

ΙΔΙΟΝΙΔΙΟΝ

MANUAL ENGLISH VERSION: 1.0

MANUAL ENGLISH VERSION: 1.0

ΙΔΙΟΝ
IDION MANUAL ENGLISH VERSION: 1.0

Institute for Language and Speech Processing

November 8, 2018

2

3

Contents

| | | |
|----------|-------------------------------------|-----------|
| 1 | A Guide to IDION | 5 |
| 2 | Installation instructions | 7 |
| 2.1 | IDION installation | 7 |
| 2.2 | IDION execution | 7 |
| 2.3 | IDION updates | 7 |
| 2.4 | Creation of a new lexicon | 8 |
| 2.5 | Opening a lexicon | 8 |
| 2.6 | Saving a lexicon | 8 |
| 2.7 | Adding a new lemma | 8 |
| 3 | General Remarks | 9 |
| 4 | IDION tabs | 11 |
| 4.1 | GENERAL TAB | 11 |
| 4.1.1 | The Meaning feature | 11 |
| 4.1.2 | The FORMS feature | 13 |
| 4.2 | USE/RELATIONS TAB | 25 |
| 4.2.1 | The USE feature | 25 |
| 4.2.2 | The RELATIONS feature | 27 |
| 4.3 | CORPUS/DIAGNOSTICS TAB | 29 |
| 4.3.1 | The CORPUS feature | 29 |
| 4.3.2 | The DIAGNOSTICS feature | 30 |
| 5 | Grundlagen | 35 |
| 5.1 | 1. Thema | 35 |
| 5.1.1 | 1. Unterthema | 35 |
| 6 | Fazit | 37 |
| | Anhang | 41 |

Chapter 1

A Guide to IDION

IDION is a lexicographic environment for the documentation of Greek idioms. The technical term *Multi Word Expressions* (MWE), with the equivalent Greek term *Πολυλεκτικές Εκφράσεις* (ΠΛΕ), is often used for this type of linguistic structure.

MWEs may assume the functions of different parts of speech: they can be nominal (1), verbal (2), adjectival (3), adverbial (4) and prepositional (5).

- (1) παιδική χαρά
childish happiness
'playground'
- (2) έκανα μαύρα μάτια να δω κάποιον
made.1SG black eyes.ACC to see.1SG someone
'I have not met someone for a long time'
- (3) ο απαλός σαν το χάδι αέρας της Αττικής
the soft like the stroke wind of Attica
'the very soft wind of Attica'
- (4) μόνο και μόνο
only and only
'only to'
- (5) κατά καιρούς
in times
'occasionally'

This is a brief guide to the usage of IDION. A detailed description of the lexicographic environment is given, along with instructions on how to edit its contents.



Chapter 2

Installation instructions

PANAGIOTI, PLEASE REVISE THE SECTION 2

There is no need for installation, remove this chapter.

2.1 IDION installation

IDION application does not need to be installed. Decompress the file [multiword-editor-app-1.0.0-SNAPSHOT.zip](#). Then click on the file [mweditor](#).

2.2 IDION execution

The application is executed by double-clicking the Then click on the file [mweditor.exe](#) in the file [bin](#).

2.3 IDION updates

The application can be updated via the menu as follows:

- Select *Help*.
- Select *Check for Updates*.
- If there are updates available, click on *Next*.

2.4 Creation of a new lexicon

A LEXICON is a collection of lemmas. Currently, IDION contains nearly 350 verbal MWEs. Lexica are just files for IDION. To create a new lexicon, in other words to open a new file, the following steps are followed:

- From *File* select *New File*.
- In the emerging window select *NLP* and the type *Empty MultiWord Lexicon*. Then press *Next*.
- In the emerging window name your file and specify the file location where it will be saved.
- Finally, press *Finish*.

2.5 Opening a lexicon

- From *File* select *Open File*
- Choose the file you wish to edit/read.

2.6 Saving a lexicon

From *File* select *Save* or *Save As*.

2.7 Adding a new lemma

To add a new lemma:

- Right-click on the list in the left side of the screen.
- In the emerging menu select *Add*.

Chapter 3

General Remarks

1. Expressions are introduced in the lemma form: *κάνω την πάπια* and NOT in the form *Ο Χ κάνει την πάπια* or any other form. If the expression:

- Has a fixed subject and a variable pronominal object then the first person is selected for the object: *με κόβει η λόρδα*
- Has a fixed subject and a variable dative genitive attached to a fixed part of the expression, including the verb, then the first person is selected for the object: *μου κόβονται τα χέρια*

2. Synonymous MWEs with minimally different fixed constituents are encoded as individual lexical entries.

- (6) a. *κόβω κάποιον λωρίδες*
cut somebody.ACC stripes.ACC
- b. *κόβω κάποιον φέτες*
cut somebody.ACC slices.ACC
'to kill somebody'
- (7) a. *πίνω το πικρό ποτήρι*
drink the bitter.ACC glass.ACC
'I experience a sad situation'
- b. *καταπίνω το πικρό ποτήρι*
swallow the bitter.ACC glass.ACC
'I experience a sad situation'

MWEs (6a) and (6b) are synonymous and have minimally different fixed constituents. Since the fixed parts are essentially different, we will consider them distinct MWEs and for each of the two MWEs a distinct entry will be encoded. Encodings of the form *κόβω κάποιον λωρίδες/φέτες* are not allowed. The synonymy relation has to be

encoded with the USE/RELATIONS tab (and the Lexical variant relation whereas it applies). Similarly, (7a) and (7b) show synonymous MWEs with different verbs and are encoded as individual lexical entries.

Caution:

Differences among fixed constituents of a MWE concerning diminutives, specific adjectives, a definite/indefinite article, alternative prepositions, different form of adverb/preposition, additional constituents that modify/emphasize¹ the conveyed meaning (ανάβω όλα τα λαμπάκια κάποιου, πίνω το πικρό ποτήρι ως τον πάτο) are encoded as alternative forms (see also FORMS under the GENERAL tab).

3. Distinct meanings of the same MWE form are encoded in separate entries: (8a) means that someone was imprisoned, while (8b) that someone has lost money and is in debt.

- (8) a. Ήθελε φακελάκια ο γιατρός και (ΕΠΙΤΕΛΟΥΣ) τον βάλανε μέσα!!!
 wanted bribe the doctor and (AT LAST) him put in!!!
 ‘This doctor wanted a bribe and (AT LAST) he was imprisoned.’
- b. Το σπίτι που έβαλε μέσα χοντρά τη Σάρα Τζέσικα Πάρκερ...
 the house.NOM which put in a lot the Sarah Jessica Parker...
 ‘This house costed Sarah Jessica Parker a lot of money, and now she is in debt..’

4. Semantic relations among MWEs are encoded with the USE/RELATIONS tab.

Relations among MWEs are encoded only in the USE/RELATIONS tab. Relations should be encoded carefully. For instance, the MWE τα παίρνω (στο κρανίο) ‘I get very angry’ is not synonym with the MWE κόβω τα φτερά κάποιου ‘I dishearten someone’. In other words, the editor should be mindful of slight or not so slight differences in meaning.

5. Last but not least, before inserting a new MWE to IDION, the editor should check with the text search facility if the expression has already been encoded in IDION, in order to avoid duplicates.

¹Emphasis is treated as an alternative form at the moment; a note is made about emphatic forms in the

Chapter 4

IDIION tabs

IDIION is organized in 8 sections, which are placed in 5 tabs. These tabs are the following:

1. GENERAL/ FORMS tab
2. USE/ RELATIONS tab
3. CORPUS/ DIAGNOSTICS tab
4. HTML comment tab
5. CHANGELOG tab

Each tab is described in detail below.

4.1 GENERAL TAB

4.1.1 The Meaning feature

At least one, and possibly multiple definitions of the MWE are encoded in the GENERAL tab in Greek, English, French and, possibly, other languages. Definitions of the encoded MWE should be conceived as different senses, which capture slight semantic differences. This can be checked, to some extent, with substitution. The MWE *παίρνω κεφάλια* is ambiguous between the senses ‘to impose sanctions to someone’ and ‘to punish someone’ therefore the example (9) may receive two interpretations. The MWE is encoded with two senses.

- (9) Πήραν «κεφάλια» στη διοίκηση της Eurobank.
took heads in the administration the.GEN Eurobank
'Members of the Eurobank administration were punished/fired.'

On the other hand, it should be kept in mind that **phrases with multiple definitions which present great differences in meaning should be split into different lemmas**. For instance, the MWE αφήνω σέκο κάποιον has the meanings ‘to knock someone off’ and ‘to leave someone stunned, astounded, speechless’ which can hardly be substituted for each other as synonyms.

Each definition should be introduced in a separate box (as a new item) along with encapsulated boxes providing the English and the French translation. If an English or French equivalent expression exists, it should be provided in the appropriate box and it should take precedence over the definition. In this case no English definition of the expression is needed.

A box for possible comments is provided, be it for the Greek or the English/ French MWEs and/or definitions.

| | | |
|----------------|--|--|
| Lemma | αρπάζω την μπουκιά (μέσα) από το στόμα κάποιου | |
| Meaning * | Definition * | αρπάζω κάτι από κάποιον την τελευταία στιγμή, πάνω που θα μπορούσε να το απολαύσει, ή να το αναλάβει |
| | Comment | |
| Language * | Greek | |
| Translations * | Definition | to get or enjoy something that someone was close to take on, or enjoy |
| | Comment | |
| | Language | English |
| | | ✘ Remove the Item |
| | Definition | enlever le pain de la bouche à quelqu'un |

Figure 4.1: The meaning tab of the MWE αρπάζω την μπουκιά (μέσα) από το στόμα κάποιου.

The following should be observed:

1. Each definition should start with a lowercase letter. The equivalent English/ French (or other language) expression (when available) should also start with a lowercase letter. No use of a full stop is needed.
2. The English/ French word, expression or the defining phrase should always be in an encapsulated box labeled with the appropriate language in the first definition box.

3. The full form of the pronouns κάποιον (someone) and κάτι (something) should be used for free NPs rather than the abbreviations κπ (sb) and κτ (sth) (10). If both κάποιον and κάτι apply, then they should be written as κάποιον/ κάτι.
4. The definition provided in the Lemma box may contain free objects represented with κάποιον/ κάτι. These objects should follow, if possible, the fixed parts of the expression. For instance, αφήνω κάποιον στα κρύα του λουτρού should be avoided and αφήνω στα κρύα του λουτρού κάποιον should be used instead. This may not be possible in the English/ French expression, but it is of crucial importance to have a uniform way of encoding.

- (10) a. βγάζω την γλώσσα μου σε κάποιον
take out.1SG the tongue.ACC my to someone
“I tease someone”
- b. βγάζω την μάσκα κάποιου
take out.1SG the mask.ACC someone.GEN
“I unmask/reveal someone”

The following sources can be used to retrieve English translations:

| |
|---|
| Έντυπες πηγές |
| Ελληνο-αγγλικό λεξικό Κοραΐς/Εκδόσεις Πανεπιστημίου Πατρών Oxford Dictionary of English Idioms/Oxford University Press |
| Ηλεκτρονικές πηγές |
| https://www.phrases.org.uk/ |
| https://idioms.thefreedictionary.com/ |
| https://dictionary.cambridge.org |
| https://www.theidioms.com/ |
| https://www.merriam-webster.com/ |
| https://www.urbandictionary.com/define.php?term=slang |
| https://americanliterature.com/english-language-idioms |
| http://www.readingmatrix.com/articles/john_iontas/article.pdf |
| https://www.translatum.gr/forum/index.php?board=81.0 |

Table 4.1: Sources for English translations.

4.1.2 The FORMS feature



Exhaustive morphological and syntactic description of the MWE is encoded in this section, which is placed under the semantics of the encoded MWE. Encoding is theory neutral and is aimed to serve as a basis for any type of parser. To this end, morphological tags are standardized (PAROLE http://nlp.ilsp.gr/nlp/tagset_examples/tagset_en/). Similarly, the encoded syntactic relations are kept theory independent, therefore minimal, and include:

- Information about free constituents
- Phrasal information for constituents that are realized with full phrasal structures
- Lexical information for constituents that are realized with weak pronouns
- Binding and control relations
- Delineation of fixed/semi-fixed strings

A tabular arrangement with five columns is used together with controlled vocabularies that help to minimize the number of errors. The label ‘tokens’ is used to cover both the lexical and the phrasal parts of a MWE. Each token occupies a row in the tabular format. New tokens can be added and existing ones can be removed with the use of the corresponding buttons.

Forms Comment

| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optional Bound | Bound Controlled By | Operations |
|--------|------|---------|----------|----------|------------|-----------|--------------------------|--------------------------|--|
| | NP-† | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | NC <input type="button" value="DELETE"/> |
| | LEM | τρώω | | VbAv | VbAv | | <input type="checkbox"/> | <input type="checkbox"/> | NC <input type="button" value="DELETE"/> |
| | LEM | με | με | AsPpSp | AsPpSp | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC <input type="button" value="DELETE"/> |
| | LEM | χρυσός | χρυσά | AjBaNePl | AjBaNePIAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC <input type="button" value="DELETE"/> |
| | LEM | κουτάλι | κουτάλια | NoCmNe | NoCmNePIAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC <input type="button" value="DELETE"/> |

Figure 4.2: The Forms section.

It is necessary to clarify that each phrase (MWE) should be encoded in the FORMS tab with its LEMMA form as it was introduced in the General Remarks and not in any other form.

Variants of a MWE are encoded with additional forms in the FORMS tab. As variants of a MWE we encode cases where the following occur:

- a diminutive as an alternative (eg. βάζω την ουρά/την ουρίτσα μου)
- a slightly different form of a lemma (eg. ανεβάζω κάποιον στα ουράνια/στα μεσουράνια)
- an alternation between a definite/indefinite determiner or additional constituents (drawn from a very small set of specified adjectives or PPs) that modify the semantics of the MWE (e.g. ανάβω όλα τα λαμπάκια κάποιου, πίνω το πικρό ποτήρι ως/μέχρι τον πάτο).

Columns 1 (Type) and 2 (Lemma):

Free, semi-fixed and fixed parts of the MWE are encoded in the order that they appear in the MWE. So in column 1 named Type a controlled vocabulary has been created in order to denote the nature of the various parts of the MWE.

| |
|--|
| LEMMA: declinable |
| WF (WORDFORM): non-declinable |
| COMPL: designates the notion COMPLEMENTIZER (να, θα, που, ...) |
| VP: any controlled but otherwise free VP normally following a complementizer |
| NP-NOM |
| NP-NOM-anim |
| NP-NOM-nonanim |
| NP-GEN |
| NP-GEN-anim |
| NP-GEN-nonanim |
| NP-ACC |
| NP-ACC-anim |
| NP-ACC-nonanim |

Table 4.2: The controlled vocabulary used in FORMS.

The controlled vocabulary provided in Column 1 (Type) is given in (4.2) (see (4.3) for an explanation of the abbreviations used in the labels).

In the second column (Lemma), the lemmatized form of each token is given, provided of course that the token type is appropriate for such labeling. Otherwise, whereas the token type doesn't allow a specific token to be provided, this box remains empty.

Binding relations (figure 4.3) and optionality of tokens (figure 4.4) are also encoded in this area with the appropriate buttons that are provided next to each token.

| Comment | Type | Lemma | WF | Tag | Tag | WWS Index | Optional/Bound |
|---------|------|--------|--------|--------|------------|-----------|-------------------------------------|
| | NP-↑ | | | | | | |
| | LEM | αφίγω | | VbAv | VbAv | | |
| | NP-↓ | παιδιά | παιδιά | NoCmNe | NoCmNePIAc | | |
| | LEM | πίσω | πίσω | AdXxBa | AdXxBa | | |
| | PnG | μου | | PnGeWe | PnGeWe | | <input checked="" type="checkbox"/> |

Operations: NO, TOKEN 1, TOKEN 2, TOKEN 3, TOKEN 4, TOKEN 5, TOKEN 6, TOKEN 7, TOKEN 8, TOKEN 9, TOKEN 10, TOKEN 11, TOKEN 12, TOKEN 13, TOKEN 14, TOKEN 15, NC, Add Token, Remove the item

Figure 4.3: Binding relations depicted in the DB.

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|-------|-------|----------|------------|-----------|-------------------------------------|
| NP-N | | | | | | <input type="checkbox"/> |
| LEM | βγάζω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | μόνος | | AjBaMaS | AjBaMaSgNm | 1 | <input checked="" type="checkbox"/> |
| PnGi | μου | μου | PnGe | PnGe | 1 | <input checked="" type="checkbox"/> |
| LEM | ο | το | AtDfNeSς | AtDfNeSgAc | 2 | <input type="checkbox"/> |
| LEM | φιδι | φιδι | NoCmNe | NoCmNeSgAc | 2 | <input type="checkbox"/> |
| LEM | από | από | AsPpSp | AsPpSp | 2 | <input type="checkbox"/> |
| LEM | ο | την | AtDfFeSς | AtDfFeSgAc | 2 | <input type="checkbox"/> |
| LEM | τρύπα | τρύπα | NoCmFe | NoCmFeSgAc | 2 | <input type="checkbox"/> |
| PnGi | μου | του | PnGeWe | PnGeWe | 2 | <input checked="" type="checkbox"/> |

Figure 4.4: Optionality of tokens in a MWE.

The encoder works simultaneously with Columns 1 and 2. Data are introduced in Column 2 while Column 1 is used to provide the appropriate label from the controlled vocabulary and to indicate optionality where it applies. More precisely (A-D):

A. Semi-fixed and fixed parts are treated as individual words and not as phrasal constituents and are all listed in their lemma form:

- (11) κάνω μαύρα μάτια να σε δω
do.1SG black eyes to see.1SG you
'I haven't seen you for a long time'

Example (11) contains two verbal heads, namely *κάνω* and the semi-fixed *δω*. Of them the first is free while the second has to be always in Past Subjunctive form and is controlled by the subject. *κάνω* will be assigned the label LEMMA and in the fourth column it will be assigned only Part of Speech and active voice, which are the only constraints that apply to it. *δω* will be assigned the label LEMMA but in the third and fourth column it will be assigned two constraints. In the third column it will be assigned the "controlled by" tag. In the fourth column it will be assigned constraints on morphology: *VbMnIdXxXxXxPeAvXx* (figure 4.5).

| Comment | | | | | | | | | |
|---------|--------|-------|--------|---------------------|-----|-----------|---------------|----------|------------|
| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound | Bound By | Operations |
| NF | | | | | | | | | NO |
| LE | κάνω | | Vb | Vb | | | | | PRO |
| LE | μαύρος | μαύρα | AjBaNi | AjBaNePIAc | 1 | | | | TOKEN 1 |
| LE | μπί | μάτια | NoCml | NoCmNePIAc | 1 | | | | TOKEN 2 |
| CC | να | να | PtSj | PtSj | | | | | TOKEN 3 |
| LE | βλέπω | | VbMni | VbMniXXXXXXXXPeAvXx | | | | | TOKEN 4 |
| NF | | | | | | | | | TOKEN 5 |
| | | | | | | | | | TOKEN 6 |
| | | | | | | | | | TOKEN 7 |
| | | | | | | | | | TOKEN 8 |
| | | | | | | | | | TOKEN 9 |
| | | | | | | | | | TOKEN 10 |
| | | | | | | | | | TOKEN 11 |
| | | | | | | | | | TOKEN 12 |
| | | | | | | | | | TOKEN 13 |
| | | | | | | | | | TOKEN 14 |
| | | | | | | | | | TOKEN 15 |
| | | | | | | | | | DELETE |
| | | | | | | | | | DELETE |
| | | | | | | | | | DELETE |

Figure 4.5: Example of a controlled verb in a MWE.

- (12) μένω ταπί και ψύχραμος
 stay broke and cool
 'I am broke'

Example 12 contains the cranberry word ταπί. Ταπί will be assigned the label WF and in the fourth column it will not be assigned any tag at all.

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|---------|------|---------|------|-----------|---------------|
| NP-I | | | | | | |
| LEM | μένω | | VbAv | VbAv | | |
| WF | ταπί | ταπί | | | | |
| LEM | και | και | CjCo | CjCo | | |
| LEM | ψύχραμο | | AjBaXxX | | | |

Figure 4.6: The cranberry word ταπί as depicted in the FORMS Tab of the DB.

In the same example (12), the adjective ψύχραμος that is controlled for number and gender by the free NP (the free subject of the MWE) will be assigned the label LEMMA in the second column and in the fourth column will not be assigned any constraint because all constraints are supplied by the control relation.

B. Prepositional heads that are indeclinable by definition, are considered fixed. They may accept fixed (για ένα κομμάτι ψωμί), semi-fixed (στο νυχάκι NP-GEN-anim) or free nominal complements that are encoded accordingly

C. One or more strings of one or more fixed wordforms may be indeclinable within the MWE: μαύρα μάτια (11) is an example of a fixed string

D. Free nominal parts of the MWE may be phrasal or lexical:

| |
|----------------|
| VP |
| hline NP-NOM |
| NP-NOM-anim |
| NP-NOM-nonanim |
| NP-GEN |
| NP-GEN-anim |
| NP-GEN-nonanim |
| NP-ACC |
| NP-ACC-anim |
| NP-ACC-nonanim |

Table 4.3: The controlled vocabulary used in FORMS.

1. As lexical are treated the free possessive pronouns that are very often bound either by the subject or the object. They are assigned the lemma *μου*. The label attached to them is standard and it is the “bleached” Parole tag that includes only the part of speech label (Pn), the case label (Ge) and the indication “weak pronoun” (We). The number and gender labels are left unspecified and will be specified at parsing parsing by the binding mechanisms.
2. As phrasal are treated:
 - (a) Free nominal constituents (often analyzed as free subjects, objects, prepositional complements and genitive NPs modifying nouns) that are labeled for their phrasal category, case and animosity. The controlled vocabulary for this purpose is given in (4.3).

Labels without specification about animosity must be understood as the disjunctive form of the corresponding labels that specify animosity. Thus NP-ACC must be understood as NP-ACC-anim OR NP-ACC-nonanim. For instance, the MWE *αφήνω κατά μέρος κάποιον/ κάτι* has a free noun phrase (the object of the verb), which will be denoted with the term NP-ACC (4.7):

| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optional |
|--------|------|-------|-------|--------|------------|-----------|--------------------------|
| | NP-† | | | | | | <input type="checkbox"/> |
| | LEM | αφήνω | | VbAv | VbAv | | <input type="checkbox"/> |
| | LEM | κατά | κατά | AsPpSp | AsPpSp | 1 | <input type="checkbox"/> |
| | LEM | μέρος | μέρος | NoCmNe | NoCmNeSgAc | 1 | <input type="checkbox"/> |
| | NP-/ | | | | | | <input type="checkbox"/> |

Figure 4.7: Example of the controlled vocabulary of the NP-ACC LABEL.

- (b) Free verbal constituents: If there are no constraints on a subordinated clause, then the following are encoded: the complementizer that introduces the clause, if it is fixed, and then an VP. Otherwise, no complementizer is specified. Figure 4.8 shows the encoding of the MWE βάζω μπρος να VP:

| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optic |
|--------|------|-------|----|--------|--------|-----------|--------------------------|
| | NP-τ | | | | | | <input type="checkbox"/> |
| | LEM | βάζω | | VbAv | VbAv | | <input type="checkbox"/> |
| | LEM | μπρος | | AdXxBa | AdXxBa | | <input type="checkbox"/> |
| | CON | να | | | | 1 | <input type="checkbox"/> |
| | VP | | | VbPe | VbPe | 1 | <input type="checkbox"/> |

Figure 4.8: A VP in an encoded MWE.

Columns 3 and 4: Marking of the fixed strings and the constrained semi-fixed strings is done explicitly on these columns. A MWE may contain one (13) or more (14) fixed strings. On Column 3 the fixed word forms are given; their PAROLE tags are specified on the fourth column (Figure 4.9).

| Type | Lemma | WF | Tag | WWS Index | OptionalBound |
|------|----------|---------|------------|-----------|-------------------------------------|
| NP-N | | | | | <input type="checkbox"/> |
| LEMI | πίνω | | VbAv | | <input type="checkbox"/> |
| LEMI | ο | το | AtIdNeSgAc | 1 | <input checked="" type="checkbox"/> |
| LEMI | αμίλητος | αμίλητο | AjBaNeSgAc | 1 | <input type="checkbox"/> |
| LEMI | νερό | νερό | NoCmNeSgAc | 1 | <input type="checkbox"/> |

Figure 4.9: The fixed word forms with their PAROLE tags.

On Column 5 each string is assigned a different number. We make sure that all the words in the same fixed string are assigned the same number and that all the fixed strings in the MWE are marked, see for instance figure (4.7). We use the term *Word_Wwith_Spaces* (WWS) to denote a fixed string. The WWSs occurring in examples (13) and (14) are indicated with subscripts. Please notice that no theoretical commitment is made as regards the syntactic function of the fixed strings.

- (13) Τάζω λαγούς₁ με₁ πετραχήλια₁.
 promise rabbits.ACC with vestments.ACC
 ‘I promise something impossible.’
- (14) κάνω μαύρα₁ μάτια₁ να₂ δω₂ NP.ACC
 do black1 eyes to see NP.ACC
 ‘I have not seen someone for a long time and I suffer.’

A semi-fixed constituent of the type discussed before (11) or of the type given below (15) is assigned the necessary constraints. (15) contains the cranberry word ταπί that will be a word form with no tag assigned and the adjective ψύχραμος that is controlled for number and gender by the free NP-NOM-ANIM (the free subject of the verb MWE) therefore no special constraints will be assigned to it.

- (15) Ο Νίκος έμεινε ταπί και ψύχραμος.
 The Nikos stayed broke and cool.ACC
 ‘Nikos was broke.’

It should be repeated that we introduce a new form if we have variations of the MWE due to the occurrence of diminutives or slightly different versions of a lemma (for instance, ουράνια-μεσουράνια), specified adjectives, an alternation of definite/indefinite determiner and an alternation of prepositions. We do not introduce a new form in order to encode differences in animacy as explained in Columns 1 and 2 above.

1. Diminutives

- (16) a. Θα έχει βάλει η Ελένη την ουρά της βέβαια.
 Will have put the Helen_i the tail her_i of course.
 ‘Of course Helen has been involved in this’
- b. Θα έχει βάλει η Ελένη την ουρίτσα της βέβαια.
 Will have put the Helen_i the small.tail her_i of course.
 ‘Of course Helen has been involved in this.’

2. Adjectives (small number)

- (17) Μου άναψε όλα τα λαμπάκια.
 me.DAT.GEN turn.on.3RD.SG all.ACC the lights.ACC
 ‘He made me completely furious.’
- (18) Να δούμε ποιος θα βγάλει το φίδι από τη φωλιά του
 to see who.NOM will pull.out.3SG the snake.ACC from the nest.ACC its
 ‘Let’s see who is going to do all the dirty work for us.’

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|---------|---------|----------|------------|-----------|-------------------------------------|
| NP-ἰ | | | | | | <input type="checkbox"/> |
| LEM | βάζω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | η | την | AtDfFeSζ | AtDfFeSgAc | 1 | <input type="checkbox"/> |
| LEM | ουρίτσα | ουρίτσα | NoCmFe | NoCmFeSgAc | 1 | <input type="checkbox"/> |
| PnG | μου | μου | PnGeWe | PnGeWe | 1 | <input checked="" type="checkbox"/> |

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|-------|------|----------|------------|-----------|-------------------------------------|
| NP-ἰ | | | | | | <input type="checkbox"/> |
| LEM | βάζω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | η | την | AtDfFeSζ | AtDfFeSgAc | 1 | <input type="checkbox"/> |
| LEM | ουρά | ουρά | NoCmFe | NoCmFeSgAc | 1 | <input type="checkbox"/> |
| PnG | μου | μου | PnPοGe | PnPοGe | 1 | <input checked="" type="checkbox"/> |

Figure 4.10: Diminutives in a MWE.

- (19) Θέλει να κόψει όλες τις γέφυρες
wants to cut all.ACC the bridges.ACC

‘He wants to be completely independent.’

3. Definite or Indefinite article

- (20) Οι χειρισμοί του έβαλαν τη χώρα στο μάτι ενός κυκλώνα.
the handlings.NOM his put.3PL the country in.the eye a.INDEF typhoon.GEN

‘His handling has put the country in an unfavorable position.’

- (21) Οι “αγορές” έβαλαν στο μάτι του κυκλώνα τη Γαλλία.
the “markets”NOM put.3PL.PAST in_ihe eye the.DEF typhoon.GEN the France

‘The Markets have put France in an unfavorable position.’

| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optional Bound | Bound | Controlled | Operations |
|--------|------|----------|---------|---------|--------------|-----------|--------------------------|--------------------------|------------|------------|
| | NP-I | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | LEM | βάζω | | VbAv | VbAv | | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | LEM | στον | στο | AsPpPa† | AsPpPaNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | LEM | μάτι | μάτι | NoCmNe | NoCmNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | LEM | ένας | ενός | NmCdMε | NmCdMaSgGeAj | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | LEM | κυκλώνας | κυκλώνα | NoCmMε | NoCmMaSgGe | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |
| | NP-, | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | NC | DELETE |

Figure 4.11: The MWE's form with an indefinite article.

| Type | Lemma | WF | Tag | Tag | WWS Index | Optional Bound |
|------|----------|---------|---------|--------------|-----------|--------------------------|
| NP-I | | | | | | <input type="checkbox"/> |
| LEM | βάζω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | στον | στο | AsPpPa† | AsPpPaNeSgAc | 1 | <input type="checkbox"/> |
| LEM | μάτι | μάτι | NoCmNe | NoCmNeSgAc | 1 | <input type="checkbox"/> |
| LEM | ο | του | AtDfNeS | AtDfNeSgGe | 1 | <input type="checkbox"/> |
| LEM | κυκλώνας | κυκλώνα | NoCmMε | NoCmMaSgGe | 1 | <input type="checkbox"/> |
| NP-, | | | | | | <input type="checkbox"/> |

Figure 4.12: The MWE's form with a definite article.

4. Alternative prepositions

- (22) βάζω το μαχαίρι στο / ως / μέχρι το κόκαλο
put the knife in.the / up to / till the bone

‘to implement radical solutions/ to investigate a case without fear of whom the investigation might touch’

- (23) παίρνω μπρος / μπροστά
take ahead / ahead.ADV

‘be activated’

| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optional | Bound | Bound By | Controlled By | Operations |
|---------|----------------------|---------|---------|--------------|--------------|-----------|--------------------------|--------------------------|----------|---------------|-----------------|
| | NP-NOM | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | βάζω | | VbAv | VbAv | | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | ο | το | AtDNeSgAc | AtDNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | μαχαίρι | μαχαίρι | NoCmNeSgAc | NoCmNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | ως | ως | AsPpSp | AsPpSp | 2 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | ο | το | AtDNeSgAc | AtDNeSgAc | 2 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | κόκκαλο | κόκκαλο | NoCmNeSgAc | NoCmNeSgAc | 2 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | | | Add Token |
| | | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | | | Remove the item |
| Comment | <input type="text"/> | | | | | | | | | | |
| Tokens | Type | Lemma | WF | Tag | Tag | WWS Index | Optional | Bound | Bound By | Controlled By | Operations |
| | NP-NOM | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | βάζω | | VbAv | VbAv | | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | ο | το | AtDNeSgAc | AtDNeSgAc | 2 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | μαχαίρι | μαχαίρι | NoCmNeSgAc | NoCmNeSgAc | 2 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | στον | στο | AsPpPaNeSgAc | AsPpPaNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | LEMMA | κόκκαλο | κόκκαλο | NoCmNeSgAc | NoCmNeSgAc | 1 | <input type="checkbox"/> | <input type="checkbox"/> | NO | NO | DELETE |
| | | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | | | Add Token |
| | | | | | | | <input type="checkbox"/> | <input type="checkbox"/> | | | Remove the item |

Figure 4.13: Alternative forms of the MWE βάζω το μαχαίρι ως/στο το κόκκαλο.

5. Alternative form of the same lemma (fig:4.14:

(24) τους αφήνω χρόνους / χρόνια
 them leave years.MASC / years.NEUT

'I am dying'

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|--------|---------|--------|------------|-----------|--------------------------|
| NP-τ | | | | | | <input type="checkbox"/> |
| PnG | μου | | PnGeWe | PnGeWe | | <input type="checkbox"/> |
| LEM | αφήνω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | χρόνος | χρόνους | NoCmMa | NoCmMaPIAc | | <input type="checkbox"/> |

| Type | Lemma | WF | Tag | Tag | WWS Index | OptionalBound |
|------|--------|--------|--------|------------|-----------|--------------------------|
| NP-τ | | | | | | <input type="checkbox"/> |
| PnG | μου | | PnGeWe | PnGeWe | | <input type="checkbox"/> |
| LEM | αφήνω | | VbAv | VbAv | | <input type="checkbox"/> |
| LEM | χρόνος | χρόνια | NoCmNe | NoCmNePIAc | | <input type="checkbox"/> |

Figure 4.14: Example of a lemma of a MWE with two possible forms.

6. Additional constituent that intensifies the meaning of the MWE

(25) ανάβω όλα τα λαμπάκια κάποιου
 turn.on all.ADJ.PL.ACC the lights.PL.ACC someone.GEN

‘to make someone furious’

(26) πίνω το πικρό ποτήρι ως τον πάτο
 drink the bitter.ADJ.SG.ACC glass.ACC to the bottom

‘to experience something sad’

| Type | Lemma | WF | Tag | Tag | WWS Index | Optional Bound | Bound Controlled By | Bound Controlled By | Operations |
|------|--------|--------|----------|------------|-----------|--------------------------|---------------------|---------------------|------------|
| NP-P | | | | | | <input type="checkbox"/> | | NC | DELETE |
| LEM | πίνω | | VbAv | VbAv | | <input type="checkbox"/> | | NC | DELETE |
| LEM | ο | το | AIDfNeSf | AIDfNeSgAc | | <input type="checkbox"/> | | NC | DELETE |
| LEM | πικρός | πικρό | AjBaNeS | AjBaNeSgAc | 1 | <input type="checkbox"/> | | NC | DELETE |
| LEM | ποτήρι | ποτήρι | NoCmNe | NoCmNeSgAc | 1 | <input type="checkbox"/> | | NC | DELETE |
| LEM | ως | ως | AsPpSp | AsPpSp | 2 | <input type="checkbox"/> | | NC | DELETE |
| LEM | ο | τον | AIDfMaSi | AIDfMaSgAc | 2 | <input type="checkbox"/> | | NC | DELETE |
| LEM | πάτος | πάτο | NoCmMa | NoCmMaSgAc | 2 | <input type="checkbox"/> | | NC | DELETE |

Figure 4.15: Example of an additional constituent that intensifies the meaning of the MWE.

4.2 USE/RELATIONS TAB

4.2.1 The USE feature

For each MWE entry, a characteristic example along with the phonetic transcription, PAROLE transliteration and English translation is provided at the USE tab.

A tabular representation of a glossed example that has been verbatim retrieved from corpora or the WEB is given (figure 4.16).

| Use | Translation | | | | |
|---------|--|---------|--------------------|--|--------|
| | Russians involved Turkey | | | | |
| Source | http://m.capital.gr/News.aspx?id=1756993 | | | | |
| Tokens | | | | | |
| Or | i | the | AIDfMaPfm | | DELETE |
| Púdoi | Rosi | Russian | NoPfmPfm | | DELETE |
| éðakon | evalan | put | VbMidPa3SPOuPeAu0x | | DELETE |
| sto | sto | in the | AsPpPaNeSgAc | | DELETE |
| paixidi | paixidi | game | NoCmNeSgAc | | DELETE |
| tin | tin | the | AIDfFesGAc | | DELETE |
| Touxiá | Toukia | Turkey | NoPfmFesGAc | | DELETE |

Add Word
✖ Remove the item

Figure 4.16: The tabular organization of the USE tab.

Caution:

Examples extracted from the corpus should be spelt and punctuated correctly e.g. the MWE *πέρνω για έννα κομμάτι ψομί* should to be rewritten as *παίρνω για ένα κομμάτι ψωμί* (figure 4.17).

| Word | Phonetic | Prefix | Gloss | Parole Tag | Suffix |
|---------|----------|--------|---------|-----------------------|--------|
| Έβαλε | evale | | put | VbMnidPat03SgXxPeAvXx | |
| το | to | | the | AtDfNeSgAc | |
| χέρια | xjeraki | | hand | NoCmNeSgAc | |
| του | tu | | his | PnPc03MaSgGeWe | |
| για | gia | | for | PtOt | |
| να | na | | to | PtSj | |
| περάσει | perasi | | promote | VbMnidXx03SgXxPeAvXx | |
| ο | o | | the | AtDfMaSgNm | |
| ΟΣΦΠ. | OSFP | | OSFP | NoPrMaSgNm | |

Figure 4.17: An example extracted from the corpus.

The sentences chosen to be analyzed and translated should have a complete meaning by themselves, if possible. This, as a consequence, will facilitate both the understanding of the source and the translated sentences. For example the sentence οι σύνδεσμοι έβαλαν μπροστά τον Ζαγοράκη does not have a clear meaning in the source language and thus its translation is problematic.

The following table was used for the Phonetic Transcription of the MWEs, which are encoded in the second column of the USE tab (“Phonetic”):

| | | | |
|---|---|-------------|-------------|
| α | a | υ | i |
| β | v | φ | f |
| γ | γ | χ | χ |
| δ | δ | ψ | ps |
| ε | e | ω | o |
| ζ | z | μπ | b |
| η | i | ντ, γκ | d,g |
| θ | θ | ει, οι, ηι | i |
| ι | i | αι | e |
| κ | k | ου | u |
| λ | l | ηυ, ... | iv,if |
| μ | m | κι,γι,χι | ki,gi,xi |
| ν | n | κοι,γοι,χοι | ki,gi,xi |
| ξ | x | κυ,γυ,χυ | ki,gi,xi |
| ο | o | κη,γη,χη | ki,gi,xi |
| π | p | κει,γει,χει | ki,gi,xi |
| ρ | r | κε,γε,χε | kie,yie,xie |
| σ | s | και,γαι,χαι | kie,yie,xie |
| τ | t | | |

Table 4.4: Phonetic transcription of Greek into English.

- The MWE encoded in the USE tab will be glossed according to the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>). Glosses are given on the fourth column of the USE tab (“Gloss”).

- Also, the MWE encoded in the USE tab will be tagged with the PAROLE tagset (http://nlp.ilsp.gr/nlp/tagset_examples/tagset_en/). Morphological tags are given on the fifth column of the USE tab (“Parole Tag”), Where the Select button allows an easy and effective definition of every word (figure 4.18):

| Parole Tag | SuffixOperations |
|-------------|------------------|
| AtDfNePINm | DELETE |
| | DELETE |
| | DELETE |
| Noun | DELETE |
| NoCmMaSgNm | DELETE |
| NoCmMaSgGe | DELETE |
| NoCmMaSgAc | DELETE |
| NoCmMaSgAc | DELETE |
| AtDfNeSgGe | DELETE |
| AjBaNeSgGe | DELETE |
| AjBaNeSgGe | DELETE |
| NoCmNeSgGe | DELETE |

✘ Remove the item

Figure 4.18: The tag button.

4.2.2 The RELATIONS feature

In the RELATIONS tab the semantic relations among MWEs are documented. The following relations can be encoded:

- Synonymous MWEs (figure 4.19):

| | |
|----------|--------------------|
| Relation | βάζω το χεράκι μου |
| Type | Synonym |

✘ Remove the item

Figure 4.19: Synonymous MWE with the MWE βάζω την ουρά μου.

- Semantic pair (MWEs with common nominal phrase but with different verb normally

standing in a causative-non causative relation) (figure 4.20). To this set typically belong pairs of the type βάζω-μπαίνω, αφήνω-μένω, είμαι-γίνομαι, etc.

Relation: μένει στη μέση κάτι

Type: Semantic pair

Remove the item

Figure 4.20: Semantic pair with the MWE αφήνω μισό κάτι.

•Opposite MWEs (figure 4.21):

Relation: κατεβάζω από τα ουράνια κάποιον

Type: Opposite

Remove the item

Figure 4.21: Encoding an opposite MWE in the entry of the MWE ανεβάζω στα ουράνια κάποιον.

•Possessive alternants: Pairs of MWEs where one of the members denotes that an entity has the property denoted by the other member MWE and has the verb έχω as a head verb (βάζω κάποιον στο βρακί μου, έχω κάποιον στο βρακί μου) (figure 4.22)

Relation: έχω έναν άσο στο μανίκι μου

Type: Possessive

Remove the item

Figure 4.22: Encoding an opposite MWE in the entry of the MWE βάζω κάποιον στο βρακί μου.

•Stative alternants: Pairs of MWEs where one of the members denotes that an entity is at the state denoted by the other member MWE and has the verb είμαι, βρίσκομαι as a head verb (βάζω κάποιον στο βρακί μου, είμαι στο βρακί κάποιας) (figure 4.23).

Relation: είμαι στο μάτι του κυκλώνα

Type: Stative

Remove the item

Figure 4.23: Encoding a stative alternant to the MWE βάζω στο μάτι του κυκλώνα κάποιον / κάτι

4.3 CORPUS/DIAGNOSTICS TAB

4.3.1 The CORPUS feature

Corpus data are introduced by creating new slots with the key `ADD`. Corpus data are accessible from the `DIAGNOSTICS` tab.

The text corpora accommodated in `IDION` contain both grammatical and ungrammatical strings featuring MWEs. Strings are marked for their grammaticality by selecting a label from the appropriate pull down menu at the far right part of the window. The grammatical strings are drawn from the Hellenic National Corpus¹ and from the `WEB` while the ungrammatical ones have been evaluated by native speakers (introspection). A string which can not be characterized either as grammatical or as ungrammatical can be marked as questionable.

Ungrammatical strings are also documented to be used as counterexamples in the `Diagnostics Tab`, when the MWEs value in a question is set to “No”, while questionable strings are used as examples when the value in one of the diagnostics is unclear and it is set to “Unknown”.

Examples of literal use of the MWE have also been added when the MWE allows for them, in order to illustrate the difference between the literal and the idiomatic meaning of MWEs. (Figure 4.24).

The screenshot displays three entries in the CORPUS tab, each with a 'Remove the Item' link at the top right. Each entry consists of the following fields:

- Example:** A text input field containing a Greek sentence.
- Acceptance:** A dropdown menu with a downward arrow.
- URL:** A text input field containing a URL.
- Literal:** A checkbox.

Entry 1: Example: "Το λουριό οριζήθηκαν λάσκα", Acceptance: "ungrammatical", URL: "Introspection", Literal: .

Entry 2: Example: "Άφηνε λίγο λάσκα το λουρι της χυρίς όμως να χάνεις τον έλεγχο γιατί είναι και δυνατό σκυλι και θα σε δυσκολεύει", Acceptance: "grammatical", URL: "https://www.dogforum.gr/community/threads/%CF%80%CE%B5%CF%83%CE%BC%CE%B1%CF%84%CE%AC%CE%BA%CE%B9%CE", Literal: .

Entry 3: Example: "Από την άλλη, υπάρχουν περιπτώσεις που αφήνει λγάκι λάσκα το λουρι στους δημιουργούς.", Acceptance: "grammatical", URL: "http://www.ragequit.gr/reviews/item/child-of-light-pc-review", Literal: .

Figure 4.24: The CORPUS tab.

The purpose of the `CORPUS` tab is to provide a concise collection of verbatim usages and of evaluated negative data that exemplify the analysis provided in the different tabs. These data may be of help in future implementations.

¹<http://hmc.ilsp.gr/>

Aiming at the homogeneity of the number of the examples in the corpus per entry, at least 10 grammatical and ungrammatical examples are provided, when possible.

Caution:

We make sure that the corpus examples feature the same meaning of the expression. For instance, if we have the expression βάζω μέσα κάποιον that means both ‘imprison somebody’ and ‘create debts for somebody’:

- we develop two different entries, one for each meaning of the expression
- we make sure that the corpus examples in each lexical entry exemplify the right meaning

4.3.2 The DIAGNOSTICS feature

With the Diagnostics tab, we investigate:

1. whether a verb MWE has a free subject or not (dedicated diagnostic)
2. the number of constituents that a MWE contains. For instance, the MWE below contains three constituents (example 19)

(27) [κάνω] [μαύρα μάτια] [να δω NP.ACC]
 [do] [black eyes] [to see NP.ACC]

The diagnostics *admission of a free XP* and *word order permutations* are used as constituency diagnostics that show whether:

3. the MWE alternates between a form with a free NP.GEN and a form with a Dative Genitive
4. the fixed parts of the MWE can be replaced with a clitic in the same predication
5. the MWE participates in the causative-inchoative alternation
6. the MWE passivizes

Each question is assigned a yes/no/unknown button. For each question it is possible to retrieve from the corpus examples illustrating the phenomenon in question.

For each question, a detailed description is given below.

- Whether a verb MWE accepts a free subject or not. This is the basic diagnostic for the existence of a free subject. If different NPs trigger agreement on the verb a free subject exists, otherwise the subject is fixed. The situation is shown with corpus examples that demonstrate subject-verb agreement with a variety of subjects (figure 4.25):

| | | |
|-------------|----------|--|
| Diagnostics | entry | Can different NPs trigger agreement on the verb? |
| | Choice | yes |
| | Examples | <ul style="list-style-type: none"> ✓ ο Σάκης Ρουβάς βγάζει έναν ακόμα άσο από το μανίκι του, ξεκινώντας τη νέα του περιοδεία ✗ Remove the item ✓ Η πεθερά μου τελικά κάθε φορά βγάζει και έναν καινούριο άσο από το μανίκι της ✗ Remove the item |

Figure 4.25: The Free/fixed subject diagnostic.

- Whether the MWE accepts a free XP constituent. The position of the free constituent delineates the constituents of the MWE.

In the examples below (28a, 28b, 28c), the free XPs *μόνος του* (28a) –an ADJP and *στην κυριολεξία* (28b) –an ADVP, indicate that the fixed string is ‘το φίδι από την τρύπα’.

- (28) a. Προτιμά να βγάζει *μόνος του* το φίδι από την τρύπα.
 prefers to pull.out alone his the snake from the nest
 ‘He prefers to do the dirty work all alone.’
- b. Όλοι αυτοί άνθρωποι βγάζουν στην κυριολεξία το φίδι από τη
 all these people pull.out literally the snake from the
 τρύπα.
 nest
 ‘All these people are literally doing all the dirty work.’
- c. *Προτιμά να βγάζει το φίδι *μόνος του/στην κυριολεξία* από τη
 prefers to pull.out the snake alone his/literally from the
 τρύπα.
 nest

- Word order permutations are exemplified by retrieving examples from the corpus. Ungrammatical word orders are also exemplified with corpus examples that have been supplied by the encoders.

| | | |
|--|----------|--|
| | entry | Word order permutations involving the subject as well |
| | Choice | yes |
| | Examples | <ul style="list-style-type: none"> ✓ Θα παραδεχόταν πως πήγε αυτόφωρο επειδή είπε ψέματα δημόσια ε; Τρώμε κουτόχορτο νομίζεις. ✗ Remove the item ✓ Κουτόχορτο έφαγε μπόλοκο το φόρουμ, ξυπνάτε τώρα. ✗ Remove the item |

[Add another item](#)

Figure 4.26: Example of a MWE’s word order permutations.

- If the MWE is fully fixed and doesn't allow for word order permutations, it is possible to provide an ungrammatical example to show its limits in word ordering (figure 4.27).



Figure 4.27: An unacceptable word order.

As shown above (figure 4.27), the word order τη γλυτώνω is strictly fixed and, therefore, a word ordering of the form “γλυτώνω τη” is unacceptable, since the clitic must be always placed before the verb.

- Cliticisation of the WWS

Cliticisation of the WWS is investigated in the context of the same MWE:

The value YES is assigned if cliticisation is possible in the context of the same MWE (21).

- (29) Έβαλε την ουρά του κι ο Τσιώλης!
 Put.3SG.PAST the tail.ACC hisPN.GE and the Tsiolis.GEN
- (30) Ναι, την έβαλε.
 yes, it.ACC put3SG.PAST

The value NO is assigned if cliticisation is not possible in the context of the same MWE.

- (31) Έταξε στην Ελένη λαγούς με πετραχήλια.
 promised3SG.PAST in Helen.ACC haresPL.ACC with stoles.PL.ACC
- (32) *Ναι, τους έταξε.
 *Yes, them promised.3SG.PAST

The value no is also assigned when the cliticisation diagnostic is irrelevant, eg. when the WWS is a PP.

The value unknown is assigned when the editor cannot decide whether the cliticisation criterion of the examined MWE applies or not.

- Alternation of a free NP-GEN/ free σε-PP/ free από-PP with Dative Genitive

The value YES is assigned if the alternation with a Dative Genitive is possible.

- (33) a. Ο Γιώργος έκοψε τα φτερά της Ελένης.
 the George1SG.NOM cut.3SG.PAST the wings.PL.ACC the helen.GE
 “George discouraged Helen.”

- b. Της έκοψε τα φτερά ο Γιώργος.
her..DATGEN cut.3SG.PAST the wings.PL.ACC the George.NOM.

The value NO is assigned if the alternation with a Dative Genitive is not possible.

- (34) a. Τρώω τα νύχια μου.
eat.1SG.PRES the nails.PL.ACC my
“I’m anxious”
b. *Μου τρώω τα νύχια.
me.DATGEN eat.1SG.PRES the nails.PL.ACC.
“I’m bored to death”

It is left to future decision whether the dative genitive form of an alternating expression will be encoded explicitly or it will be derived with some rule from the without-the-dative-genitive encoded version. In fact, once we know whether it alternates or not and the structure of the MWE (both are tagged in the DB), we can apply one of the two rules of dative genitive alternation that can be defined—NP-NOM and Y are optional: 1. If the MWE alternates and is of the form [NP-NOM, X, σε κπ], it receives the form [NP-NOM, μου, X] 2. If the MWE alternates and is of the form [NP-NOM, X,NP-GEN,Y], it receives the form [NP-NOM, μου, X,Y]

● Causative-inchoative alternation The value YES is assigned if the alternation is possible.

- (35) a. ανάβω τα λαμπάκια σε κάποιον
turn on.1SG the lights.ACC to someone
“To make someone furious”
b. μου ανάβουν τα λαμπάκια
me.DATGEN turn on.1PL the lights.NOM
“Something makes me furious”

The value NO is assigned if the alternation is NOT possible:

- (36) a. Κόβω τις φλέβες μου
cut.1SG.PRES the veins.ACC my
“I’m bored to death”
b. *Οι φλέβες μου κόβουν.
the veins.NOM.SUBJ my cut.3PL.PRES
“I’m bored to death”

● Passivisation. Whether the verb of the MWE passivises or not (figure 4.28):

| | |
|----------|---|
| entry | Passivisation |
| Choice | yes |
| Examples | <ul style="list-style-type: none"> ✓ Η πολωνική ομάδα αφέθηκε στα κρύα του λουτρού από τον έμπειρο γκολκίπερ, αφού είχε διώξει όλο |

Figure 4.28: Passivisation of a MWE.

The value YES is assigned if the alternation is possible:

- (37) a. *πάρθηκαν μέτρα*
 take.3PL.PAST.PASS measures.3PL.NOM
 “Measures were taken”

The value NO is assigned if the alternation is NOT possible:

- (38) a. **αφέθηκαν παιδιά πίσω του*
 let.3PL.PAST.PASS children.ACC behind him.PN.GE
 “The deceased had children”

Chapter 5

Grundlagen

5.1 1. Thema

5.1.1 1. Unterthema

- (39) This is an example.
- (40) This is the first example.
- (41) This is the second example.
- (42) This is the third με λίγα ελληνικά.

- (43)
 - a. This is a sub-example.
 - b. This is a second sub-example.
 - c.
 - i. This is a sub-sub-example.
 - ii. This is a second sub-sub-example.

Chapter 6

Fazit

Bibliography

Anhang