22. The relation of non-Western approaches to linguistic typology

1. Background

In the non-Western world there have been long-lasting linguistic traditions of high scientific value in India (= the Sanskrit tradition and, to a somewhat lesser extent, the Tamil tradition) and in the Islamic cultural sphere (= the Classical Arabic tradition). Also the relatively short tradition of indigeneous Japanese linguistics deserves to be mentioned in this context. Of course, some forms of linguistic thinking have existed in other cultures too.

The traditions which are under scrutiny here are characterized by the fact that the oldest extant grammar has had a decisive influence on subsequent linguistic research. Each of these grammars concentrates on one language only, namely Sanskrit, Tamil, or Classical Arabic. In fact, it seems to be a universal truth, also confirmed by the history of Western linguistics, that linguistic theorizing starts, and must start, by concentrating on one single language (which in the West was either Classical Greek or Latin). This explains the lack of the typological aspect in the Sanskrit, Tamil, and Classical Arabic traditions (as well as in the Japanese tradition).

It follows that if one wishes to investigate the non-Western approaches to linguistic typology, one has to make use of an 'indirect' method. An attempt has to be made to evaluate to what extent the various descriptive frameworks have been determined by the typological peculiarities of those languages for the description of which they have been devised and, inversely, to what extent they would lend themselves to the description of typologically dissimilar languages.

2. Sanskrit

Pāṇini's (c. 400 B.C.) grammar *Aṣṭādhyāyī* ('Eight Chapters') constitutes the essence of the Sanskrit tradition. According to the expert opinion, it still qualifies as the best grammar of a single language ever written. In the present context it is not possible to go into the technical details that justify this opinion. A very thorough introduction to Pāṇini is presented in the dozen volumes in which S. D. Joshi and J. A. F. Roodbergen have edited and explicated central passages from Patañjali's (c. 150 B.C.) 'Great Commentary' (see e.g. Joshi & Roodbergen 1975). Book-length introductions are provided by Sharma (1987) and Cardona (1988). Shorter overall accounts are given e.g. by Pinault (1989), Itkonen (1991: 5–87), and Kiparsky (1993).

It is generally agreed that Pāṇini's grammar constitutes a 'device' which takes semantic representations as its input and aims at generating – via a great number of intermediate stages – the 'surface forms' of all and only correct sentences of Classical Sanskrit as its output. The semantic representation centers around the notion of 'action' (*kriyā*), which is divided into three (semantic) subcategories anticipating the (formal) distinctions between the active, personal passive, and impersonal passive endings of the verb. An action involves participants that exemplify one or more of the following semantic roles (*kāraka*): agent, patient, recipient, instrument, location, source. These roles represent a 'shallow' semantic level in the sense that they correspond, roughly, to the case endings of Sanskrit (excluding the genitive). However, the roles of agent and patient require a special treatment: the agent may be expressed by the active endings of the verb (plus the nominative) or by the instrumental (in connection with the passive endings) or by the genitive (in connection with deverbal nouns), whereas the patient may be expressed by the passive endings of the verb (plus the nominative) or by the accusative (in connection with the active endings) or by the dative (in the sense of 'goal') or by the genitive (in connection with deverbal nouns). In addition, all semantic roles may be encoded by derivational affixes as part of the meaning of the word to be derived.

Because the six roles represent a 'shallow' level of semantics, there is no need for an 'intermediate' level corresponding to the grammatical functions 'subject' and 'object'. Sanskrit is a free word-order language, which

means that the word order, being dependent on the speaker's communicative intention, is no concern for the grammarian.

The semantic representation containing e. g. the information 'Devadatta is cooking rice' is the starting point for several possible derivations, because – depending on which grammatical rule is chosen first – the derivation may ultimately result in an active sentence or a passive sentence or a nominal sentence (with *paktā/pācakaḥ* = 'cooker', and no copula) (see 1a–c).

(1) (a) Sanskrit
Devadatta-ḥ odana-m- pac-ati
D.-NOM rice-ACC cook-ACT
'D. is cooking rice'

(b) *Devadatte-na odana-ḥ pac-yate*
D.-INSTR rice-NOM cook-PASS
'Rice is being cooked by D.'

(c) *Devadatta-ḥ odana-sya*
D.-NOM rice-GEN
paktā/pācaka-ḥ
cooker(NOM)/cooker-NOM
'D. is a cooker of rice'

The phenomena which could be handled by the notion of syntactic governance are handled here indirectly, namely by means of semantic roles: these are determined by the meaning of the verb plus the choice of one of the three semantico-syntactic categories of the verb (cf. 1a–c). The morphosyntax of verbal complements is in turn determined by rules assigning forms to semantic roles.

The formalism developed by Pāṇini follows the structure of Sanskrit rather closely. Kiparsky's (1993) insistance that Pāṇini's grammar be considered 'generative' in the technical sense of this word seems to indicate, however, that it could have been developed to deal with typologically dissimilar languages to the same extent as current generative grammars can.

3. Tamil

The grammar *Tolkāppiyam* ('Old Book'), which was composed around 100 B. C., describes the earliest stage of Old Tamil (cf. Lehmann 1994); at the same time, it is the oldest extant document of this languae. It consists of three parts which deal with phonology, morphology-syntax-semantics, and poetics, respectively. It is probable that the third part has been composed three or four hundred years later that the first two. The preface of the book expresses indebtedness to a Sanskrit research tradition, but this tradition is definitely non-Pāṇinian in character.

Commentaries on *Tolkāppiyam* that have been written between 1000–1700 are available. According to Chevillard (1996: 23–24), the commentary tradition has in several respects improved upon the theory expounded in *Tolkāppiyam*.

The second part of the grammar has been analyzed by Sastri (1945) and Chevillard (1996). It is centrally concerned with the agreement between nouns and verbs. Nouns are divided into a 'high class' and a 'low class'; the high class has three subcategories: 'masculin' and 'feminin' in singular and 'human' in plural; the low (= 'non-human') class has two subcategories: singular and plural; thus, there are five subcategories in all. The basis for these subcategories is purely semantic. They are systematically expressed by the verbal inflection, and much less systematically by the nominal inflection. Thus, the verb agrees with the semantic subcategory of the subject noun, except that – since the singular form of the noun may stand for the plural – the verb may in such cases be said to express the plural, rather than to agree with a (pre-existent) noun in plural.

The basic difference between nouns and verbs is that they inflect in case and in tense, respectively. Apart from the vocative, there are seven cases (although the 'case' for location is actually expressed by 19 post-positions). Case endings may be either interchanged or deleted without any semantic motivation (for confirmation of this rather surprising fact, see Lehmann 1994: 29, 39, 42). There are seven semantic roles, roughly corresponding to the Pāṇinian *kārakas*, on the one hand, and to the (basic) Tamil case endings, on the other. Verbs are either 'non-independent' (= non-finite) or 'independent' (= finite). The former are either ad-verbal or ad-nominal. The latter are either 'full' verbs (i. e. express the tense explicitly) or 'suggestive' verbs (i. e. contain the tense only implicitly). The finite verb is the only 'complete word' because no other word can constitute a sentence all alone. To be sure, it too 'desires' other words, but these need not be overtly expressed. For instance, the verb form *uṇṭāṉ* may either stand alone or be complemented, depending on the context (see 2a–d).

(2) (a) Tamil
uṇ-ṭ-āṉ vs.
eat-PRET-3.P.M
'He ate' (Lit. 'Ate-he')
uṇ-ṭ-āḷ
eat-PRET-3.P.F
'She ate' (Lit. 'Ate-she')

(b) *cāttaṉ uṇ-ṭ-āṉ*
S.(M) eat-PRET-3.P.M
'Sāttan ate'

(c) *cōṟṟ-ai uṇ-ṭ-āṉ*
rice-ACC eat-PRET-3.P.M
'He ate rice'

(d) *cāttaṉ cōṟṟ-ai uṇ-ṭ-āṉ*
S.(M) rice-ACC eat- PRET-3.P.M
'Sāttan ate rice'

In sum, the second part of *Tolkāppiyam* represents a non-abstract analysis of the Old Tamil sentence structure based, essentially, on the noun vs. verb distinction. Old Tamil is a rather strictly verb-final language, subordinate clauses with non-finite verbs being placed before the one finite verb that concludes the sentence. Interestingly, the commentators give several examples that deviate from the verb-final order.

4. Classical Arabic

The Arab linguistic tradition starts with Sībawaihi's (d. 793) grammar *Al-kitāb* ('The Book'), which analyzes Classical Arabic, i. e. the language stage characteristic of the Koran and of ancient Bedouin poetry. The representatives of this tradition have taken its homogeneity and continuity for granted. Recently it has been claimed, however, that Sībawaihi originally provided his theory with 'actionist' or 'operational' underpinnings which were not taken up by his successors. Yet even those who advocate this view admit that Sībawaihi's phonology and morphosyntax "are basically identical to that of the later grammarians" (Bohas & Guillaume & Kouloughli 1990: 48).

The endings of the three cases of Classical Arabic, i. e. nominative, accusative, and genitive, are provided by the three (short) vowels of the language, namely *-u, -a,* and *-i,* respectively. Moreover, the distinction between the indicative (of the 'imperfect' tense, equalling the present) and the subjunctive is expressed by the distinction between *-u* and *-a*. It is the central task of the Arab tradition, or at least of its morphosyntactic part, to account for this variation (see 3 a–c).

(3) (a) Classical Arabic
yaktub-u zayd-un risālat-an
write-IND.IMPF Z.-NOM letter-ACC
'Z. is writing a letter' (Lit. 'Zayd he-is-writing a letter')

(b) ... *ḥattā yaktub-a*
so-that write-SUBJU
'... so that he should write'
(Lit. '... so that he-should-write')

(c) *marra zayd-un bi rajul-in*
went Z.-NOM past man-GEN
'Zayd went past a man'

Inspection of examples like these made it natural to think that the verb 'causes' – in a metaphorical sense – the *-u* vs. *-a* distinction in the following nouns, and that the 'particle' (here the preposition *bi*) 'causes' the occurrence of the *-i* ending. In just the same way the 'particle' (i. e. conjunction) *ḥattā* 'causes' the occurrence of the subjunctive marker *-a*. Such observations gave rise to a notion of 'governance' and to a corresponding three-way classification of the word classes: a) nouns are governed and do not govern; b) verbs govern and are governed; c) particles may or may not govern, but are not governed. (A member of a given word class cannot govern another member of the same class.) Expectedly, there are borderline cases which must be related – by means of (theoretical) analogy (*qiyās*) – to the 'clear cases'. For instance, participles are verb-like nouns whereas (indeclinable) pronouns are particle-like nouns; the perfect of the copula is a particle-like verb, whereas the emphatic marker *'inna* is a verb-like particle (cf. Owens 1988, Bohas & Guillaume & Kouloughli 1990). Although the Arabs claimed to have no interest in other languages, the three-way classification of word classes was occasionally asserted to be universal (cf. Versteegh 1995: 26, 40–41).

The notion of governance is only the more natural because in a VSO language like Classical Arabic the 'causes' (= verbs and prepositions) precede their 'effects', as they should. From the VSO order it also follows that transitive verbs are taken to be connected more closely to the nominative (= 'subject') than to the accusative (= 'object'). Thus, the notion of 'verb phrase' would be impossible in this framework. The VSO order is regarded as canonical although 'pragmatically' moti-

vated deviations from it are acknowledged as fully permissible.

This rather simple picture is made more complex by the other basic sentence type, namely the nominal sentence (with no copula in the 'imperfect') (cf. 4 a).

(4) (a) Classical Arabic
zayd-un rajul-un
Z.-NOM man-NOM
'Zayd is a man'

First, it is not easy to apply the notion of governance to a sentence like this. Second, a sentence like this exemplifies a 'topic vs. comment' structure which cannot be easily applied to 'verbal sentences' like those given above. However, the force of theoretical analogy ultimately led the Arab grammarians to make both of these moves, i. e. postulating 'invisible governors' for nominal sentences and imposing the 'topic vs. comment' structure on verbal sentences.

There is some 'external evidence' to support the view that it is indeed the formal variation between case endings which constitutes the core of the Arab tradition. Because biblical Hebrew had lost the corresponding endings, the descriptive framework of Arab linguistics could be applied only unsystematically in the first complete grammar of biblical Hebrew, which was written in the 11th century (cf. Kouloughli 1989: 288–289).

5. Japanese

It is very interesting to note that in spite of a cultural tradition of high quality that started already in the second millenium B. C., there has been, apart from phonetics and lexicology, no genuine linguistics in China. The first indigeneous grammar of Chinese was written only at the end of the 19th century. Apparently those interested in linguistic matters were absorbed by the intricacies of the Chinese characters. To be sure, such broad distinctions as 'full' vs. 'empty' words and words for actions vs. words for states were known since the 14th century (cf. Casacchia 1989: 446–447).

For a long time, Japan was under the Chinese cultural influence. At the end of the 17th century there arose a nationalist movement that attempted to consolidate the independence of Japanese thinking also in the domain of language. The most important grammarian was Nariakira Fujitani (1738–1779), who developed a descriptive framework for Japanese that is even today accepted as largely adequate (cf. Saeki & Yoichiro 1981). Its core is a four-fold division of word-classes: *na* (= nouns), *yosoi* (= verbs and adjectives), *kazashi* (= demonstrative pronouns, adverbs, conjunctions, interjections), and *ayui* (= postpositions, auxiliaries, suffixes of the verbal or adjectival predicate). Of these terms, the last three mean ornaments for the body, the head, and the feet, respectively. In this way the basic word order is encoded into the names of the word-classes. This apparatus may be illustrated by means of example (5 a); the morpheme boundary in the verb follows Fujitani's analysis:

(5) (a) Japanese
kinoo niwa de otoko ga
yesterday garden LOC man SUBJ
inu o nagu-tta
dog OBJ hit-PRET
'Yesterday a man hit a dog in the garden'

Here *kinoo* is a *kazashi*; *niwa*, *otoko*, and *inu* are *na*; *nagu-* is a *yosoi*; *de*, *ga*, *o*, and *-tta* are *ayui*. The class of *ayui* is further divided into 50 subclasses on the basis of distributional criteria. It may be added that the topic marker *wa*, which plays an important role in the Japanese sentence structure, is just one *ayui* among others (see 6 a).

(6) (a) *kinoo niwa de otoko wa*
yesterday garden LOC man TOP
inu o nagu-tta
dog OBJ hit-PRET
'The man hit a dog yesterday in the garden'

6. Conclusion

Tamil and Japanese are SOV languages; Classical Arabic is a VSO language; Sanskrit is a language with a (relatively) free word order. In spite of these differences, all traditions reviewed here share the common feature that the language which they investigate exhibits systemic suffixal (and in the case of Japanese, also postpositional) variation. Grammatical description seems to have originated in the wish to capture the nature of this variation. Classical Arabic is the prototypical instance of a language with word-**internal** (= 'Ablaut'-type) change, and yet – as was noted above – the central question in the corresponding tradition has been how to account for the oc-

currence of the three case **endings**. The interest in formal variation quite naturally leads to the establishing of distinct word classes. The word-class systems are very explicit in the Arabic and Japanese traditions. In the Tamil tradition, the original fourway classification given in *Tolkāppiyam* was not very coherent, apart from the noun vs. verb division, but it was amended by later commentators. In the Pāṇinian tradition nouns and verbs are named by means of abbreviations standing for the collections of their respective endings, and indeclinable words are considered as nouns by default, their 'endings' being automatically deleted. These systems should be compared with the four-way classification of Marcus Terentius Varro, based on the presence vs. absence of case and/or tense, or with the eight-way classification of Apollonius Dyscolus (cf. Itkonen 1991: 199, 202–203).

On the basis of the available evidence, it seems natural to assume that the lack of a systematic variation between grammatical morphemes might constitute an obstacle to the emergence of linguistic research. With the aid of this assumption, it could conceivably be explained why there never developed a genuine linguistic tradition in China. The speculative character of this type of 'explanation' goes without saying.

7. Special abbreviations

ACT	active
IND	indicative
INSTR	instrumental
P	person
SUBJU	subjunctive

8. References

Bohas, G. & Guillaume, J.-P. & Kouloughli, D. E. 1990. *The Arabic linguistic tradition.* London and New York: Routledge.

Cardona, George. 1988. *Pāṇini: His work and traditions, Vol.* I. Delhi u. a.: Motilal Banarsidass.

Casacchia, Giorgio. 1989. "Les débuts de la tradition linguistique chinoise et l'âge d'or de la linguistique impériale". In: Auroux, Sylvain (ed.). *Histoire des idées linguistiques, tome* 1. Liège-Bruxelles: Mardaga, 431–448.

Chevillard, Jean-Luc. 1996. *Le commentaire de Cēṇāvaraiar sur le Collatikāram du Tolkāppiyam.* Pondicherry: Institut français de Pondichéry.

Itkonen, Esa. 1991. *Universal history of linguistics: India, China, Arabia, Europe.* Amsterdam and Philadelphia: Benjamins.

Joshi, S. D. & Roodbergen, Jouthe Anton Foke. 1975. *Patañjali's vyākaraṇa-mahābhāsya. Kārakāhnika.* Poona: University of Poona.

Kiparsky, Paul. 1993. "Pāṇinian linguistics". In: R. E. Asher (ed.). *The encyclopedia of language and linguistics.* Oxford: Pergamon Press, vol. 1: 6, 2918–23.

Kouloughli, Djamel. 1989. "Le débuts de la grammaire hébraïque". In: Auroux, Sylvain (ed.). *Histoire des idées linguistiques, tome* 1. Liège-Bruxelles: Mardaga, 283–292.

Lehmann, Thomas. 1994. *Grammatik des Alttamil.* Stuttgart: Steiner.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory.* Amsterdam: Benjamins.

Pinault, Georges-Jean. 1989. "Le système de Pāṇini". In: Auroux, Sylvain (ed.). *Histoire des idées linguistiques, tome* 1. Liège-Bruxelles: Mardaga, 371–400.

Saeki, Tetsuo & Yoichiro, Yamanai (eds.) 1981. *Kokugogaisetsu.* Osaka.

Sastri, Subrahmanya. 1945. *Tolkāppiyam-collatikāram.* Annamalainagar: Annamalai University.

Sharma, V. Venkatarama. (ed.). 1982. *Taittirīyaprātiśākhya: with the bhāsya padakramasadana of Māhiseya.* New Delhi: Meharchand Lachhmandas.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧaaǧi's theory of grammar.* Amsterdam and Philadelphia: Benjamins.

Esa Itkonen, University of Turku (Finland)