

**LEMLAT**  
**Wordforms analysis**  
**Database description**

**Glossary**

- o LES: the invariable part of the inflected forms;
- o SM (Segmento Mediano): the middle part of the inflected forms;
- o SF (Segmento Finale): the final part of the inflected forms;
- o SI (Segmento Iniziale): the initial part of the inflected forms;
- o SPF (Segmento Post Finale): a segment added on the right side of the final part of a wordform;
- o COD LES: it is the code assigned to each LES; each COD LES refers to a particular type of inflexion;
- o COD LEM: it is the code assigned to each output lemma; each COD LEM refers to a general type of inflexion;
- o FE (Forma Eccezionale): exceptional wordform. A wordform inflected in an exceptional way that cannot be regularly segmented and recognised;
- o LE (Lemma Eccezionale): exceptional lemma. A lemma created in an exceptional way that cannot be automatically created;
- o CLEM (Costellazione LEMmatica): contains all the LES related to a common lemma, or common dictionary entry; it is referred to through a unique N\_ID
- o Ipolemma: intermediate lemma produced in output, not referring to a dictionary entry;
- o Iperlemma: lemma produced in output referring to a dictionary entry
- o N\_ID: alphanumeric code applied to all the LES. More LES can share the same N\_ID: all the LES related to a common lemma, or common dictionary entry are registered with the same N\_ID (forming a CLEM)
- o CodLE: numeric code of LE, related to pattern(s) of 7 EAGLES codes bringing morphological information about the wordforms

- o EAGLES (Expert Advisory Group on Language Engineering Standards): standard coding of morphological, morpho-syntactic and semantic information of the words. In LEMLAT, 3 EAGLES codes are related to lemmas, and 7 to wordforms

## Analysis of a wordform

Receiving in input a wordform, if it is suitable to be analysed, LEMLAT produces in output:

- the corresponding lemma(s);
- a code expressing the inflexional paradigm of the lemma(s) (codlem)
- the n\_id of the lemma(s) CLEM (see table "lessario")
- 3 EAGLES codes (converted by codlem) related to the lemma (one pattern of 3 EAGLES codes for each lemma produced in output), with information about (see "cod\_morf" table):
  - o P(osition)1: PoS
  - o P2: Type (different possible types of each PoS; for instance, a noun can have Type "common", or "proper")
  - o P3: Flexional Category (declension, conjugation,...)
- pattern(s) of 7 EAGLES codes related to the wordform, with information about (see "cod\_morf" table):
  - o P4: mood
  - o P5: tense
  - o P6: case
  - o P7: gender
  - o P8: number
  - o P9: person
  - o P10: degree

This analysis is obtained through a process of segmentation/recognition of input wordforms.

For each input wordform, LEMLAT operates a number of segmentation attempts.

When one of these attempts is found consistent with LEMLAT data about wordforms possible segments, the analysis of the wordform is performed in output.

There are three possible segmentation structures:

1. LES + SF
2. LES + SM + SF
3. LES + SM + SM +SF

Each of these structures can be preceeded by a SI and followed by a SPF.

In addition to segmentation process, a wordform can be also recognised (and, thus, analysed) with no segmentation, in the following cases:

- Input wordform is a FE
- Input wordform is a LE
- Input wordform is a les with codles "i" (invariables)
- Input wordform is a les with codles "n" (uninflected nouns)

- Input wordform is a les with codles "v" (verbs not related to a specific conjugation)
- Input wordform is a les with codles "pr", or "p1-p9", or "p18" (not segmented pronominals)

Also each of these structures can be preceded by a SI and followed by a SPF.

A segmentation is valid if its segments are found as each other compatible (on left and/or on right side). The compatibility of the segments is coded along with the segments itself (see "lessario", "tabsf", "tabsm", "tabsi", "tabspf" tables).

For instance, a structure such as

LES + SM + SF

is found valid if:

- left compatibility of SM corresponds to codles (that is, with right compatibility of LES)
- right compatibility of SM corresponds with left compatibility of SF

In order to produce output information:

A)

If the input wordform is segmented:

- lemma and codlem (3 EAGLES lemma codes): produced according to codles (see "eagles" table and annex -2-)
- pattern(s) of 7 EAGLES wordform codes: from SF (and SM) coding (see "tabsf" and "tabsm" tables)

B)

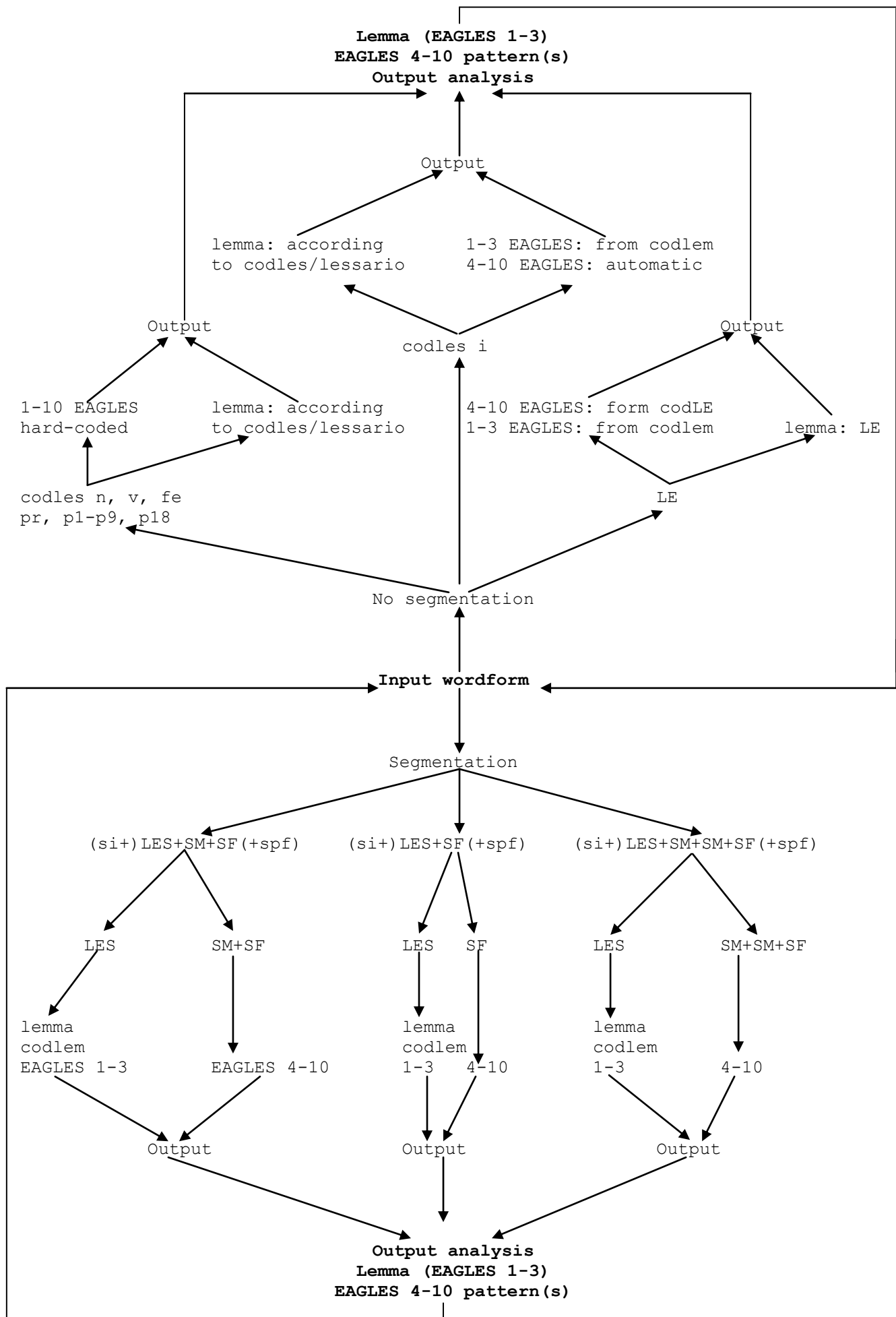
If the input wordform is not segmented:

- in case of LE:
  - o codlem (3 EAGLES lemma codes): according to codles (see "eagles" table and annex -2-)
  - o pattern(s) of 7 EAGLES wordform codes: from codLE (each LE is related to a codLE, that brings the seven EAGLES codes pattern(s) of the wordform; see "cod\_le" and "tabl\_le" tables)
  - o lemma: LE itself (possibly, reduced to an iperlemma)
- in case of LES with codles "i":
  - o patterns of 10 EAGLES codes (3 lemma codes + 7 wordform codes): 1-3 converted from codlem (see "eagles" table); 7-10 automatically assigned as "-----"
  - o lemma: produced according to codles (see annex -2-), or to information related to concerned les on table "lessario"
- in case of les with codles "FE", "n", "v", "pr", "p1-p9", or "p18":
  - o pattern(s) of 10 EAGLES codes (3 lemma codes + 7 wordform codes): from hard-coding of each les with codles "FE", "n", "v", "pr", "p1-p9", or "p18" (see "forme\_ecc" table)
  - o lemma: produced according to codles (see annex -2-), or to information related to concerned les on table "lessario"

Each segmentation can produce analysis related to more than one lemma.

When a segmentation is found valid and the analysis is performed, LEMLAT does not stop the process, but produces other segmentation/recognition attempts: a wordform can be segmented (and analysed) in more than one way. Equally, the same wordform can be analysed through segmentation and through no-segmentation (see the case of a wordform showing homography between a regular - segmented - one and, for instance, a FE - not segmented -).

The analysis of a wordform performed by LEMLAT can be summarised according to the following schema:



## **Database Tables**

- o lessario
- o cod\_le
- o cod\_morf
- o eagles
- o forme\_ecc
- o teb\_le
- o tabsf
- o tabsm
- o tabspf
- o tabsai
- o tabsi
- o graph\_vars
- o lemma\_ending
- o lemmario

### **lessario**

List of the les.

- n\_id
  - o clem identification number
  - o values:
    - letter (first letter of the lemma)
    - four numbers
- gen
  - o gender
  - o values:
    - see "cod\_morf" table, field "field\_pos", value "7"
    - \*: means "any possible value"
- clem
  - o in a clem containing more than 1 les, identifies the les through which the lemma has to be created
  - o values:
    - v: identifies the les through which the lemma has to be created
    - i: for superlative and comparative forms of irregular participle, irregular gerundive, and irregular forms (graphical variants) of I/II class adjectives, the second lemma created (participle, gerundive, or adjective at positive degree) is an ipo- and not an iperlemma
    - k: stops the creation of the iperlemma (value "v" is inhibited)
- si (Segmento Iniziale)
  - o initial alteration h
  - o value:
    - h: the les appears also with an initial h
- smv (Segmento Mediano Verbale)
  - o automatic insertion/exclusion of smv
  - o values:

- +: adds a smv to the les, to automatically create the regular basis for perfect and future participle, and perfectum
    - -: adds a smv to the les, to automatically create the regular basis for comparative, superlative, present participle, gerund and gerundive
    - blank: no smv to be added (irregular inflections)
  - spf (Segmento PostFinale)
    - o adds/cuts a spf to les
    - o values:
      - 3: exclusion of -que (enclitic)
      - see "tabspf" table, field "comp\_cod"
  - les
  - codles
    - o values: see annex -1-; see table "eagles", field "codles"
  - lem
    - o LE:
      - a complete form

NOTE: in case of homography between two, or more lemmas, if the only difference among them is the length of a vowel, this is recorded in LE as follows:

      - one quote (') after the involved vowel: the vowel is short
      - two quotes (') after the involved vowel: the vowel is long
    - or
    - a SF to be added to les
    - or
    - =: the lemma is identical to the les

if more than one LE is concerned, the LE are divided by a slash

  - o if no LE is recorded, the lemma is created through through automatically adding a SF to the les, rule depending on codles.; see annex -2-
- s\_omo
  - o omographic lemma
  - o values:
    - A: omographic lemma A
    - B: omographic lemma B
- più
  - o more les in the same clem, but none with "v" in clem field
  - o values:
    - +
- codlem
  - o manually recorded if cannot be automatically assigned according to codles
  - o see annex -3-; see table "eagles", field "codlem" for the correspondance codles/codlem
- type
  - o manually recording of Type
- codLE



- o in case of LE, exclusion of the 7-10 position codes in output patterns
  - o values: see "cod\_le" table
- pt
  - o pluralia tantum
  - o values:
    - x: exclusion of patterns with code "s" in position 8
- a\_gra
  - o graphic alteration
  - o values: see "tabsai" table
- gra\_u
  - o les possibly divided in two parts
  - o values
    - x
- notes
- pr\_key
  - o identification number of the les
- ts
  - o Time Stamp: last time when the line has been modified
- src
  - o source of the lexical entry
  - o values
    - B: base (original lexical basis of LEMLAT)
    - O: onomasticon (Forcellini's onomasticon)

### **cod\_le**

List of codes and values for LE analysis.

- cod\_LE
  - o codLE: in the analysis of an LE, adds the codes from c04 to c10. See "cod\_morf" table for codes values
- c04
  - o codes in position 4
- c05
  - o codes in position 5
- c06
  - o codes in position 6
- c07
  - o codes in position 7
- c08
  - o codes in position 8
- c09
  - o codes in position 9
- c10
  - o codes in position 10
- pr\_key
  - o identification number of the codLE
- ts
  - o Time Stamp: last time when the line has been modified

### **cod\_morf**

Description of codes/values/attributes occurring in the 10 positions output patterns.

- field\_pos
  - o position in the pattern
  - o values: 1-10
- field\_descr
  - o description of the field value
- value\_descr
  - o description of the attribute for each field
- value
  - o description of the code for each attribute/field
- ts
  - o Time Stamp: last time when the line has been modified

### **eagles**

Conversion codles/codlem/1-3 position codes (lemma codes)

- codles
  - o codles list
- codlem
  - o codlem corresponding to codles recorded on the same line
- c01
  - o codes in position 1
- c02
  - o codes in position 2
- c03
  - o codes in position 3

### **forme\_ecc**

Hard-Coding of exceptional wordforms pattern(s).

- les\_id
  - o link to corresponding line in lessario table (pr\_key field)
- add\_lem
  - o link to a second lemma through pr\_key field in lessario table
- enc
  - o presence of an enclitic
- c01
  - o codes in position 1
- c02
  - o codes in position 2
- c03

- o codes in position 3
- c04
  - o codes in position 4
- c05
  - o codes in position 5
- c06
  - o codes in position 6
- c07
  - o codes in position 7
- c08
  - o codes in position 8
- c09
  - o codes in position 9
- c10
  - o codes in position 10
- pr\_key
  - o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **tab\_le**

List of LE recorded along with its own codLE

- lemma
  - o list of LE
- codLE
  - o codLE
  - o Value: see "cod\_le" table, field "cod\_LE"
- les\_id
  - o link to corresponding line in lessario table (pr\_key field)
- pr\_key
  - o identification number of the line

### **tabsf**

List of SF and related codes patterns.

- segment
  - o SF
- comp\_cod
  - o codles compatible on SF left side
- c01
  - o codes in position 1
- c02
  - o codes in position 2
- c03
  - o codes in position 3
- c04

- o codes in position 4
- c05
  - o codes in position 5
- c06
  - o codes in position 6
- c07
  - o codes in position 7
- c08
  - o codes in position 8
- c09
  - o codes in position 9
- c10
  - o codes in position 10
- ex
  - o example
- pr\_key
  - o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **tabsm**

List of SM and related codes patterns.

- segment
  - o SM
- pm
  - o +: if "+" is recorded in field "smv" (table "lessario"), automatically creates ipolemmas of perfectum, supine, future and perfect participle
  - o -: no ipolemma is created through the SM occurring in field "segment"
- comp\_cod\_prec
  - o codles compatible on SM left side
- comp\_cod\_succ
  - o codles compatible on SM right side
- c01
  - o codes in position 1
- c02
  - o codes in position 2
- c03
  - o codes in position 3
- c04
  - o codes in position 4
- c05
  - o codes in position 5
- c06
  - o codes in position 6
- c07
  - o codes in position 7

- c08
  - o codes in position 8
- c09
  - o codes in position 9
- c10
  - o codes in position 10
- ex
  - o example
- pr\_key
  - o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **tabspf**

SPF compatibility.

- segment
  - o SPF
- comp\_cod
  - o compatibility on SF left side
- pr\_key
  - o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **tabsai**

Initial graphic alteration. Related to "a\_gra" field in lessario.

- segment
  - o Initial alteredated segment
- comp\_cod
  - o compatibility code
- pr\_key
  - o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **tabsi**

Initial segment. Related to "si" field in lessario.

- segment
  - o Initial segment
- comp\_cod
  - o compatibility code
- pr\_key

- o identification number of the line
- ts
  - o Time Stamp: last time when the line has been modified

### **graph\_vars**

Graphical variation in the les.

- gv\_code
  - o Code of graphical variation; recorded in field "a\_gra" of "lessario" table
- gv\_pos
  - o Ordinal number of occurrence position (in the les) of the letter after/before which the variation is applied
  - o Numeric code:
    - 1: the variation is applied after/before the first occurrence of the letter concerned by the variation (this letter is recorded in "gv\_out" field)
    - 2: the variation is applied after/before the second occurrence of the letter concerned by the variation (this letter is recorded in "gv\_out" field)
    - ...
- gv\_in
  - o graphical form appearing in input wordform
- gv\_out
  - o output graphical variation (letter concerned by graphical variation)
  - o graphical variation is applied to input wordform to retrieve the involved les in table "lessario" (the les is recorded with no graphical variation)
- ts
  - o Time Stamp: last time when the line has been modified

### **lemma\_ending**

Automatic creation of lemma according to COD LES.

- codles
  - o COD LES of the last segment
- in\_ending
  - o matching condition on the end of the wordform, without the last segment. Blank: no condition
- out\_ending
  - o ending of the lemma (to be concatenated to the wordform without the last segment)
- ts
  - o Time Stamp: last time when the line has been modified

### **lemmario**

List of lemmas of the lexical basis of LEMLAT.

It includes both the lemmas from the original lexical basis of LEMLAT and those from the Onomasticon of Forcellini.

- id\_lemma
  - o unique identification number of the line (i.e. lemma)
- lemma
  - o name of the lemma
- codlem
  - o codlem
- gen
  - o gender
- codmorf
  - o EAGLES morphological tags for the lemma (positions 1-2-3)
- n\_id
  - o n\_id of the lemma in "lessario" table
- lemma\_reduced
  - o verbs with more than one endings reduced to one (for practical matters of derivational\_db)
    - example: punio/-ior -> punio
- src
  - o source
    - B: base (original lexical basis of LEMLAT)
    - O: onomasticon (Forcellini's onomasticon)
    - F: fictional lemmay (for wordformation purposes)

### **lemmas\_wfr**

Wordformation-based relations between input and output lemmas.

- wfr\_key
  - o unique identification number of the wordformation rule
- o\_id\_lemma
  - o id of the output lemma (see id\_lemma in table "lemmario")
- i\_id\_lemma
  - o id of the input lemma (see id\_lemma in table "lemmario")
- i\_ord
  - o order of the input lemma in the output (used in case of compounding). Default value: 1
- category
  - o category of the wordformation rule in terms of input PoS and output PoS. Format: PoSinput-To-PoSoutput (for derivation); PoS1+PoS2=PoSoutput (for compounding)
- type
  - o type of the wordformation rule: Derivation (Suffix, Prefix, Conversion); Compounding
- affix
  - o name of the affix involved in the wordformation rule

## New codings

### Adding a new les

To add a new les into an already existing clem:

- table "lessario": identify the clem into which the new les has to be added
- add a new line in the table
- write clem n\_id in "n\_id" field
- write the new les in "les" field
- compile the "codles" field

Compiling these three field is obligatory; the others should be compiled according to the inflection of the les to be added.

Remind:

A)

- if the new les has been added into a clem previously formed by only one les

and

- if the wordforms formed with the newly added les should be lemmatized under the lemma created with the previously registered les,
- thus, on the line of the previously registered les add the code "v" into the field "clem"

B)

- if the codles of the new les is "FE", or "v", or "n", or "pr", or "p1-p9", "p18"
- write the code pattern(s) related to the analysis of the new les in "forme\_ecc" table, linking the two tables ("lessario" and "forme\_ecc") pasting in field "les\_id" (in "forme\_ecc" table) the numeric value appearing in the field "pr\_key" of the new les line in "lessario".

C)

- if the new les should be lemmatised under a new LE (recorded in field "lem")
- this LE has to be recorded in "tab\_le" table along with its own codLE

### Adding a new clem

To add a new clem:

- table "lessario": identify a clem n\_id still available
- add a new line in the table
- write clem n\_id in "n\_id" field
- write the new les in "les" field
- compile the "codles" field and other necessary fields

### Adding a new SF



To add a new SF along with its compatibility code(s) and EAGLES codes pattern(s):

- in "tabsf" table: add a new line
- in "segment" field, write the new SF
- in "comp\_cod" field, write codles compatible on SF left side
- in "c01-c10" fields, write the code pattern(s) related to new SF. Note: only the codes recorded in "c04-c10" are active: in output analysis, the first 3 codes are, in fact, converted from codlem. The fields "c01", "c02", "c03" are equally reported for a more comfortable recording and data view: otherwise, they can be hidden.

#### Adding a new SM

To add a new SM along with its compatibility code(s) and EAGLES codes pattern(s):

- in "tabsm" table: add a new line
- in "segment" field, write the new SM
- in "comp\_cod\_succ" and "comp\_cod\_prec" fields, write codles compatible on SM left and right sides
- in "c01-c10" fields, write the code pattern(s) related to new SM. Note: only the codes recorded in "c04-c10" are active: in output analysis, the first 3 codes are, in fact, converted from codlem. The fields "c01", "c02", "c03" are equally reported for a more comfortable recording and data view: otherwise, they can be hidden.

In SM coding the code "=" means that in the final analysis of the input wordform, the code to appear in this position of the pattern is the coding appearing in the same position of the pattern in the coding of the SF occurring in that wordform (on the right side of SM).

#### Adding a new SI

To add a new SI along with its compatibility code:

- in "tabsi" table: add a new line
- in "segment" field, write the new SI
- in "comp\_cod" field, write compatible code appearing in "si" field of "lessario" table

#### Adding a new SPF

To add a new SPF along with its compatibility code:

- in "tabspf" table: add a new line
- in "segment" field, write the new SPF

- in "comp\_cod" field, write compatible code appearing in "spf" field of "lessario" table

#### Adding a new codLE:

To add a new codLE along with its EAGLES codes pattern(s):

- in "cod\_le" table: add a new line
- in "cod\_le" field, write the new code (pay attention to use an available code, not already used)
- in "c-4-c10" fields, write code pattern(s) related to the new codLE

#### Adding a new morphological code (EAGLES)

To add a new morphological code:

- in "cod\_morf" table: add a new line
- in "field\_pos" field, write the position of the new code in the analysis pattern
- in "field\_descr" field write the value related to the chosen position
- in "value\_descr" field write the attribute related to the new morphological code
- in "value" field write the new code. Pay attention that the code is not already used

#### Adding a new codles

To add a new codles along with related codlem and first 3 EAGLES codes:

- in "eagles" table: add a new line
- in "codles" field, write the new codles
- in "codlem" field, write the codlem related to the new codles
- in "c01/c02/c03" write the first three EAGLES codes (lemma codes) corresponding to codlem related to new codles

## Examples

### I declension noun

- lexical entry: abiga, -ae
- inflection: regular; I declension
- gender: feminine
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)  
remember that the first letter of n\_id is the same of the first letter of lemma
- write "f" (feminine) in field "gen"
- write "abig" in field "les"
- write "n1" in field "codles"

### I conjugation verb with some fe

- lexical entry: amo, -are
- inflection: regular; I conjugation;  
presence of the following fe:
  - amarei: present passive infinitive
  - amassint: active perfect congiunctive, plural third person; active past future indicative, plural third person
  - amassis: active perfect congiunctive, singular second person; active past future indicative, singular second person
  - amasso: active past future indicative, singular first person
  - ameminor: passive future imperative, plural second person
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)
- write "v" in field "clem", to use the data on this line to create the iperlemma
- write "+" in field "smv", to automatically create the regular basis for perfect and future participle, and perfectum
- write "am" in field "les"
- write "v1r" in field "codles"
- add another line
- write in field "n\_id" the same n\_id used for the previous line
- write "amarei" in field "les"
- write "fe" in field "codles"
- add another line
- write in field "n\_id" the same n\_id used for the previous line

- write "amassint" in field "les"
- write "fe" in field "codles"
- write "amaui" in field "lem", to create ipolemma "amaui" (basis of perfectum)
- write "vp" in field "codlem", to assign codlem "vp" to ipolemma "amaui"
  
- add another line
- write in field "n\_id" the same n\_id used for the previous line
- write "amassis" in field "les"
- write "fe" in field "codles"
- write "amaui" in field "lem", to create ipolemma "amaui" (basis of perfectum)
- write "vp" in field "codlem", to assign codlem "vp" to ipolemma "amaui"
  
- add another line
- write in field "n\_id" the same n\_id used for the previous line
- write "amasso" in field "les"
- write "fe" in field "codles"
- write "amaui" in field "lem", to create ipolemma "amaui" (basis of perfectum)
- write "vp" in field "codlem", to assign codlem "vp" to ipolemma "amaui"
  
- add another line
- write in field "n\_id" the same n\_id used for the previous line
- write "ameminor" in field "les"
- write "fe" in field "codles"
  
- table "forme\_ecc"
- in field "les\_id", copy/paste the number occurring in field "pr\_key" in table "lessario" on the line of fe "amarei"
- in the fields "c01-c10", write the following codes: VmFh1----- (I declension verb, present passive infinitive). See table "cod\_morf" for details about codes and positions
  
- the same for the other fe

### III conjugation verb with irregular perfect/future participle

- lexical entry: abigo, -ere
- inflection: III conjugation;
  - variant: abago, -ere
  - perfectum basis: abeg
  - perfect participle basis: abact
  
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)

- write "v" in field "clem", to use the data on this line to create the iperlemma
- write "abig" in field "les"
- write "v3r" in field "codles"
  
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "abag" in field "les"
- write "v3r" in field "codles"
  
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "abeg" in field "les"
- write "v7s" in field "codles"
  
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "abact" in field "les"
- write "n41" in field "codles" (for supine)
  
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "i" in field "clem"
- write "abact" in field "les"
- write "n6p1" in field "codles" (for perfect participle)
  
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "i" in field "clem"
- write "abactur" in field "les"
- write "n6p2" in field "codles" (for future participle)

### III declension noun with lemma created through substitution of codles ending

- lexical entry: raucedo, -inis
- inflection: III declension -o, -inis
- gender: feminine
  
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)
- write "f" (feminine) in field "gen"
- write "raucedin" in field "les"
- write "n31" in field "codles"

### III declension noun with initial graphic alteration

- lexical entry: abscessio, -inis
- inflection: III declension -o, -inis

- variant: apscessio, -onis
- gender: feminine
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)
- write "f" (feminine) in field "gen"
- write "abscession" in field "les"
- write "n31" in field "codles"
- write "b02" in field "a\_gra" (for details: see "tabsai" table)

#### I declension noun with a graphical variation

- lexical entry: carruca, -ae
- inflection: I declension
  - variant: carrucha, -ae
  - caruca, -ae
  - carucha, -ae
- gender: feminine
- table: lessario
- add a new line
- write in field "n\_id" a new n\_id (not already used)
- write "f" (feminine) in field "gen"
- write "carruc" in field "les"
- write "n1" in field "codles"
- write "h12" in field "a\_gra" (for details: see "graph\_vars" table)
- add a new line
- write in field "n\_id" the same n\_id used for the previous line
- write "f" (feminine) in field "gen"
- write "caruc" in field "les"
- write "n1" in field "codles"
- write "h12" in field "a\_gra" (for details: see "graph\_vars" table)

Code "h12":

- "h12": variation "ch("gv\_in")/c("gv\_out")"
- "h12": graphical variation is related to the second occurrence ("2") of "c" in the les
  - o "c(1)arruc(2)": graphical variation "ch/c" is related to "c" (2)
  - o input les "carruch" ("ch": "gv\_in") is transformed in "carruc" ("c": "gv\_out") to retrieve the involved les in table "lessario"

## Annexes

### -1- codles

List of available codles along with its codlem and semantic description.

#### Not related to SF

- fe (codlem: iperlemma codlem): exceptional wordforms
- i (codlem: i): invariables
- n (codlem: n): uninflected nouns
- pr; p1-p9; p18 (codlem: pr): pronominals
- v: verbs not related to a specific conjugation
- blk: blank following FE, LE, I, N, V, PR, P1-P9, P18
- blk1: blank following SM "-ns-" in Present Participle forms
- blk2: blank following SM "-ior-" or "-ius-" in comparative forms
- blk3: blank following SM "-ius-" in comparative forms of only neuter irregular gerundive

#### Related to SF

I declension nouns (codlem: n1)

- n1: I declension nouns
- n1e: I declension irregular nouns

II declension nouns (codlem: n2)

- n2: II declension nouns (masculine, and feminine)
- n2e: II declension irregular nouns
- n2i: II declension masculine nouns in -ius
- n2n: II declension neuter nouns
- n2ni: II declension neuter nouns in -ius

Gerund (codlem: n2g)

- n21: gerund

Only neuter perfect participle (codlem: n2np)

- n2np: only neuter perfect participle

III declension nouns (codlem: n3b)

- n3: III declension nouns (masculine, and feminine) with plural genitive in -um/-ium
- n31: III declension nouns (masculine, and feminine) with plural genitive in -um
- n32: III declension nouns (masculine, and feminine) with plural genitive in -ium
- n3e: III declension irregular nouns; singular ablative in -e

- n3n: III declension nouns (neuter) with plural genitive in -um/-ium
- n3n1: III declension nouns (neuter) with plural genitive in -um
- n3n2: III declension nouns (neuter) with plural genitive in -ium
- n3p: III declension nouns with singular ablative in -e/-i

IV declension nouns (codlem: n4)

- n4: IV declension nouns

Supine (codlem: n4s)

- n41: supine

V declension nouns (codlem: n5)

- n5: V declension nouns

I class adjectives; perfect and future participles; gerundives (codlem: n2/1)

- n6: I class adjectives; perfect and future participles; gerundives
- n6i: I class adjectives in -ius
- n6r: I class adjectives in -er
- n6s: I class superlative degree adjectives

I class pronominal adjectives (codlem: pr)

- n6p: I class pronominal adjectives with singular genitive in -ius and singular dative in -i
- n6p3: I class pronominal adjectives inflected as regular first class adjectives

Perfect, future irregular participles; irregular gerundives (codlem: n1/2)

- n6g: irregular gerundives
- n6p1: perfect irregular participles
- n6p2: future irregular participles

Only neuter irregular gerundive (codlem: n2np)

- n6gn: only neuter irregular gerundive
  - o Superlative forms: SM + SF n6n
  - o Comparative forms: SM + SF n7cn

II class adjectives (codlem: n3a)

- n7: II class adjectives with singular nominative masculine and feminine ending in -is, neuter in -e, and singular ablative in -i
- n71: II class adjectives with singular nominative the same for masculine, feminine and neuter (-s; -x; -r; -l), and singular ablative in -e/-i
- n72: II class adjectives with singular nominative feminine ending in -is, masculine in -er, neuter in -e, and singular ablative in -i
- n7c: II class comparative degree adjectives



II class pronominal adjectives (codlem: pr)

- n7p: II class pronominal adjectives

Present irregular participle (codlem: n3p)

- n7p3: present irregular participle

Pronominals (codlem: pr)

- p10-p17; p19-p23: see table "tabsf"

## Verbs

### A) Infectum

Each codles beginning with a v- can have, in fourth position, one of the following letters:

#### Infectum

- v\*\*a: compatibility with present indicative SF
- v\*\*b: compatibility with present conjunctive SF
- v\*\*c: compatibility with future indicative SF
- v\*\*d: compatibility with imperfect indicative SF
- v\*\*e: compatibility with imperfect conjunctive SF
- v\*\*f: compatibility with present imperative SF
- v\*\*g: compatibility with present infinitive SF

#### Perfectum

- v\*\*a: compatibility with active perfect indicative SF
- v\*\*b: compatibility with active perfect conjunctive SF
- v\*\*c: compatibility with active perfect future indicative SF
- v\*\*d: compatibility with active plusperfect indicative SF
- v\*\*e: compatibility with active plusperfect conjunctive SF
- v\*\*g: compatibility with active perfect infinitive SF

I conjugation verbs (codlem: v1)

- v1d: I conjugation deponent verbs
- v1e: I conjugation verbs, impersonal
- v1i: I conjugation verbs, intransitive
- v1r: I conjugation verbs, transitive
- v1s: I conjugation verbs, only active diathesis

II conjugation verbs (codlem: v2)

- v2d: II conjugation deponent verbs
- v2e: II conjugation verbs, impersonal
- v2i: II conjugation verbs, intransitive
- v2r: II conjugation verbs, transitive
- v2s: II conjugation verbs, only active diathesis

III conjugation verbs (codlem: v3)

- v3d: III conjugation deponent verbs
- v3e: III conjugation verbs, impersonal

- v3i: III conjugation verbs, intransitive
- v3r: III conjugation verbs, transitive
- v3s: III conjugation verbs, only active diathesis

#### IV conjugation verbs (codlem: v4)

- v4d: IV conjugation deponent verbs
- v4e: IV conjugation verbs, impersonal
- v4i: IV conjugation verbs, intransitive
- v4r: IV conjugation verbs, transitive
- v4s: IV conjugation verbs, only active diathesis

#### e/i conjugation verbs (codlem: v5)

- v5d: e/i conjugation deponent verbs
- v5e: e/i conjugation verbs, impersonal
- v5i: e/i conjugation verbs, intransitive
- v5r: e/i conjugation verbs, transitive
- v5s: e/i conjugation verbs, only active diathesis

#### Not regular conjugation verbs (codlem: va)

- v6d: not regular conjugation deponent verbs
- v6i: not regular conjugation verbs, intransitive
- v6r: not regular conjugation verbs, transitive
- v6s: not regular conjugation verbs, only active diathesis
- v61a: not regular conjugation verbs; compatibility with present indicative SF
- v62a: not regular conjugation verbs; compatibility with active imperfect conjunctive SF
- v63a: not regular conjugation verbs; compatibility with active future perfect indicative SF
- v64a: not regular conjugation verbs; compatibility with present active conjunctive SF
- v65a: not regular conjugation verbs; compatibility with imperfect active indicative SF
- v66a: not regular conjugation verbs; compatibility with perfect active conjunctive SF
- v67a: not regular conjugation verbs; compatibility with passive future perfect indicative SF
- v68a: not regular conjugation verbs; compatibility with present conjunctive SF
- v69a: not regular conjugation verbs; compatibility with present indicative SF (passive: only SF -tur)

#### B) Perfectum (codlem: vp)

- v7s: perfectum
- v7e: impersonal perfectum
- v8s: syncopated perfectum

**automatic creation of lemma**

If in field "lem" of "lessario" table no LE is recorded, the lemma is created through automatically adding a SF to the les, rule depending on codles.

<u>codles</u>	<u>automatic SF</u>
n1	-a
n1e	-a
n2	-us
n2e	-us
n2i	-ius
n2n	-um
n2ni	-ium
n21	-i
n2np	-um
n3	-is
n31	-is
n32	-is
n3e	-is
n3n	-is
n3n1	-is
n3n2	-is
n3p	-is
n4	-us
n41	-um
n5	-es
n6	-us
n6i	-ius
n6r	-us
n6s	-us
n6p	-us
n6p3	-us
n6g	-us
n6p1	-us
n6p2	-us
n6gn	-us
n7	-is
n71	-is
n72	-is
n7c	-is
n7p	-is
n7p3	-is
p1-p23	LE (always)
v1d	LE (always)
v1e	-at
v1i	-o
v1r	-o
v1s	-o
v2d	LE (always)
v2e	-et

v2i	-eo
v2r	-eo
v2s	-eo
v3d	LE (always)
v3e	-it
v3i	-o
v3r	-o
v3s	-o
v4d	LE (always)
v4e	-it
v4i	-io
v4r	-io
v4s	-io
v5d	LE (always)
v5e	-it
v5i	-io
v5r	-io
v5s	-io
v6d	LE (always)
v6i	-o
v6r	-o
v6s	-o
v61a	-o
v62a	-o
v63a	-o
v64a	-o
v65a	-o
v66a	-o
v67a	-o
v68a	-o
v69a	-o
v7s	-i
v7e	-it
v8s	-i
i	=les
n	=les
v	=les
pr	=les
fe	to hyperlemma
blk	to hyperlemma
blk1	=input form
blk2	=input form / if input form "-us", change "-us" in "-or"
blk3	=input form

With codles n3\* and n7\*:

<u>les ending</u>	<u>substituted with automatic SF</u>
-in	-o
-on	-o
-c	-x
-g	-x
-d	-s
-t	-s

-3-  
**codlem**

List of available codlem along with its semantic description:

- enc: enclitics
- i: invariables
- n: uninflected nouns 1
- n1: I declension nouns
- n1/2: perfect and future participles; gerundives
- n2: II declension nouns
- n2/1: I class adjectives
- n2g: gerunds
- n2np: only neuter gerundives and only neuter past participles
- n3a: II class adjectives and only neuter gerundive comparative degree
- n3b: III declension nouns
- n3p: present participles
- n4: IV declension nouns
- n4s: supins
- n5: V declension nouns
- pr: pronominals
- nx: uninflected nouns 2
- ny: uninflected adjectives
- v: verbs not related to a specific conjugation
- v1: I conjugation verbs
- v2: II conjugation verbs
- v3: III conjugation verbs
- v4: IV conjugation verbs
- v5: e/i conjugation verbs
- va: not regular conjugation verbs
- vp: verbs at perfectum