

DANIEL HOLE (Berlin)

Heuristics and typology

Summary

In this note some issues in methodology and heuristics are addressed from the perspective of an ongoing typological research project in the domain of reflexivity, emphatic reflexives and focus particles.¹ I will have nothing to say about issues of sampling, statistics and branches of typology which aim at quantifying their results. Instead I want to (i) defend our way of doing factorial typology against the position of an holistic typology; (ii) highlight the differences between heuristic and justificational procedures; (iii) illustrate a fruitful heuristic strategy oscillating between onomasiology and semasiology; (iv) determine in what respects heuristic pitfalls in typology and generative grammar differ or coincide. By strictly distinguishing methodology from heuristics I am trying to show that the strengths of cross-linguistic research lie in the realm of heuristics. A short characterization of our research project precedes the theoretical discussion.

1. A typological research project on intensifiers, reflexivity and focus particles

The sentences in (1) illustrate different uses of *herself*, (2) gives two examples of the German third person reflexive pronoun *sich*.

- (1)
 - a. *Sue_i was proud of [her*(self)]_i.*
 - b. *Sue_i noticed a letter in the drawer. It was addressed to [her(self)]_i.*
 - c. *Sue's brother remained calm, but Sue herself got mad at the policeman.*
 - d. *Sue didn't call the plumber, she fixed the sewer herself.*
 - e. *Sue couldn't give her brother any money; she was broke herself.*
- (2)
 - a. *Susanne war stolz auf sich (selbst).*
'Susanne was proud of herself.'
 - b. *Susanne schämte [sich (*selbst)].*
'Susanne felt ashamed.'

(1a) and (1b) are both examples of *herself* occupying an argument position; in the other examples this is not the case. But whereas (1a) is an ordinary case of reflexivization in which *-self* may not be left out if a co-referential interpretation of *Sue* and the prepositional object

¹ The complete title of this project is *Typologische Untersuchungen zu den Intensifikatoren, zur Reflexivität und zu den Fokuspartikeln*. The financial support from the Deutsche Forschungsgemeinschaft is gratefully acknowledged. The positions defended here do not necessarily reflect the position of each member of our research group which, at present, comprises the following persons: EKKEHARD KÖNIG, VOLKER GAST, PETER SIEMUND, RENATE RAFFELSIEFEN, and myself. Special thanks go to KAY-EDUARDO GONZÁLEZ-VILBAZO who has heavily influenced my way of thinking about methodology.

is to be expressed, *-self* in (1b) is optional if co-reference with *Sue* in the preceding clause is intended. The latter use of reflexives is often called "logophoric", i.e. a certain proposition is presented from the viewpoint of the referent to which the logophoric pronoun refers. In (1c–e) we have cases with what we call adnominal and adverbial intensifiers: *herself* in (1c) relates Sue to her brother in a way which makes Sue appear to be central with regard to her brother; the use of adverbial *herself* in (1d) is good if other people performing the relevant job are under discussion; *herself* in (1e) is syntactically similar, but it gives rise to a different reading which is often captured as being similar to the use of an additive focus particle like *also/either*. The German examples in (2) show that the use of the German cognate of *-self*, *selbst*, is optional in reflexive contexts (2a) and even ungrammatical in cases of lexical reflexivity (2b). Moreover, German has a reflexive pronoun which does not contain an ordinary pronoun as a part. The intensifier uses as illustrated in the English cases in (1c–e) have comparable German equivalents with *selbst*, as the reader may verify for himself. In the ongoing research project, similar data for roughly 75 languages from almost all parts of the world have been collected so far. On a descriptive level we are trying to state the limits within which the expression of reflexivity and intensification may vary (categorical status of reflexive markers and intensifiers, inflectional properties of intensifiers, distribution of intensifiers in specific marginal contexts), find out about historical paths of development leading to and originating from intensifiers and reflexivity markers, and provide semantic and/or pragmatic analyses of the observed facts. We also review theoretical discussion in the relevant areas of grammar and test hypotheses brought up in the context of these discussions against the background of our data. For a survey of this research cf. KÖNIG (to appear).

2. Issues of research practice and methodology

2.1. Factorial typology vs. holistic typology

By pursuing a research project as described in the preceding section, we are committing ourselves to a factorial typology as opposed to a typology of holistic language types (cf. the tradition dating back to WILHELM VON HUMBOLDT and other romanticist thinkers which was taken up by GEORG VON DER GABELENTZ 1891[1971], among others, and which has been influential in European typology with researchers such as EUGENIO COSERIU and PAOLO RAMAT; see COSERIU 1988 and RAMAT 1987, but also COMRIE 1981: 37 for a critical assessment of holistic language typology). That means we compare certain traits of languages cross-linguistically without assuming aprioristically that these traits must depend on certain other linguistic characteristics. We think that factorial typology should be given precedence over holistic conceptions of language types, both for theoretical and for practical reasons. One obvious problem with a holistic approach to language types is the fact that it is hard to decide on which phenomena a holistic typology should be based. One cannot compare languages as a whole, and it is not clear whether determining, for instance, the morphological type of a language allows more essential conclusions about a holistic language type than, let's say, the statement of its system of morphophonemics. Thus, for very practical reasons, a holistic typologist will always end up doing factorial typology, at least for a start. A general problem with holistic language typology follows from the indisputable fact of cross-cutting dimensions of comparison. A holistic typologist who is both interested in morphological types and

the expression of anaphoric relations will probably come up with an isolating and an inflecting language type (among other types), and he will find that in some languages pronouns may be dropped while in others they may not. These phenomena yield orthogonal classifications. A holistic classification of languages should, however, yield a hierarchically organized system of language types, i.e. there may be sub-types of more general types, but the relationship between languages and most general types must be many-to-one, not many-to-many (for a sketchy outline of this position see COSERIU 1988: ch. XII). Factorial typology will, on the other hand, usually deal with many-to-many relations between languages and types. If holistic typology admits cross-classification (i.e. many-to-many relations between languages and types), it is really factorial typology in disguise.

2.2. *Heuristic procedures vs. justificational procedures*

The question whether one should make use of inductive reasoning or of deductive reasoning in linguistic research dominates many methodological discussions. A typical situation looks like this: a typologist defends the view that good hypotheses on language universals can only be stated on a broad basis of cross-linguistic data; a generative grammarian strongly argues in favour of an approach in which some theory is the guiding line through the data; the data are there to (dis)confirm a universal hypothesis. Such discussions must be frustrating because they are not to the point. What the typologist is talking about when he requires many languages to be in the data base is the issue of forming a reasonable hypothesis which does not run the risk of being disconfirmed because obvious facts have been ignored. What the generativist has in mind when he stresses the secondary status of language data is not the forming of hypotheses, but rather their testing. In other words: the typologist thinks of elaborate ways to come up with an hypothesis, while the generativist usually has the justification of his hypothesis (in a methodological sense) in mind. Viewed from this perspective the typologist and the generativist do not necessarily disagree any longer. If, in the course of further research, the typologist finds a language which runs against his hypothesis – and he should try hard to find such a language – he will dismiss or modify his hypothesis (unless the new data can be explained away). That is exactly what a generativist does when he finds counter-evidence. Thus both linguists must justify their hypotheses by trying to prove that they are wrong. (Following POPPER 1971, I assume that in empirical research there is simply no way to test whether a hypothesis which states a general law is actually true. Only its falsehood can be proven.) The conclusion here is that the generativist and the typologist often do not disagree on methodological issues, with methodology taken to cover only the matters of justifying hypotheses or theories, but rather on heuristics, i.e. the issue of what is a good way to find an hypothesis. Typologists will insist on the point that one should state an hypothesis after a lot of data have been reviewed, whereas generativists are either not interested in how one comes to state a hypothesis if only it can be expressed in terms of general principles, or they follow POPPER's guiding line by deliberately making the strongest generalization which is still supported by the data so far considered, because this makes it easier to find counterexamples and modify the theory accordingly. If we take for granted that the justification of hypotheses always involves deductive reasoning no matter whether we are carrying out typological research or studies in the generative tradition, we should take a closer look at the heuristics applied in typological research. It is exactly in this area, and not in the area of (justificational) methodology proper, where the strengths of typological research are

~~apparent~~^{obvious}. To see this let us investigate the relationship between onomasiology and semasiology in cross-linguistic comparison.

2.3. Onomasiology and semasiology: the case of intensifiers and reflexivity

Let us start with a (highly compressed and simplified) illustration based on the practice in our research project on intensifiers and reflexivity to see how flipping back and forth between onomasiology (analysis leading from the signifié to the signifiant, or: 'how are given notions expressed?') and semasiology (analysis leading from the signifiant to the signifié, or: 'what does a given form express?') is a fruitful heuristic strategy. A cross-linguistic (onomasiological) comparison of words which are comparable in function with German *selbst* as in *der Kanzler selbst* 'the chancellor himself' yields a first (semasiological) distinction between languages: in many languages, and probably in the majority of languages, translational equivalents of German *selbst* are also used to express co-referential arguments in reflexive sentences. In German and in most other continental European languages, this is not the case, but it is true of English; cf. the German sentences and their English translations in (3) and (4).

- (3) *Der Kanzler selbst wird das Schiff taufen.*
'The chancellor himself will christen the ship.'
- (4) *Der Kanzler betrachtete sich im Spiegel.*
'The chancellor was looking at himself in the mirror.'

In the German sentences the reflexive pronoun and the intensifier *selbst* are distinct in form, but in English only a distributional difference exists between intensifiers and reflexive pronouns (reflexive pronouns are arguments, intensifiers are non-argument expressions). But now the scope of the investigation should broaden: it should not only encompass translational equivalents of German *selbst* any longer, but also reflexive pronouns. A semasiological distinction between two classes of reflexive pronouns follows, namely one class which is (partially) made up of an intensifier like English *x-self*, and a second class which does not contain any such element (German *sich*, for instance). Further investigations of verbal contexts in which the German and the English reflexive pronouns are used yields the well-known result that contexts in which reflexive pronouns are used in English are a subset of the contexts in which reflexive pronouns are used in German.² While it is difficult to generalize semantically over the complete distribution of German *sich*, which also comprises uses of derived intransitivity, the English situation is quite clear: reflexive pronouns are only used if typically other-directed actions are carried out reflexively (cf. KIPARSKY 1990 or KÖNIG/SIEMUND 1999). Killing being an action usually done to others, *to kill oneself* is a perfect context for the English reflexive pronoun. On the other hand, with regard to the typically non-other directed action of shaving, *to shave oneself* as opposed to *to shave* is a marked option which may be used in contexts in which the shaving could also be performed by somebody else, e.g. if one is sick and has to stay in hospital. Usually the unmarked option *to shave* is preferred, just as with other verbs of grooming.

² If this generalization is to be maintained, the occurrence of *herself* as in (1b) must be categorized as non-reflexive. Without being able to go into the details of this argument here it is clear that we are not dealing with a syntacticized occurrence of the *self*-form in (1b) because its use is not mandatory.

All of this is pretty much descriptive, the only theoretical prerequisites being that there is some cross-linguistic consistency between functions that are encoded by linguistic items in different languages and that languages can be compared at all. The results of this small-scale comparison are, however, quite challenging. If we want to have a theory of reflexivity in German and English, or in other languages which display a similar split, and if we want to develop a typology of expressive means in this area, which values do we want to keep constant, and which ones may vary? Is the semantics of reflexive pronouns the same in English and German? If this were so, what could account for the different distributions? Is typical other-directedness and typical non-other-directedness a lexical or a pragmatic notion, or are we dealing with a syntactic phenomenon? If the semantics of reflexive pronouns in English and German were not the same, how exactly would they differ, i.e. how could we incorporate the notion of (non-)other-directedness in the representation of pronouns? In English, do we want to derive the meaning of reflexive pronouns from the meaning of intensifiers, or should the meaning of intensifiers boil down to reflexivity in the end? Or is the formal identity of intensifiers and reflexive pronouns in English and other languages just a case of homonymy? In this paper, no attempt is made to answer any of these questions. What matters here is the fact that a certain heuristic strategy, namely flipping back and forth between onomasiological and semasiological procedures in a comparative perspective, leads to the statement of problems and rough hypotheses which may then be the subject of linguistic research in a narrow sense.³ Just by reviewing means to express notions like German *selbst* in other languages a large area of investigation has opened up. Without restricting our viewpoint by adopting a very specific hypothesis on intensifiers or reflexivity right from the start, we are now in a position to develop hypotheses on particular issues in this domain, but we can constantly keep the broader domain in mind. I am emphasizing this point because this is exactly what many generative grammarians don't do: they start out from a particular problem derived from a very general hypothesis or theory which should hold universally. The empirical data to support their claims have usually been stripped off their systematic context, both cross-linguistically and within a single language. In our case, a typical question in generative grammar would be: if (reflexive) binding is a basic syntactic phenomenon, how can we account for the observed differences in different languages. Recall that this is also one of the possible hypotheses a typologist will probably ponder over, but for the cross-linguistically informed typologist it is only one of several possible research tasks. We should note here that nothing in the methodological design of generative grammar keeps generativists from widening the perspective, and that a cross-linguistic approach is not a guarantee to avoid fallacies in the course of investigation. The typologist, in dealing with many different languages, is just more likely to be aware of systematically cooccurring phenomena in a variety of languages, and to separate them from incidental cooccurrences, while such incidental cooccurrences may be hard to detect if one is only guided by a powerful theory which makes predictions about a single phenomenon.

We may thus summarize as follows: nothing in typology as such guarantees better hypotheses; nothing in generative grammar as such prevents a researcher from doing cross-linguistic

³ HIMMELMANN's *holistic methodology* (this issue) is the very same thing as my flipping back and forth between onomasiology and semasiology in a comparative perspective. According to my terminological choices (see section 2.2.) I refrain from subsuming this research strategy under the heading of methodology proper. Methodology in the sense in which it is used here only determines the validity of justificational procedures. HIMMELMANN's *holistic methodology*, and also my flipping back and forth between onomasiology and semasiology, could thus be renamed *holistic heuristics*.

tic research and from considering concomitant factors; the large amount of data from various languages, hopefully considered in their systematic context, makes it easier for the typologist to avoid the formation of obviously untenable universal hypotheses and to uncover regularly cooccurring phenomena.

2.4. *Being aware of the larger context: reflexives and the nature of some emphatic elements*

In this section I would again like to exemplify how the application of the heuristics presented above may lead to generally interesting research tasks that are not even noticed if heuristic matters do not count as important. Having illustrated the heuristics favoured here in the preceding section, I would also like to pay attention to the factors that discourage a similar heuristic process in many generative approaches. The topic is again taken from our research project.

If one studies expressive means to express co-referential arguments of a single core predicate in different languages, and if one simultaneously considers the wider distribution of the elements that figure in these constructions (i.e. if one is aiming at the complete semasiological picture or if one applies, in HIMMELMANN's terms, a holistic methodology/heuristics), it is difficult not to come up with the following generalization:

- (5) If a nominal reflexivization strategy in a given language involves an emphatic element this element also expresses adnominal and/or adverbial intensification.⁴

If this is so, we will immediately ask ourselves what implications this cross-linguistic fact has. Note that the regularity with which (5) gets confirmed language after language hints at a synchronic link between intensification and the expression of co-reference in a local domain. If we were dealing with two phenomena that are merely linked by historical development, we would expect to find languages in which either the reflexivization strategy or the expression of intensification has changed. For the synchronic account of nominal reflexivization strategies involving emphatic elements and of adnominal and/or adverbial intensification it follows that we should aim at showing what the shared properties of the two phenomena are. The research in our project has not (yet) given the ultimate response to this challenge (cf., however, steps in the direction of making this link explicit e.g. in KÖNIG/SIEMUND 1996, 1998). What matters here is the fact that given the complete semasiological picture there is no way to evade the issue. In generative grammar it is all too easy not to address it, and the fact that among the hundreds and probably thousands of generative studies on reflexivity only a handful, at best, deal with it (e.g. BROWNING 1993), simply shows what consequences the lack of interest in heuristic strategies and semasiological completeness has in this framework. Since there is *per se* no ban on careful semasiological generalizations in generative grammar let us try to identify the factors that probably contribute to a situation which looks as if such a ban actually existed. Here I will discuss three factors that may help to account for the observed situation. The first one is a consequence of the history of linguistic schools. Generative grammar started out as a polemic counter-movement to American taxonomic structuralism. In the course of doing away with vacuous distributionalist concerns the taxonomic goose was killed with the golden egg of holistic heuristics; i.e. structur-

⁴ Recall that, in the context of our research project, *intensifier* and *intensification* are technical terms with a denotation that is more restricted than the meaning these terms may have in other frameworks or in every-day language.

alism was discarded together with the genuine structuralist idea that linguistic analysis should always take the systematic whole (of a phenomenon, of a language) into account. Now it was possible for generative grammarians to carry out extremely reductionist investigations into very small sub-areas of grammar without having to worry about whether these areas really constituted systematic units of individual grammars. A second reason for the generative neglect of the systematic tie between intensifiers and nominal reflexive pronouns with emphatic elements may be seen in the marginal status which the lexicon was assigned in many branches of generative grammar: researchers did not study the formal parallelism between reflexive pronouns and intensifiers simply because reflexivity was deemed a grammatical phenomenon, whereas intensifiers are just words which may usually be left out without influencing grammaticality. From this perspective it is quite obvious how the virtually complete emphasis on reflexivity as opposed to the neglect of intensification could come about in generative grammar. The third reason that comes to my mind is really complementary to the more general second one: starting with CHOMSKY (1981) and throughout the decade of government and binding theory, accounts of reflexivity in terms of syntactic binding were part of the core concerns of generative grammar, whereas the ways to deal with word meaning in semantic branches of generative grammar were not yet very developed, let alone standardized (cf. the overview of competing generative lines of research in this field in CHERCHIA/McCONNELL-GINET 1990: ch. 8). The syntactic treatment of reflexivity facts usually centered around allegedly innate axiomatic binding principles, and no attention was paid to the lexical sources from which emphatic components of reflexive pronouns derived. This is also true of the tradition within generative grammar which distinguishes morphologically complex from simplex reflexive pronouns, usually called SELF-anaphors vs. SE-anaphors, and even of REINHART/REULAND (1993) who approach the whole issue of reflexivity and binding from a semantic perspective.⁵

The non-generativist typologist may shrug now and ask: "What am I to worry about the failures of generative grammar?" In my opinion, this refusal is precipitate. For one thing, descriptive traditions in generative grammar are often sufficiently widespread and influential to shape descriptive generalizations in other traditions as well; so we might sometimes be doing typology on a generative and insufficient descriptive basis without even being aware of it. Secondly, as was pointed out in section 2.3, a typological perspective alone is not a guarantee of having better hypotheses in the end. Studying the shortcomings of other schools will help typologists to avoid them in their own work. We have identified the loss of a systemic view of language, the lexical disinterestedness and the axiomatic view of reflexive binding as the main factors which have led to the descriptive situation concerning reflexivity and intensification as we find it in generative grammar. I think typologists cannot be said to have no interest in the lexicon, but a similar danger is constituted by the fact that for large language surveys information from grammars must be used without there being a way to investigate the larger distribution of the elements one is interested in, simply because no other grammars or dictionaries exist, the language is extinct, or fieldwork is impossible. The danger of losing a systemic view of language is lurking just as much in typology as in generative grammar, I think. The third factor, namely taking for granted certain very specific axi-

⁵ Although the terminology of SE-anaphors vs. SELF-anaphors may suggest just this, the distinction is in fact orthogonal to our differentiation between nominal reflexivization strategies with emphatic elements and other nominal reflexivization strategies. What gets to be classified as a SE-anaphor in generative grammar may actually be an intensifier in our terminology (cf. Chinese *ziji* which is a SE-anaphor, but also has uses as an adnominal and as an adverbial intensifier).

oms – which are actually only quasi-axiomatic working hypotheses – right from the start, is equally relevant in typology. The only difference seems to be that the axioms in many typologies are not of a formal or syntactic nature, but rather of a functional kind: since every language must somehow express communicative function *f*, comparing expressive means in the domain of *f* is justified without considering the extension of the expressive means into other functional domains. I believe the core notion with respect to these two factors, and also the complement of holistic heuristics, is modularity. While I would like to remain neutral here about the role modularity should play in the design of a linguistic theory, I definitely claim that it is harmful if observed in the process of searching for hypotheses, i.e. if we make modularity also the guiding line in heuristics.⁶ Typology is weak, or at least not any stronger than competing research traditions, when it over-determines its specific research topics. Typology is strong when it first opens doors between spaces of grammar that were not even suspected to be adjacent.

References

- BROWNING, M. A. (1993): Adverbial reflexives, in: *NELS* 23,1, 83–94.
- CHIERCHIA, GENNARO and MCCONNELL-GINET, SALLY (1990): *Meaning and grammar: an introduction to semantics*. Cambridge, MA & London: MIT Press.
- CHOMSKY, NOAM (1981): *Lectures on government and binding*. Dordrecht: Foris.
- COMRIE, BERNARD (1981): *Language universals and linguistic typology*. Oxford: Blackwell.
- COSERIU, EUGENIO (1988): *Einführung in die Allgemeine Sprachwissenschaft*. Tübingen: Francke.
- CROFT, WILLIAM (1995): Autonomy and functionalist linguistics, in: *Language* 71,3, 490–532.
- GABELENTZ, GEORG VON DER (1891[1971]): *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel Nachf. [Reprint: Tübingen: Niemeyer 1971].
- KIPARSKY, PAUL (1990): Strong reflexives in Germanic, *Diachronic Syntax Conference*, York.
- KÖNIG, EKKEHARD (to appear): Intensifiers and reflexive pronouns, in: MARTIN HASPELMATH, EKKEHARD KÖNIG, WULF OESTERREICHER and WOLFGANG RAIBLE (eds.), *Language typology and language universals/Sprachtypologie und Sprachliche Universalien/La typologie des langues et les universaux linguistiques: An international handbook/Ein internationales Handbuch/Manuel international*, Berlin etc.: Mouton de Gruyter.
- KÖNIG, EKKEHARD and SIEMUND, PETER (1996): On the development of reflexive pronouns in English: a case study in grammaticalization, in: BÖKER, UWE and SAUER, HANS (eds.), *Anglistentag 1996: Proceedings*, 95–108.
- KÖNIG, EKKEHARD and SIEMUND, PETER (1998): On the development of reflexive anaphors in English, in: *ZAS Working Papers*.
- KÖNIG, EKKEHARD and SIEMUND, PETER (1999): Intensifiers and reflexives: a typological perspective, in: ZYGMUNT FRAJZYNGIER (ed.), *Reflexives: forms and functions*, Amsterdam: Benjamins.
- POPPER, KARL R. (1971): *Logik der Forschung. 4., verbesserte Auflage*. Tübingen: Mohr.
- RAMAT, PAOLO (1987): *Linguistic Typology*. Berlin etc.: Mouton de Gruyter.
- REINHART, TANYA and REULAND, ERIC (1993): Reflexivity, in: *Linguistic Inquiry* 12, 657–720.

DANIEL HOLE
Institut für Englische Philologie
Freie Universität Berlin
Goßlerstraße 2–4
14195 Berlin
hole@zedat.fu-berlin.de

⁶ For a thorough treatment of modularity or autonomy hypotheses of different strengths from a typological-functionalist perspective cf. CROFT (1995).