

Chapter 1

Conventions for citing and glossing multilingual examples of multiword expressions

Stella Markantonatou

Institute for Language and Speech Processing, Athena RIC, Greece

Carlos Ramisch

Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France

Agata Savary

University of Tours, LIFAT, France

Veronika Vincze

University of Szeged, Hungary

This document extends the Language Science Press conventions for citing and glossing multilingual examples, so as to adapt them to multiword expressions.

Papers and chapters about multiword expressions may address a large number of languages, as in the case of books published in the *Phraseology and Multiword Expressions* (PMWE) series at *Language Science Press*. We put forward notational conventions which might become a standard for citing and glossing multilingual MWE examples, both in the PMWE series and in other publications. We illustrate the proposed conventions by the *numbered examples* (1) to (4). Each numbered example contains:

1. a sample use of the VMWE, followed by the 2-letter ISO 639-1 language code

2. a transcription, if the language of the example is written with a script different from the one used for the main text,¹
3. a gloss following the Leipzig Glossing Rules,²
4. a literal translation, followed by an idiomatic translation in single quotes.

For English examples, items (2)–(4) are irrelevant or optional but idiomatic translation might sometimes be useful to ease the comprehension by non-native readers. For right-to-left languages (e.g. Farsi or Hebrew), item (1) is spelled right-to-left, item (4) left-to-right and items (2)–(3) left-to-right within components, and right-to-left from one component to another. Lexicalized components of the VMWE, i.e. those which are always realized by the same lexeme are highlighted in bold face.

- (1) She reluctantly **took on** this task. (EN)
 ‘She reluctantly agreed to be in charge of this task.’
- (2) Ida **skriva glavo v pesek**. (SL)
 Ida hide.3.SG head in sand
 Ida hides her head in the sand. ‘Ida pretends not to see a problem.’
- (3) Η Ζωή παίρνει μία απόφαση. (EL)
 i Zoi perni mia apofasi
 the.FEM.SG Zoe.FEM.SG take.3.SG a decision
 Zoe takes a decision. ‘Zoe makes a decision.’
- (4) به قدر کافی برای من خواب دیده است (FA)
 ast **dide khab** man **baraye** kafi qadre be
 is seen sleep me for enough quantity to
 He had enough sleep for me. ‘He has many plans for me.’

These examples can be typeset by the dedicated LSP commands, e.g. example (3) is typeset like this:

```
\ea \label{preface:el:take-decision}
\settowidth \jamwidth{(EL)}
\glll Η Ζωή \lex{παίρνει} μία \lex{απόφαση}. \l
```

¹For instance, transcription is needed for Bulgarian, Greek, Farsi and Hebrew examples in this volume. Conversely, examples in English, or any other language using Latin script, would require transcriptions in texts written in Cyrillic, Greek, Arabic or Hebrew script.

²<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>

```
i zoi perni mia apofasi \\
the\textsc{.fem.sg} Zoe\textsc{.fem.sg} take.\textsc{3.sg} a decision \\
\jambox{(EL)}
\glt Zoe takes a decision. ‘Zoe makes a decision.’
\z
```

In-line examples, used for brevity, are preceded by the 2-letter language code, contain items (1), (2) if relevant, and (4) only, and the idiomatic translation (if any) is introduced by a double arrow ‘ \Rightarrow ’. For instance, an in-line example corresponding to numbered example (2) would be the following: (SL) *Ida skriva glavo v pesek* ‘Ida hides her head in the sand’ \Rightarrow ‘Ida pretends not to see a problem’. If the language under study is written with a non-Latin alphabet, the inline example should not be in italics, and the transliteration should be included in parentheses, e.g. (EL) Η Ζωή *παίρνει* μία *απόφαση* (I Zoi perni mia apofasi) ‘The Zoe takes a decision’ \Rightarrow ‘Zoe makes a decision’. To keep such examples reasonably short, the first item can be omitted and only the transliterated example is kept: (EL) I Zoi perni mia apofasi ‘The Zoe takes a decision’ \Rightarrow ‘Zoe makes a decision’. The literal or the idiomatic translation are sometimes superfluous or too verbose, and can be skipped, as in: (EL) I Zoi perni mia apofasi ‘Zoe makes a decision’.

The typesetting commands for both numbered and in-line examples for \LaTeX can be found in the GitHub repository containing the source codes of this volume, accessible from its webpage (file `multilingual-mwe-examples.tex`).³

³<http://langsci-press.org/catalog/book/204>