Experimental Data vs. Diachronic Typological Data:
Two Types of Evidence for Linguistic Relativity

Wiltrud Mihatsch

SFB 441 – University of Tübingen, Germany

wiltrud.mihatsch@uni-tuebingen.de

Overview

The aim of this talk is to discuss two types of evidence for the possible influence of grammatical typological structures on categorization processes in line with recent work supporting a moderate version of linguistic relativity. Evidence from nonverbal classification tasks will be compared to evidence from lexical change, which is very likely to reflect categorization processes too.

Lexical change has become an important focus of cognitive linguistics in the past few years since lexico-semantic change is assumed to reveal fossilized categorization processes that are based on universal cognitive factors along with cultural and areal influence. Unlike in syntax or morphology, there seem to be no typologically shaped structures in the lexicon.

It will be shown that obligatory vs. optional number marking influences not only nonverbal categorization of speakers as shown by recent studies (Imai and Gentner 1997, Lucy 1992a and b, Lucy and Gaskins 2001), but also the way nouns for certain types of objects are created in different languages. We thus have to assume that typological factors do influence lexical change and that there are important links between nonverbal categorization and lexical change.

A renewed interest in linguistic relativity

Lucy (1992a) and Imai/ Gentner (1997) showed that different noun systems influence to some extent nonverbal classification tasks. Their subjects tend to classify objects

with a complex internal structure and shape (such as a wooden whisk) on the basis of shape, substances rather on the basis of the material.

Interestingly, in the case of simple objects, i.e. objects made of homogeneous material with a simple shape such as a cork pyramid, the native language of the subjects seems to determine the classification process. Speakers of English (where count nouns that are assumed to draw attention to shape prevail) prefer shape, whereas speakers of Japanese and Yucatec (where nouns are transnumeral, need a unitizer for counting, and direct attention rather to material) prefer the material for sorting simple objects.

Thus the noun system shifts the boundary between shape and material-based classification in the middle ground between strong individuals and substances.

Nonverbal categorization vs. lexical change: converging evidence?

In this talk I will compare the experimental data based on nonverbal categorization tasks with data from lexical change in a world-wide sample of about 30 languages in the domain of body part nouns – a universally named conceptual domain. Although body parts are of course not individual objects, but parts, they are perceptually on a continuum of individuation. There are substance-like parts such as hair, complex "objects" with a clear shape such as the eyeball and simple "objects" in between such as the eyebrow or eyelash, that can be conceptualized as a kind of hair or as a kind of arc.

Data from experimental nonverbal classification tasks and lexical change are of course two very distinct data types, although lexicologists assume that lexical change reflects fossilized categorization processes.

Unlike experimental data, diachronic data are quite elusive. In this talk, the emphasis will be on the comparability of these two data types. What are the main differences?

Obviously, diachronic data are hardly controllable for sociological factors such as age, sex, schooling etc.: we usually do not know who invented a new label for a concept!

Lucy (1992a) and Imai/ Gentner (1997) used specially designed objects in their experiments, but in the case of diachronic data one has to find existing concepts that are universally labelled and fulfil the criteria of the different object types as distinguished by Imai/ Gentner (1997). In particular, how can we find widespread concepts that correspond to simple objects?

Body parts are good candidates, but have the disadvantage of being parts, not whole objects. Besides many body-part lexemes are extremely stable (s. Swadesh 1955),

e.g. those denoting HEAD, HAIR, EYE, EAR, SKIN, BLOOD. They usually do not reveal any diachronic paths. We therefore have to analyse less stable body part nouns such as nouns designating EYEBALL, EYEBROW, EYELASH or EYELID instead of EYE or HAIR. Even then, it is not always easy to detect lexical changes. Many languages do not have any written records and therefore we do not have access to diachronic data. Here, polysemy and morphological transparency are indirect sources for diachronic information.

Furthermore, lexical change is not a binary decision between shape and substance. Other conceptual sources have to be taken into account as well as areal and cultural influences that go beyond simple perceptual properties of objects. Unlike nonverbal classification tasks lexical change is also partly determined by the types of word formation processes available in a language (cf. Koch 2001).

Since unlike Imai/ Gentner (1997) and Lucy (1992a) I compare 30, not 2 languages, the situation as to plural marking is more complex and I find a greater variety of noun types. There are languages where count nouns prevail and where the plural is obligatorily marked such as English. Other languages sometimes have an obligatory plural only for higher segments of the animacy hierarchy or a facultative plural (Corbett 2000:56f., cf. Lucy 1992a:69ff.) or no plural at all. Some of these languages have numeral classifiers that transform transnumeral nouns into countable units, e.g. Japanese or Lahu. Where do we draw the line? It will be argued that since the linguistic patterns that are likely to influence thought are unconscious and obligatory distinctions (see Lucy 1992b:194), there are two groups of languages as to plural marking: languages with an obligatory plural for most nouns and the cross-linguistically more widespread type of language (Rijkhoff 2002:38) with no obligatory plural marking, at least in the segment of the animacy hierarchy we are interested in.

Of course we also have to be aware that the lexicon contains many traces of lexical change dating from different periods. Thus a change of language type at some stage might have to be taken into account. So the problem of the noun type has to be examined very carefully.

Despite all the aforementioned problems with diachronic data, these data also have certain advantages, namely a greater quantitative relevance than the experimental data since in the case of lexical change a whole speech community has adopted the new conceptualization path.

Even more importantly, the processes of change are not produced by artificial experimental situations, but presumably by natural everyday conceptualization.

On a more practical level, experiments are very costly and time-consuming, whereas lexical data are already there and waiting for analysis – although the analysis of lexical change is not a trivial task either.

All these differences and problems will be discussed on the basis of four concepts and their diachronic sources in a sample of 30 languages and the results obtained by Lucy (1992a) and Imai/ Gentner (1997).

Surprising results

Considering all the problems mentioned so far one would expect that data from lexical change would give us a rather messy picture. Remarkably, despite the differences between these two types of evidence and their respective limitations the results from nonverbal classification tasks and lexical change match rather well:

Worldwide, the sources tend to be based on shape for lexemes meaning EYEBALL and on material for those meaning HAIR, whereas the greatest diversity of paths can be found for simple objects such as EYELASH in correlation with the noun system. For example, in a language where count nouns prevail, such as English, we find *eyelash* (lit.:EYE + WHIP-LASH) based on shape, in Tzeltal, a language with transnumeral nouns, we find *stsotsel sit* EYELASH (lit. HAIR + EYE) based on material, quite comparable with the observations made by Imai/ Gentner (1997) and Lucy (1992a).

This study thus explores for the first time the interaction of perceptual and typological factors in lexical change and shows that the two data types, experimental and diachronic data, very clearly converge and point to a moderate version of linguistic relativity.

References

Corbett, G. (2000). *Number*. CUP. Cambridge (Cambridge Textbooks in Linguistics).

Imai, M. and D. Gentner (1997). A Cross-Linguistic Study of Early Word Meaning: Universal Ontology and Linguistic Influence. *Cognition*, 62: 169-200.

Koch, P. (2001). Lexical typology from a cognitive and linguistic point of view. In M. Haspelmath, E. König, W. Oesterreicher, W. Raible, eds., *Language Typology and Language Universals. An International Handbook,* De Gruyter, Berlin /New York (Handbücher der Sprach- und Kommunikationswissenschaft, 20.2), pp. 1142-1178.

Lucy, J. (1992a). *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. CUP, Cambridge.

Lucy, J. (1992b). *Language diversity and thought. A reformulation of the linguistic relativity hypothesis*. CUP, Cambridge (Studies in the Social and Cultural Foundations of Language; 12).

Lucy, J. and S. Gaskins (2001). Grammatical categories and the development of classification preferences: a comparative approach. In M. Bowerman and S. Levinson, eds., *Language Acquisition and Conceptual Development*, CUP, Cambridge, pp. 257-283.

Rijkhoff, J. (2002). *The Noun Phrase.* OUP, Oxford (Oxford Studies in Typology and Linguistic Theory).

Swadesh, M. (1955). Towards greater accuracy in lexico-statistic dating, *IJAL* 21: 121-137.