected tend to be placed together.
3. What is at the moment uppermost in the speaker’s mind tends to be first expressed.

According to Croft (p. 202), all three principles are iconic, even if the iconicity of principle 3 is “metaphorical” rather than “direct”. Now, “by principle 1, new information should follow old information” [whereas] “by principle 3, new information should
precede the old information” (p. 66). This self-contradictory result confirms our view, expressed above, that non-ontological iconicity is of secondary importance and should perhaps not be used at all. Notice also that principle 2 is not equivalent to ontological cohesion/distance. It is extralinguistic reality as perceived (and not ‘ideas’) that explains the word order in a sentence like *There is a blond woman in the park walking a black dog*.

To conclude, iconicity, if duly restricted, is a much more viable notion that Croft (2003) makes it out to be.

2. Iconicity and Analogy

Iconicity has been defined here as structural similarity between language and extralinguistic reality. It is well known that structural similarity is the general definition of *analogy*. It follows that iconicity is a special case of analogy. Some implications of this fact have been explored in Itkonen (1991: 5.5.3, 1994, and 1996).

Traditionally, words and sentences have been said to be analogous if they share a common structure. As exemplifications of this more abstract structure, they must be structurally similar to each other:

“New words may be consciously created from these fundamental elements on the analogy of old ones, but hardly new types of words. In the same way new sentences are being constantly created, but on strictly traditional lines. [...] For instance, these two sentences fit precisely the same pattern, [...] differing only in their material trappings: [...] *The farmer kills the duckling, The man takes the chick*” (Sapir 1921: 37).

“[...] we feel that these two sentences [i.e. *John gave Mary an apple* and *My uncle lent the joiner five shillings*] are analogous, that is, they are made after the same pattern. [...] Now, how do such [sentence] types come into existence in the mind of a speaker? [...] from innumerable sentences heard and understood [the child] will abstract some notion of their structure which is definite enough to guide him in framing sentences of his own” (Jespersen 1965 [1924]: 19).

“[...] the speaker who knows the constituents and the grammatical pattern, can utter [speech forms] without ever having heard them; [...] A grammatical pattern [...] is often called an analogy” (Bloomfield 1933: 275).

Let us note in passing that in Sapir’s, Jespersen’s, and Bloomfields accounts the notion of *pattern* plays the same role that the notion of *construction* does in today’s ‘construction grammar’ (cf. Goldberg 1995).

Occasionally, ‘isomorphism’ and ‘(diagrammatic) iconicity’ are interpreted widely enough to make them effectively coincide with the notion of analogy as it has traditionally been employed in linguistics and was illustrated in the preceding quotations: “In fact, the principle of isomorphism is the basis on which new compound signs such as phrases, clauses, and sentences are formed and is the reason why these new signs can be understood. Thus, diagrammatic iconicity is ubiquitous in language, and the only types of signs for which diagrammatic-isomorphic iconicity does not hold are non-compound signs (e.g., mono-morphemic words)” (Waugh 1994: 57).

Iconicity (as defined in 1.2) can be viewed as ‘vertical’ analogy, i.e. analogy between extralinguistic reality and language. By contrast, the traditional linguistic notion of analogy is an intralinguistic and hence ‘horizontal’ relationship. The functioning of vertical and horizontal analogy can be illustrated as follows.

According to Givón (1995: Ch. 4), it is meaningful to distinguish between two ‘megamodalities’, viz. ‘fact’ and ‘non-fact’, expressed by affirmative-indicative markings and negative and/or subjunctive markings, respectively. Identifying
subjunctive with ‘non-fact’ is unproblematic. By contrast, identifying negative too with ‘non-fact’ might seem somewhat problematical. From the logical point of view, $p$ and not-$p$ are considered as symmetrical, and therefore if the former expresses ‘fact’, it may be difficult to accept that the latter expresses ‘non-fact’. From the standpoint of the psychology of logic, however, $p$ and not-$p$ are not symmetrical. Rather than expressing a fact, a true negative sentence (in the indicative mood) is thought to express the “denial of a falsehood”, which is something more complicated than, and hence different from, a simple fact (cf. Evans 1982: 28). Thus, there is reason to accept the notion of ‘megamodality’.

In a given language the markings for the two megamodalities are either asymmetric or symmetric. Both cases can be explained by analogy and, what is more, only by analogy.

In the asymmetric case, we have to do with vertical analogy or iconicity: from the ontological/conceptual point of view, ‘fact’ is more differentiated than ‘non-fact’, and linguistic structure reflects this difference. The asymmetric way to express the two megamodalities is exemplified by Modern Tamil. The affirmative-indicative verb, which expresses ‘fact’, inflects in person/number/gender and in three (basic) tenses. By contrast, the negative-indicative verb has four distinct forms (= a, b, c, d), of which the first three are uninflected: (a) expresses any tense and any person/number/gender; (b), being based on nominalized present tense, expresses habituality and any person/number/gender; (c), being based on nominalized past tense, expresses the past and any person/number/gender; only (d), which expresses the future, inflects in person/number/gender (cf. Asher 1985: 175). Apart from the negation, the megamodality of ‘non-fact’ is expressed either by the conditional mood or by one of several modal auxiliaries. The corresponding verb-forms, whether affirmative or negative, never inflect in person/number/gender, and only rarely in tense (ibidem, pp. 165–166, pp. 181–182). — Although, as argued in 1.2, negation itself is not an ontological entity, or “does not exist in the world”, it is clear enough that the difference between existence and non-existence is of ontological nature and can legitimately be made use of in ontological-iconic explanations.

When the two megamodalities are expressed symmetrically, we have to do with horizontal (or language-internal) analogy: the structure of affirmative-indicative is extended to that of negative and/or subjunctive. This case is exemplified — approximatively — by Latin. Affirmative-indicative and negative-indicative are fully symmetric (the latter being expressed by negative particles, conjunctions, or pronouns). Both in the indicative and in the subjunctive, there is the same number of non-future tenses (= present, imperfect, perfect, pluperfect). However, in the subjunctive there are no counterparts to the two future tenses (which means that the symmetry between fact and non-fact is not complete). It is quite obvious that, semantically, the four tenses of the subjunctive have become more or less confused. The reason is, of course, that their ontological/conceptual motivation is less than that of the indicative tenses. That they are maintained nevertheless, must be due to the analogical influence exerted by the latter.

At this point, it may seem that analogy ‘explains too much’. But what would be the alternative explanation? There is none. To be sure, there remains the ulterior question why one language chooses the vertical analogy (= ontological motivation) rather than the horizontal analogy (= morpho-syntactic motivation).

3. Typological Explanation

3.1. Implicational Universals, Markedness, and Empathy
Typological research makes extensive use of so-called implicational universals (= IU’s) which are of the form ‘For all languages x, if x has A, then x has B’, or more simply ‘if A, then B’. In propositional logic the (‘material’) implication ‘if A, then B’ is given a truth-functional interpretation according to which it is false when A is true and B is false, and true otherwise. Corresponding to this implication, there is a ‘tetrachoric table’ that displays the four alternatives A & B, A & not-B, not-A & B, not-A & not-B. The use of IU’s is criticized in Itkonen (1998) for a number of reasons.

First, interpreting IU’s in the strictly truth-functional fashion leads to the view that non-falsification equals confirmation: ‘if A, then B’ is ‘confirmed’, except when A is true and B is false. For instance, claims about sign languages are ‘confirmed’ in each and every case when spoken language data make A false.

Second, ‘if A, then B’ is equivalent to its contrapositive ‘if not-B, then not-A’ in propositional logic. But assuming the former sentence to express a law, it is in general not the case that the latter sentence too expresses a law. Why not? Because, in general, both not-A and not-B stand for a set of several alternatives, and there can be no nomic or lawlike connections between such disparate conglomerations of facts.

Third, in the special case where A and B stand for binary properties, the corresponding IU’s can be thought to express a genuine law, for instance: ‘For all languages x, if x has a nonzero morpheme for the singular (= A), x has a non-zero morpheme for the plural (= B)’. Even here, however, the use of UI’s has some pitfalls. As was noted above, the UI in question is not falsified by the alternatives A & B, not-A & B, not-A & not-B, which are exemplified, respectively, by Zulu, English, and Minor Mlabri. But non-falsification does not equal confirmation. An implication is genuinely confirmed only by what makes both the antecedent and the consequent true. This means that, even here, ‘if A, then B’ and ‘if not-B, then not-A’ are not perfectly equivalent. The former is genuinely confirmed by Zulu, whereas the latter is genuinely confirmed by Minor Mlabri. But the curious thing is that English, which represents here the standard or ‘unmarked’ case (= zero in the singular, non-zero in the plural) genuinely confirms neither ‘if A, then B’ nor ‘if not-B, then not-A’. This shows that implication is an inaccurate form of representation to start with.

Fourth, if we demand that an accurate form of representation must present the standard or ‘unmarked’ case (= English) as primary and the non-standard or ‘marked’ cases as secondary, the following might be suggested:

(i) Primary needs of formal differentiation are satisfied before secondary needs (= English).

(ii) Therefore, if secondary needs are satisfied, primary needs are too (= Zulu).

(iii) And, inversely, if primary needs are not satisfied, neither are secondary needs (= Minor Mlabri).

This is just a general explanatory schema. As far as our example is concerned, item (i) is given empirical content by *quantity-based iconicity*: the asymmetry of the linguistic singular vs. plural marking (= ‘short vs. long’) is motivated by the ontological distinction ‘less vs. more’. (Remember that, according to Croft, it is Zulu and Minor Mlabri, and not English, which are here iconically motivated; cf. 1.3.) This schema meets the desideratum of presenting the typical case (i) as non-implicational and primary, and the non-typical cases (ii) and (iii) as implicational or entailed by the typical case. That is, (i) entails (ii) and (iii) just like ‘A precedes B’ entails both ‘If B happens, A has already happened’ and ‘If A has not happened, B has not happened either’. Here
we have the ‘markedness theory’ in a nutshell.

Notice that although (ii) and (iii) are logically equivalent, and equally entailed by (i), they are not synonymous. This is a good illustration of the limitations of truth-conditional semantics.

In our explanatory schema we referred to the speaker’s and/or the hearer’s needs. But how can we know what those needs are? This is where empathy comes into the picture.

‘Explanation’ is generally thought to be the central methodological notion. In today’s academic world it happens to be the case that some discipline A (e.g. quantum physics or evolutionary biology) is often regarded — rightly or wrongly — as more highly developed than some other discipline B (e.g. linguistics). It follows that would-be methodologists of B may try to borrow their own notion of ‘explanation’ from A and to impose it upon practitioners of B. For reasons explained at length in Itkonen (1978) and (1983), I prefer the opposite approach. What is or is not ‘explanation’ e.g. in typological linguistics, must be determined by the actual practice of typological linguists, i.e. what they themselves — in the middle of their daily work, as it were — call ‘explanation’.

In the present context, the following qualifies as a prototypical instance of typological explanation. Mithun (1988) notes that, cross-linguistically, there are relatively few instances of the conjoined-nominals construction. Rather than simply accepting this fact, she wants to explain it: “This rarity [of N-and-N constructions] is not altogether inexplicable: several factors converge to minimize the need for conjoined nominals in connected speech. Most important, speakers typically introduce only one major piece of information into discourse at a time. Conceptually distinct entities [referred to by nominals] are introduced by distinct intonation units, often separate clauses. [...] Once they have been introduced individually, sets of entities can be referred to by plural pronouns, so the need for conjoined noun phrases is bypassed” (p. 337; emphasis added; notice the connection with the ‘needs of formal differentiation’ mentioned in our explanatory schema above).

Let us next give an analysis of the preceding quotation, and more precisely an analysis formulated in terms of problem-solving:

Problem: Why are there, cross-linguistically, so few cases of N1-and-N2?

Solution: In general, referents of N’s are introduced in separate clauses: X&V1&N1, Y&V2&N2 (where X and Y represent arbitrary material, and V1 and V2 are verbs connected with N1 and N2, respectively); and later, if needed, the referents of N1 and N2 are re-identified by the pronouns PRO-1 and PRO-2, respectively. Now, if the referents of N1 and N2 have to be re-identified together, this is done by PRO-1&2 (meaning, roughly, ‘they’). Thus, at no stage is there any need for an expression like N1-and-N2.

Comment: This explanation is achieved by means of empathy, or by adopting the speaker’s perspective. The goal of the speaker is to re-identify two entities in a situation where they have already been introduced by N1 and N2. Now he is confronted with the following problem: Which means should he choose to achieve his goal? His solution is to choose PRO-1&2 (assuming that a corresponding pronominal system is available). There is no need for him to choose N1-and-N2. Why is there no need for N1-and-N2 when PRO-1&2 is available? Because the latter is a more economical means for achieving the goal than the former. Thus, what we have here is a rational (= means-to-end) explanation (since it is inherent to rationality to choose — ceteris paribus — the more economical alternative). The preceding account (rationally) explains the rarity of N1-and-N2 (when a pronominal system is available). Of course, the rationality involved
is of *unconscious* nature; it is the same type of rationality that governs not only human but also animal behavior. Our example shows that *empathy culminates in rational explanation*.

As a second example, let us consider *converb* constructions. Implicit-subject converbs and same-subject converbs tend to coincide, and so do explicit-subject converbs and different-subject converbs. This form – meaning correlation is explained by Haspelmath (1995) as follows:

“The functional motivation [i.e. explanation] for these connections should be apparent: when the subject is mainly implicit, only the same subject reference *ensures* that its subject can be identified. When the subject is necessarily different from superordinate clause constituents, only explicit expression *ensures* that its reference can be identified” (p. 11; emphasis added).

On the one hand, suppose you hear a sentence containing an implicit-subject converb like *Listening to the conversation, John felt embarrassed*. Who did the listening? It requires little imagination to for you realize that it must be John (i.e. that this is a same-subject construction), and the speaker too knew that you would realize it; and now I, as a linguist, know these facts about you and the speaker by means of *empathy*. It does not make sense to assume that it was e.g. Mary who did the listening (i.e. that this is a different-subject construction), because it would be impossible to recover this information from the sentence-form. On the other hand, suppose the speaker intends to say that John felt embarrassed while, or because, Mary was listening (i.e. the speaker intends to convey a different-subject meaning). Then, for reasons just explained, he must an explicit-subject converb, like *With Mary listening to the conversation, John felt embarrassed*. Notice, in particular, that here the proposed explanation assumes that the *speaker* can solve his problem of shaping the linguistic form only by considering what is needed for the *listener* to solve his own problem of understanding the linguistic form-as-shaped (i.e. identifying the subject of the converb construction). This reveals the inherently *social* (i.e. minimally *dialogical*) nature of language.

Finally, let us consider the process of *grammaticalization*. The proposed explanations of this process presuppose *imagining* how the speakers may have *(re)interpreted* a certain construction, or how the linguist *himself* would have *(re)interpreted* it, if he had been one of the speakers. Once again, we have to do with *empathy*. Consider the following traditional example, discussed by Paul (1975 [1880]: 299):

Reanalysis:  Ich sehe das: er schläft → Ich sehe, dass er schläft
Extension:  Ich sehe, dass er schläft → Ich bin überzeugt, dass er schläft

Reanalysis and extension are generally assumed to be neither convenient descriptive fictions nor some merely-neurological phenomena, but really existent *cognitive processes*: “The essence of grammaticalization is [the cognitive process of] metaphoric induction of new members into the category. Such changes involve the *redefinition* of the characteristic properties and their relative ranking” (Givón 1984: 19 emphasis in the original). “[The explanation invokes] the least costly *inferential extension* from semantic to pragmatic that speakers are likely to make” (*ibidem*, p. 183; emphasis added).

All this is well understood. What is less well understood is that I, as a linguist (and you as well) can identify such processes only because I can *imagine* performing them *myself*. This is why I can understand how the demonstrative pronoun *das* is *reanalyzed* as the conjunction *dass* and how, in the sequel, the use of the *dass*-clause is *extended* to new contexts, i.e. to contexts where the demonstrative pronoun *das*
could not have occurred. If we cannot imagine ourselves performing a putative process of grammaticalization, we reject it.

3.2. Empathy and Iconicity

Above, empathy was characterized as the method of reaching rational (means-to-end) explanations. An action which fails to serve as a means for achieving the goal is ruled out. This explains why implicit-subject converbs (e.g. Listening to the conversation, John felt embarrassed) cannot express different-subject meanings (like ‘While Mary listened to the conversation, John felt embarrassed’). On the other hand, when two actions serve equally well as means for achieving the goal, it is the more economical one that is chosen. This explains why same-subject meanings are preferably expressed by implicit-subject converbs (= Listening to the conversation, John felt embarrassed), and not by explicit-subject converbs (= “John listening to the conversation, John felt embarrassed”). For the same reason, when the referents of \( N_1 \) and \( N_2 \) have to be re-identified together, \( Pro-1&2 \) is preferred over \( N_1-and-N_2 \).

As shown by these examples, rational explanations are based on considerations of economy, which means that ‘economy’ comes close to being a self-explanatory concept. The same seems to be true of iconicity. It is taken for granted that similarity between A and B eo ipso makes it easier to understand one as an expression of the other. This methodological intuition was already expressed by the medieval grammarians, called ‘Modistae’, who argued that the structure of language (= ‘modi significandi’) is explained by showing that it reproduces the structure of extralinguistic reality (= ‘modi essendi’) by the intermediary of mind (= ‘modi intelligendi’) (cf. Itkonen 1991: 226–237). The same view underlies Wittgenstein’s (1969 [1921]) ‘picture theory’ (Abbildtheorie) of language, which embodies a ‘philosophically purified’ version of iconicity (cf. Itkonen 1970: 125–130).

Whether iconicity is accessible to empathy, seems open to discussion. In its present form, iconicity represents a theoretical insight whose legitimacy has been flatly denied by some linguists. But once the existence of iconicity has been accepted, it is easy to (consciously) imagine how, and why, language users should have (unconsciously) resorted to it. It is simply natural to describe events in the same order that they have happened, to refer to what is ontologically less vs. more by means of what is linguistically less vs. more, to express direct causation directly and indirect causation indirectly, and so on. As a result, we come to appreciate the functioning of unconscious rationality that is capable of bringing all this about.

Here we have been dealing with ‘economy’ and ‘iconicity’. It is good to point out that, in general, these two notions do not come into conflict. The mistaken view that economy and iconicity constitute ‘competing motivations’ arises from confusing iconicity with the 1M1F principle (cf. below).

3.3. Discussion: Croft (2003) on Implicational Universals and ‘Deeper Explanations’

“A standard implicational universal is a generalization over a tetrachoric table in which three types are attested and one type is not (or is extremely rare). The pattern of attested and unattested language types in a tetrachoric table (or larger table) is the central fact; an implicational universal is a hypothesis of an inductive generalization over that pattern” (Croft 2003: 56). Following Greenberg (1966), Croft introduces the notions of dominance and harmony, when discussing IU’s of word order (pp. 60–62). Consider the IU ‘If demonstrative follows the noun, adjective follows the noun’, or ‘if
NDem, then NA’, which is — by contraposition — equivalent to ‘if AN, then DemN’. These IU’s are based on the tetrachoric table with the three attested alternatives DemN & AN, DemN & NA, NDem & NA and the one (supposedly) unattested alternative NDem & AN. The word orders DemN and NA are called ‘dominant’ for the very simple reason that each of them occurs twice among the attested alternatives (while the inverse orders NDem and AN, which occur only once, are called ‘recessive’). For obvious reasons, a relation of ‘harmony’ is claimed to obtain between DemN and AN, on the one hand, and between NDem and NA, on the other.

Because a tetrachoric table merely summarizes the data, and because both dominance and harmony “can be directly read off from a tetrachoric table” (p. 60, 62), they do not qualify as explanatory notions. Therefore it is surprising to see that Croft nevertheless tries to furnish them with some sort of explanatory force. On p. 78, UI’s are contrasted with such “deeper concepts” as dominance and harmony. And on p. 86, again, UI’s are contrasted with “more far-reaching relationships between the word order parameters, such as dominance and harmony”. These are in turn contrasted with a “deeper (possibly external) explanation for the relationship”, often provided by “competing motivation analyses”. And yet, dominance and harmony have been claimed to be capable of not just being read off from tetrachoric tables but also (p. 67) of providing competing motivation analyses.

In any case, Croft realizes that there is a need to go beyond UI’s and to search for deeper explanations. If the account given here in 3.1–2 is correct, such explanations must, unknowingly or knowingly, refer to empathy, economy, and/or iconicity. Let us see whether this is indeed the case.

‘Deeper explanations’ are explicitly offered on four different occasions, namely for the following phenomena: a) word order and affix order universals (pp. 69–80); b) economy and iconicity (pp. 116–117); c) hierarchies and categories (pp. 137–139); d) the typology of grammatical relations (pp. 178–183). These four topics will now be examined in order.

a) With reference to Hawkins’ and Dyer’s (supposedly analogous) processing models, it is suggested (p. 74) that universals of word order can be explained by the attempt to minimize difficulties in parsing. Roughly put, Hawkins’ (2001) principle of ‘Minimize Domains’ assumes that constituents with adjacent elements are recognized more easily than those with discontinuous elements. This idea is clearly a manifestation of (even if it may not be exhausted by) cohesional iconicity. A diachronic explanation is also suggested (p. 77, 79): ‘A and B are similar because B has evolved from A’; but its explanatory force is not overly strong. In any case, “it is not clear what word orders are in fact harmonious with what other orders. Most word order universals have significant exceptions [...]” (p. 77).

b) It is suggested that economy and ‘iconicity’ must be ultimately explained as being “an adaptive response to functional pressures” (p. 117). This claim actually says less than it seems to do since it is supposedly true of each and every aspect of human and animal behavior; and, in any case, what is at issue, is not iconicity but the 1M1F principle. More interestingly, it is suggested that (pace Greenberg) frequency as such is not a sufficient explanation of asymmetries in structural coding: “Of course one must still ask why some situations are talked about more frequently than others. Such situations are presumably more perceptually or culturally normal or salient. [...] However, it could be that salience or expectedness directly motivates structural coding, rather than via frequency and entrenchment” (pp. 115–116). Now, at least some amount of empathy is clearly required in order to understand what the members of
some alien culture might find perceptually or culturally normal or salient or expected.

c) “A conceptual space can itself be given a deeper explanation. [...] in those cases where conceptual spaces have been proposed, the similarity in form — the linking of conceptual categories in conceptual space — reflects in turn similarity of function” (p. 137). This is a perfect example of iconicity. (Notice that we are now dealing with universally valid concepts, not with language-specific meanings). Moreover, Croft is discussing here the animacy hierarchy ‘1/2SG > 3SG > human > animate > inanimate’, which is also, and preferably, known as ‘empathy hierarchy’. Surely, the link between empathy hierarchy and empathy is so close as to be self-evident.

d) Pronominal indexation is explained by the “high salience or topicality” of the entities spoken about (pp. 178–179). The formal distinction between subject vs. object prototypes is variously explained (pp. 179–181) by the high vs. low animacy of the referents, by the speaker’s wish not to mislead the hearer, and by discourse considerations (= explicitly mentioning only the new information). Finally, the correlation between high transitivity and foregrounding is explained, once again, by “cognitive salience” (pp. 181–183).

To summarize: There is no way that a typologist can do his or her job without ‘rethinking people’s thoughts’. This process was called ‘re-enactment’ by Collingwood (1946). In the German hermeneutic philosophy it was known as Einfühlung (for discussion, cf. Itkonen 1993, 2003: Ch. 11 and Appendix 5).

4. Conclusion

During his career, Wittgenstein proposed first a ‘picture theory’ and then a ‘use theory’ of language. It is often thought that the latter was a repudiation of the former. However, it has been pointed out, e.g. by Kenny (1975: Ch. 12), that there is no contradiction involved here. Language (i.e. language-form) is a picture-like instrument that we use in accordance with socially or intersubjectively valid rules (or norms), i.e. rules which have in part been motivated by the picture-like nature of the very instrument whose use they govern. It is interesting to note that these two aspects — instrument and picture — are present already in the oldest extant treatise on language in the Western tradition, namely Plato’s Cratylus. First, language is defined as an organon, i.e. instrument. Second, it is argued that, if the phusei thesis is correct, language is also a picture of the reality:

“Still, Socrates is not prepared to give up the phusei view entirely. His final opinion seems to be that perhaps there is something like the correctness of names; but even if there is not, at least there should be. In this ambiguous statement one can see the origin of the ‘ideal language’ doctrine which was later to lead to the creation of logical semantics: ‘I quite agree with you that words should as far as possible resemble things [...] For I believe that if we could always, or almost always, use likenesses, which are perfectly appropriate, this would be the most perfect state of language, as the opposite is the most imperfect’ (435C)” (Itkonen 1991: 171).

REFERENCES