

Syntactic REAP.PT: Exercises on Passive Transformation

Sandra Maria de Babo Lourenço

IST - Instituto Superior Técnico
L²F - Spoken Language Systems Laboratory - INESC-ID Lisboa
Rua Alves Redol 9, 1000-029 Lisboa, Portugal
`sandra.lourenco@tecnico.ulisboa.pt`

Abstract. Technology plays a very important role in education and Intelligent Computer Assisted Language Learning (ICALL) has emerged as an alternative to the conventional teaching of languages. REAP.PT is an ICALL system that aims to teach Portuguese to both native and foreign students. The goal of this work is to expand the exercises already available on REAP.PT by adding new exercises on the passive transformation, a very complex topic, for which there are very few CALL and ICALL resources available.

Keywords: Automatic Exercise Generation, Intelligent Computer-Assisted Language Learning, Passive Transformation, Portuguese, Syntactic Exercises

1 Introduction

With the increasing use of technology in our daily lives, it is only expectable that technology will also be used to the teaching and improvement of languages. With the evolution of technology, Intelligent Computer Assisted Language Learning (ICALL) has emerged as an alternative to conventional student-teacher interaction.

REAders-specific Practice (REAP)¹ is an ICALL system that was initially developed at Carnegie Mellon University (CMU) by the Language Technology Institute (LTI). Its main goal is to teach English by using automatically generated exercises, which focus on vocabulary learning and on the practice of reading skills. All exercises involve real texts taken from different sources (corpora, web) and are adapted to the student's interests and learning level.

REAP.PT is the Portuguese adaptation of the original REAP, which has been developed and improved over the years by Marujo [8], Correia [3], Marques [7], Silva [10], Freitas [5] and Figueirinha [4].

Goals. The goal of the present work is to extend the functionality already available in the REAP.PT system by developing a new module of exercises.

¹ <http://reap.cs.cmu.edu/>

Those exercises will focus on the transformation of sentences from the passive to the active voice. This is a topic that is frequently present in traditional textbook exercises and is taught in Portuguese classes, starting in the sixth grade and being continuously taught until the twelfth grade. The passive transformation can involve really complex constraints and morphosyntactic changes. Given its importance in the Portuguese curriculum it is important that there are resources that automatically generate and evaluate this kind of exercises, so that students can autonomously practice them, thus saving an important work load to language teachers.

2 State of the Art

REAP.PT. READER-specific Practice PorTuguese (REAP.PT) is the Portuguese version of REAP, a tutoring system whose main goal is to support the teaching of Portuguese as a second language. This system's architecture consists of several components. The Web Interface component provides the user interaction with the system; it is responsible for getting definitions from the Infopédia² online dictionary and for the communication between the database module and DIXI. The Database component consists of two different databases that keep track of the state of the system during execution, it contains references to the Andrew File System (AFS) and stores the content of the lexical resources used. The Daily REAP component uses the Nutch web crawler to extract from the web texts from Portuguese newspapers and transcriptions from television newscasts from the past three days. A filter chain is used to ensure that those texts are fit for educational purposes. After the filter chain is used, the texts are classified based on their topic and their readability level and then they are stored in the database. The Question Generation module automatically generates the exercises that are presented to the students. To ensure that the vocabulary presented to the students is appropriate, the Portuguese Academic Word List (P-AWL)[1] is used.

REAP.PT exercises. REAP.PT already provides several exercises, based on different aspects of the Portuguese language. Those exercises are generated according to the user's level of proficiency, his/hers topics of interest (provided by the user) and the history of all the user's completed activities, which is stored on the system. The Lexical Mahjong[7] exercise is based on the Mahjong Solitaire game. The user has to establish a correspondence between the lemma and the definition of a word. The Nominal Determiners[7] exercise has the goal of teaching the user the constraints between determinative nouns and the nouns which they determine (e.g. *bando de aves* (flock of birds), *vara de porcos* (herd of pigs)). The Collective Nouns[7] exercise is a subset of the nominal determiners exercise involving only collective nouns. The choice of mood in subordinate

² <http://www.infopedia.pt/>

clauses[7] exercise teaches the user the syntactic restrictions imposed by the subordinative conjunctions on the mode of the subordinate clause they introduced. The Pronominalization[5] exercise’s intent is to substitute a constituent with the appropriate clitic pronoun, in a given sentence. The crosswords[4] exercise is based on the structure of the usual crossword puzzles, however, unlike those puzzles who are filled with letters, the goal of this exercise is to fill each cell of the puzzle with syllables. The Prefixation[4] exercise has as its main focus the identification of words that are/seem to be, superficially, formed through prefixal derivation but constitute exceptions to the general rules of prefixation.

CALL systems overview. Table 1 presents an overview of the CALL systems currently available, in terms of their public availability and whether their exercises and feedback are automatically generated or not.

	Publicly Available	Automatically Generated Exercises	Dynamic Feedback
Ciberescola da Língua Portuguesa	X		
Centro Virtual Camões	X		
Observatório da Língua Portuguesa	X		
TAGARELA			X
WERTi	X	X	X
The Alpheios Project	X	X	X
FAST		X	X
ArtikIturri		X	X

Table 1. CALL systems overview

3 Passive Sentence Types

In the case of the action predicates, the subject of an active sentence has the semantic role of *agent/complement*, while the direct complement is the *object* or the patient of the action, the element that undergoes the action. Example (from [9]): *Pedro Nunes inventou o nóvio*. (Pedro Nunes invented the vernier).

Passive sentences are characterized by having a different alignment than that from active sentences. Example (from [9]): *O nóvio foi inventado por Pedro Nunes*. (The vernier was invented by Pedro Nunes). Both types of sentences have the same sentence type, describe the same situation, the arguments of the verb have the same semantic roles and both are characterized by the same illocutory force. In Portuguese, only active sentences where the verb selects a direct complement can undergo the passive transformation.

Five types of passive sentences can be considered. (i) Verbal passive sentences are sentences in which occurs a complex verbal group started by the auxiliary verb *ser* (be) followed by a participle corresponding to the verb of the active sentence. They typically describe situations in which one of the entities involved goes through some kind of change. (ii) Resultative passive sentences

describe a situation that is the result of a change in state, place or possession. In this kind of passive sentences, it is the verb *ficar* (become/stay) that usually occurs. Example (from [9]): *Os museus de Bagdade ficaram destruídos (como resultado dos bombardeamentos americanos)*. (The Baghdad museums were destroyed (as a result of the American bombing)). (iii) Stative passive sentences are passive sentences that are syntactically close with the passive resultative sentences but instead of focusing on the result of the change, they focus on describing stative situations whose meaning does not have anything related with the change of state. Example from [9]: *Este autor é muito conhecido*. (This author is well known). (iv) In pronominal passive sentences the passive voice is expressed through reflexive pronouns. Example: *As camisas lavam-se sempre a frio*. (The shirts are always washed cold). (v) Neutral or median passive sentences are sentences where the *object* (or *theme*) appears aligned in the subject position; the active sentence subject has a semantic role of *cause* and is either dropped or is introduced by proposition *com* (with) or other causal connectors, (e.g. *por causa de* (because of)): *As condutas entupiram (com o lixo)*. (The pipes clogged (with the garbage)).

4 Approach - Passive to Active Transformation

To automatically perform the transformation from the passive to the active voice, there are some steps that need to be done: selection of the sentences that will be used, morphological transformation of the passive verb and syntactical transformation of the sentence itself. The sentences used come from the CETEMPúblico corpus [2] and were previously processed using the STRING processing chain.

Sentence Selection. The first step for the creation of the exercises that deal with the transformation from the passive voice to the active voice is the selection of the sentences that will be used and the decision of how those sentences will be split into three levels of difficulty (easy level, medium level, difficult level).

The following constraints are used to ensure that the selected sentences have the best possible quality to be used for educational purposes: only one dependency of type COMPL_POST (representing the passive agent, which may be absent) and only one dependency of type SUBJ_PRE (representing the passive sentence subject) can be present to ensure that the sentence only has one passive verbal chain; only one dependency of type VDOMAIN (representing the passive verb) is allowed; no dependencies of type POSS (that represent possessive determiners) are allowed; the prepositional phrase (PP) that starts with the word *por* (by) needs to have a head noun that is a human noun; the number of verbs connected by the VLINK dependency is different for each difficulty level; the main verb in the verbal chain captured by the VDOMAIN dependency has to be part of the list of verbs that accept that passive transformation with the auxiliary *ser* (be), which comes from ViPER [12].

A sentence is considered to belong to the easy level if it has only one VLINK dependency and less than 55 dependency tags in the XML corpus' file. There were 6549 sentences found.

A sentence is considered to belong to the medium level if it has two VLINK dependencies and less than 65 dependency tags in the XML corpus' file. There were 1364 sentences found.

A sentence is considered to belong to the difficult level if it has three VLINK dependencies and less than 75 dependency tags in the XML corpus' file. There were 80 sentences found.

Morphological Verb Transformation. In the easy level, the verbal chain has only two verbs: the auxiliary verb (the passive auxiliary *ser* (be)); and the main verb, in the past participle form; there is only one VLINK dependency. It is the main verb that will be transformed to the active form, taking into account the tense of the verb *ser* (be). The transformation is done with the help of LexMan [11], by searching for the form of the verb that is in the past participle that has the same tense and mood as the tense of the verb *ser* (be) and in the same number as the noun that was identified as the sentence's agent.

In the medium difficulty level, the verbal chain has three verbs. The second verb is always the passive auxiliary *ser* (be) and the third verb is always the main verb (in the past participle form). There are two VLINK dependencies. The first verb will stay in the same tense, but may need to be changed from the singular form to the plural form or the opposite. This is done with the help of LexMan. The main verb is the verb that will be transformed to the active form, with the help of LexMan, by searching for this verb in the tense of the second auxiliary in the sentence. The main verb will be in the same tense as the verb *ser* (be) in the passive sentence.

In the difficult level, the verbal chain has four verbs; the third verb is always the passive auxiliary *ser* (be). There are three VLINK dependencies. The first verb will be transformed in the same way as in the medium level. The second verb is usually in the infinitive form and will stay the same in the active sentence. The main verb will be transformed to the active form in the same way as it is in the medium level.

Syntactical Transformation from Passive to Active. The transformation from the passive voice to the active voice is done by following the grammar rules explained in section 3 for the passive sentences with the auxiliary *ser* (be). However, not all sentences' transformations are straightforward and some additional steps may have to be performed.

In some passive sentences, words appearing inside the verbal chain have to be moved to an equivalent position in the active sentence. Example: (a) *O discurso parece ter sido bem entendido pelos destinatários.* (The speech seems to have been **well** understood by the recipients.). (b) *Os destinatários parecem ter entendido bem o discurso.* (The recipients seem to have understood **well** the speech.). In this case the adverb *bem* (well) is moved to after the main verb.

If the passive sentence starts with a connective adverb, usually that word has to stay in the beginning of the active sentence. The rest of the sentence will be transformed following all the basic rules. Example: (a) *Mas a proposta do coordenador da Comissão Permanente do PS foi recusada pelos sociais-democratas.* (**But** the proposal of the coordinator of the Permanent Commission of the PS was rejected by the social-democrats). (b) *Mas os sociais-democratas recusaram a proposta do coordenador da Comissão Permanente do PS.* (**But** the social democrats rejected the proposal of the coordinator of the Permanent Commission of the PS).³

When transforming passive sentences that are in the negative form, special care is needed to ensure that the active sentence also remains in the negative form and that the words that identify the negative sentence are put in their correct places. Example: (a) *A administração da RTP não pode continuar a ser nomeada pelos governos.* (The board of direction of RTP **cannot** continue to be nominated by governments). (b) *Os governos não podem continuar a nomear a administração da RTP.* (Governments **cannot** continue to nominate the board of direction of RTP). In the case of indefinite-negative determiner *nenhum* (no/none), when the negation is part of the subject of the passive sentence, the active counterpart involves the insertion of the negation adverb *não* (no): (a) *Porém, nenhuma decisão concreta foi tomada pelos britânicos em Toulouse.* (However, **no** concrete decision was taken by the British in Toulouse). (b) *Porém, os britânicos em Toulouse não tomaram nenhuma decisão concreta.* (However, the British in Toulouse took **no** concrete decision).

Some of the passive sentences have one or two dashes near the beginning, indicating, in an interview, the turn change. When this occurs, only the part of the sentence after the dashes is considered for transformation. Example: (a) *R -- O PC devia ser vestido pela Ana SALAZAR.* (**R --** The PC should be dressed by Ana SALAZAR). (b) *R -- A Ana SALAZAR devia vestir o PC.* (**R --** Ana SALAZAR should dress the PC).

In every sentence resulting from the passive-active transformation, it is ensured that the first word starts with an uppercase character, regardless of its case in the passive.

Some passive sentences have a comma or colon somewhere in the first part of the sentence (before the verbs). In most of the cases, the words behind the comma (and the comma itself) have to remain in the same place in the final sentence (the same for sentences with colon). The rest of the sentence will be transformed as usual. Example: (a) *Na recta final, o empresário foi ultrapassado, em capacidade financeira, por António Champalimaud.* (**In the final stretch**, the businessman was surpassed in financial capacity, by António Champalimaud). (b) *Na recta final, António Champalimaud ultrapassou o empresário, em capacidade financeira.* (**In the final stretch**, António Champalimaud surpassed the businessman in financial capacity).

³ While *mas* (but) is typically a coordinative conjunction, in this case, it is used in much the same way as a connective adverb, like *porém* (however).

In some cases, the passive sentences have fractional numbers. This has to be taken into account to ensure that the numbers from both sides of the decimal separator⁴ stay together in the correct side of the sentence.

When the passive sentence is enclosed by parenthesis, it is necessary to make sure that those parenthesis do not end up in the middle of the active sentence and that the active sentence is also enclosed by the parenthesis. The same thing has to be done when the whole passive sentence is enclosed by quotes. Example: (a) “*O sistema de alvarás tem sido regulado pelos empreiteiros*”. (“The permits system has been regulated by contractors”). (b) “*Os empreiteiros têm regulado o sistema de alvarás*”. (“The contractors have regulated the permits system”).

When only part of the passive sentence is enclosed by parenthesis, it is necessary to make sure that the part between the parenthesis is also between the parenthesis in the active sentence. The same thing has to be done when there is some part of the passive sentence enclosed by quotes. Example: (a) *O estudo foi desenvolvido pelo Departamento de Planeamento Estratégico (DPE) da autarquia*. (The study was developed by the Department of Strategic Planning (DPE) of the municipality). (b) *O Departamento de Planeamento Estratégico (DPE) da autarquia desenvolveu o estudo*. (The Department of Strategic Planning (DPE) of the municipality developed the study).

The active sentence has to end with the same punctuation mark as the passive sentence.

At the very end of the process, the final sentence may need to be corrected: since all the contractions in the original passive sentence taken from the corpus had been previously resolved, in the final active sentence those contractions have to be reconstructed.

5 Evaluation

The evaluation was performed on a random sample of sentences from each of the three difficulty levels. This sample consisted of 100 sentences from the easy level, 100 sentences from the medium level and 40 sentences from the difficult level. The number of sentences from the difficult level is smaller because less sentences from this level were found in the corpus than from the other two levels. Each pair of passive-active sentences was evaluated by an expert, who determined the sentences that had errors and the nature of the errors present.

Error Analysis. One of the most common errors found happens when the past participle in the passive sentence appears coordinated with another one, the active sentence is incorrect as the system does not contemplate this type of coordination yet. Example: *Este foi julgado e condenado pelo Tribunal de Círculo de Matosinhos a três anos de prisão*. (He was **tried and convicted** by the Matosinhos District Court to three years in prison).

⁴ In Portuguese orthographic convention, the decimal separator is the comma (,), while the thousands' separator is the dot (.).

A similar problem occurs with complex coordinated agent complements, as only the first PP element keeps the proposition *por* (by) and the remaining are parsed as NPs, thus coordinating rules are not propagated from these to the initial PP: 1. *O mercado é feito pelos produtores, distribuidores e exibidores.* (The market is made by the **producers, distributors and exhibitors**). 2. **Os produtores fazem o mercado, distribuidores e exibidores.* (The **producers** make the market, **distributors and exhibitors**). 3. *Os produtores, distribuidores e exibidores fazem o mercado.* (The **producers, distributors and exhibitors** make the market). In this example, the full agent PP includes *distribuidores e exibidores* (distributors and exhibitors).

In other cases, a sentence was identified as a passive sentence even though it is not. This happens due to a case of subject inversion. Example: *Todavia, não foram muitas as saídas deixadas pelos socialistas para uma remota resposta de Cavaco.* (However, there were not many exits left by the Socialists to a remote response from Cavaco). The base order would be: *Todavia, as saídas deixadas pelos socialistas para uma remota resposta de Cavaco não foram muitas.* (However, the exits left by the Socialists to a remote response from Cavaco were not many).

In some cases, the use of commas in the passive sentence renders the passive-active transformation inadequate. In this case, the agent PP is separated from the main clause by comma, as if it were a sort of apposition. The resulting sentence 2, though interpretable, is incorrect, as this comma, in the active sentence, ends up between the direct and the indirect (dative) complement. This happens due to an error in part of the transformation algorithm. 1. *O caso só foi contado às autoridades há três dias, por três sobreviventes da tribo.* (The case was only told to the authorities three days ago, by three survivors of the tribe). 2. **Três sobreviventes da tribo só contaram o caso, às autoridades há três dias.* (Three survivors of the tribe only told the case, to the authorities three days ago). 3. *Três sobreviventes da tribo só contaram o caso às autoridades há três dias.* (Three survivors of the tribe only told the case to the authorities three days ago).

There were also some other problems related to the use of comma and the position of adverbials in the sentences, due to the transformation algorithm still incipient treatment of adverbial complements. In sentence 1 the temporal adverbial should have been outside the scope of the transformation algorithm. Since this adverbial has been moved away from its initial position without separating it from the remainder of the clause by commas, a comma has been incorrectly placed between the main clause and the direct object, rendering the sentence ungrammatical. 1. *A partir do próximo ano lectivo, o Externato Paulo VI, em Gondomar, vai deixar de ser administrado pelos padres Capuchinhos.* (**Starting next school year**, the Externato Paul VI, in Gondomar, will no longer be administered by the Capuchin priests). 2. **Os padres Capuchinhos vão deixar de administrar a partir do próximo ano lectivo, o Externato Paulo VI, em Gondomar.* (The Capuchin priests will no longer administer **starting next school year**, the Externato Paul VI, in Gondomar). 3. *A partir do próximo ano*

lectivo, os padres Capuchinhos vão deixar de administrar o Externato Paulo VI, em Gondomar. (**Starting next school year**, the Capuchin priests will no longer administer the Externato Paul VI, in Gondomar).

Since the transformation algorithm does not take into account relative clauses, all sentences with relative clauses are transformed incorrectly. Relative clauses are sentences that are introduced by the relative pronoun, e.g. *que* (that), and modify a noun. The passive transformation should only be applied to the relative clause, that is, after *que*, and whatever is before the pronoun should be kept unchanged. Example: 1. *um dos pontos em que o relato de Champalimaud é analisado por Freire Antunes.* (one of the points where Champalimaud's report is analyzed by Freire Antunes). 2. **Freire Antunes que analisa um dos pontos no relato de Champalimaud.* (Freire Antunes that analyzes one of the points of Champalimaud's report). 3. *um dos pontos em que Freire Antunes analisa o relato de Champalimaud.* (one of the points where Freire Antunes analyzes Champalimaud's report).

Passive sentences that have apposition phrases or clauses may have errors when transformed to the active voice because these appositions are hard to detect automatically and failing to detect them makes the algorithm leave them in place, away from their antecedent. Example: 1. *A informação da RTP vai passar a ser dirigida por Manuel Rocha, até agora responsável pelo pelouro no Norte.* (The information of RTP will now be led by **Manuel Rocha, until now responsible for this area in the north**). 2. **Manuel Rocha vai passar a dirigir a informação da RTP, até agora responsável pelo pelouro no Norte.* (Manuel Rocha will now lead the information of RTP, **until now responsible for this area in the north**). 3. *Manuel Rocha, até agora responsável pelo pelouro no Norte, vai passar a dirigir a informação da RTP.* (Manuel Rocha, **until now responsible for this area in the north**, will now lead the information of RTP).

When there are personal pronouns in a sentence, the verb has to be transformed in a different way, taking into account those pronouns [5]. Since pronouns were not taken into account in the algorithm here developed, some sentences with personal pronouns are transformed incorrectly. Example: 1. *Mais, ela tem sido considerada, por numerosos historiadores, eventualmente com razão, como o verdadeiro motor de civilização.* (More, **she has been considered**, by many historians, perhaps with reason, as the driving force of civilization). 2. **Mais, numerosos historiadores têm considerado ela, eventualmente com razão, como o verdadeiro motor de civilização.* (More, many historians **have considered her**, perhaps with reason, as the driving force of civilization). 3. *Mais, numerosos historiadores têm-na considerado, eventualmente com razão, como o verdadeiro motor de civilização.* (More, many historians, **have considered her**, perhaps with reason, as the driving force of civilization).

When the passive sentence is in the negative form, different steps have to be performed in the transformation to the active voice. The transformation algorithm takes into account the basic cases (sentences with *não* (not) or *nunca* (never)) but does not take into account the more complex cases. Example: 1. *Até*

sábado, Anabela ainda não tinha sido vista por nenhum médico. (Until Saturday, Anabela had **not yet** been seen by **any** doctor). 2. **Até sábado, nenhum médico não tinha visto ainda Anabela.* (Until Saturday, **no** doctor had **not yet** seen Anabela). 3. *Até sábado, ainda nenhum médico tinha visto Anabela.* (Until Saturday, **still no** doctor had seen Anabela).

If a verb is prefixed, and the prefixed form has the same construction as the base form, STRING considers only one lemma and marks the prefixed form with the base lemma plus the information that it features a prefix. Since the algorithm does not take this information into account, it uses the (unprefixed) lemma to form the active sentence thus producing an incorrect paraphrase.

6 Conclusion and Future Work

Technology plays a very important role in education and in particular in language teaching, and it will continue to do so in the future. CALL and ICALL systems like REAP and REAP.PT have an important role because having exercises tailored to the user's interests and proficiency level improves the pedagogical experience. The exercises developed in this work focus on the automatic transformation of sentences from the passive to the active voice. This is a complex topic for which there are very few CALL and ICALL resources available. This complexity is also why it is difficult to obtain perfect results when automatically performing this transformation. We believe that the work developed could be the start of a useful tool in the teaching and the learning of the Portuguese language, for both native speakers and non native speakers.

Future Work. This work can be integrated with REAP.PT, as one of its syntactical games. For this to happen, the following things need to be done: adaptation of REAP.PT's graphical interface; implementation of a help feature; implementation of a score system; implementation of automatic feedback.

Taking into account the errors detected during the evaluation process, the following steps can be taken to improve the transformation of sentences from the passive to the active voice: correcting the placement of commas in the sentences; taking relative sentences into account; detecting more cases in which passive sentences cannot be transformed to the active voice, even though they may have all the dependencies that identify a sentence as a passive; improving the transformation of sentences with adverbials and adverbs; improving the transformation of sentences with personal pronouns; correcting the identification of the number of the subject of the active sentence in the case of coordinated agent PPs; improving the transformation of sentences that are in the negative form, or with negative determiners.

References

1. Baptista, J., Costa, N., Guerra, J., Zampieri, M., Cabral, M., and Mamede, N. (2010). P-AWL: Academic Word List for Portuguese. In Proceedings of the 10th

- International Conference on Computational Processing of the Portuguese Language, *volume 6001 of PROPOR 2010*, pages 120-123. Springer-Verlag, Berlin, Heidelberg.
2. Santos, D. & P. Rocha (2001). Evaluating CETEMPúblico, a free resource for Portuguese. In Proceedings of the 39th Annual Meeting on the Association for Computational Linguistics, ACL '01, Morristown, NJ, SA, pp. 450-457. Association for Computational Linguistics.
 3. Correia, R. Automatic Question Generation for REAP.PT Tutoring System. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2010.
 4. Figueirinha, P. Syntactic REAP.PT - Exercises on Word Formation. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2013.
 5. Freitas, T. Syntactic REAP.PT - Exercises on Clitic Pronouncing. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2012.
 6. Mamede, N., Baptista, J., Diniz, C., & Cabarrão, V. (2012). STRING: An Hybrid Statistical and Rule-Based Natural Language Processing Chain for Portuguese. In Proceedings of the 12th International Conference on Computational Processing of the Portuguese Language - Demo sessions <http://www.propor2012.org/demos/DemoSTRING.pdf>, PROPOR 2012.
 7. Marques, C. Syntactic REAP.PT. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2011.
 8. Marujo, L. REAP em Português. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2009.
 9. Raposo, E., Nascimento, F. B. do, Mota, M. A., Segura, L., Mendes, A., Vicente, G. Gramática do Português I. Fundação Calouste Gulbenkian, 2013.
 10. Silva, A. Pictorial REAP.PT. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2011.
 11. Vicente, A. LexMan: um Segmentador e Analisador Morfológico com Transdutores. Master's Thesis, Instituto Superior Técnico - Universidade Técnica de Lisboa, 2013.
 12. Baptista, J. ViPEr: uma base de dados de construções léxico-sintáticas de verbos do Português Europeu. Universidade do Algarve/CECL & INESC ID Lisboa/L²F, 2012.