

Jean-Pierre Colson

Editorial: Phraseology at the intersection of grammar, culture and statistics

In the ruins of the Roman city of Timgad (Algeria), a stone was found bearing the following Latin inscription: “*Venari, lavari, luderi, rideri, occ (=hoc) est vivere*”: to hunt, to bathe, to play, to laugh, that is life!

The translation of *hoc est vivere* might also be *this is the life!*, a phrase used in English when you are very much enjoying the situation you are in. It is fascinating to see that a phrase used in Latin has more or less survived the centuries and is still used in English, and also in other European languages (for instance in French: *c'est la belle vie*, *c'est la grande vie*). A simple phraseological unit such as *this is the life* suffices to illustrate how grammar, culture and statistics are inextricably intertwined in phraseology.

In the first place, the grammatical structure of *this is the life* is special: the definite article is used in the construction, yielding a phraseological unit whose meaning is different from *that's life*, a phrase that usually means *c'est la vie*, also used in English: that's how things happen, you can't do anything about it.

The link with culture and history is also very clear, as we have noted that a very similar phrase seemed to exist in Latin. The Timgad inscription may also be an ironical variant of a quotation from Greek or Roman philosophers, as we can for instance read in Seneca “*esse, bibere, frui patrimonio, hoc est vivere, hoc est se mortalem esse meminisse*” (*Epistulae ad Lucilium*, Book 20, CXXIII, 10): eating, drinking, enjoying your estate, that is living, that is the reminder that you are mortal (translation by Catharine Edwards, *Death in Ancient Rome*, Yale University Press, 2007). Phraseological units related to *life / living* in English and in European languages did not appear out of the blue: there is probably a link with the Greek and languages, and therefore with the vision of life in ancient times.

Phraseology is deeply rooted in culture and history. The clear implication is that phraseology should take into account the amazing diversity of languages and cultures. Although there is no complete agreement on the ranking of the most spoken languages of the world, Chinese (Mandarin), English, Spanish and Hindi are often cited as the four most spoken languages, and we should never forget that three of them are Indo-European languages, probably sharing many common phraseological features.

If phraseology is to become a theory, or if we wish to make claims about the theoretical underpinnings of phraseological hypotheses, they should not only be tested against English and other European languages, but against very different languages selected from the whole

palette of linguistic diversity, from the very agglutinating Inuit languages at one extreme, to the most isolating languages at the other end, with for instance Chinese languages and Vietnamese. Simply greeting someone in European languages, as in *How are you?* may seem quite ordinary, but it is child play to find very different ways of expressing this greeting in the languages of Asia, as in (Mandarin) Chinese 你好 (nǐ hǎo, literally: “you good”) but also (你)吃了吗 (Nǐ chī le ma, literally *Have you eaten?*).

Having stressed that phraseology is closely linked to grammar and culture is not all: we also have to explain why and how native speakers are able to find the natural associations between words. In the case of many phraselogical units such as *This is the life*, it is also amazing to think that an algorithm can extract them from a large corpus, without receiving any information as to their meaning: some aspects of phraseology have to do with the statistical associations between words, a domain that is still largely a *terra incognita*: how much of language structure and meaning can be predicted? What is the contribution of phraseology to the statistical extraction of meaning?

The articles presented in this volume reflect the diversity and richness of research on phraseology: syntactic and semantic issues, pedagogical aspects, historical and cultural elements, statistical associations. We also welcome an attempt to shed new light on the *chengyu*, a specific type of phraseological unit in Chinese.