University of Helsinki
Faculty of Arts
Department of Digital Humanities

Master's thesis

Computer-Assisted Language Learning based on Authentic Texts: applications to Italian

Elena China-Kolehmainen
012629961

Master's Programme in Linguistic Diversity and Digital Humanities

Supervisor: Roman Yangarber 29.7.2021

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty Faculty of Arts	Koulutusohjelma – Utbildningsprogram – Degree Programme Master's Programme in Linguistic Diversity and Digital Humanities	
Opintosuunta – Studieinriktning – Study Track Language Technology		
Tekijä – Författare – Author Elena China-Kolehmainen		
Työn nimi – Arbetets titel – Title Computer-Assisted Language Learning based on Authentic Texts: applications to Italian		
Työn laji – Arbetets art – Level Master’s thesis	Aika – Datum – Month and year July 2021	Sivumäärä– Sidoantal – Number of pages 53 [4]

Tiivistelmä – Referat – Abstract

Computer-Assisted Language Learning (CALL) is one of the sub-disciplines within the area of Second Language Acquisition. Clozes, also called fill-in-the-blank, are largely used exercises in language learning applications. A cloze is an exercise where the learner is asked to provide a fragment that has been removed from the text. For language learning purposes, in addition to open-end clozes where one or more words are removed and the student must fill the gap, another type of cloze is commonly used, namely multiple-choice cloze. In a multiple-choice cloze, a fragment is removed from the text and the student must choose the correct answer from multiple options. Multiple-choice exercises are a common way of practicing and testing grammatical knowledge.

The aim of this work is to identify relevant learning constructs for Italian to be applied to automatic exercises creation based on authentic texts in the Revita Framework. Learning constructs are units that represent language knowledge. Revita is a free to use online platform that was designed to provide language learning tools with the aim of revitalizing endangered languages including several Finno-Ugric languages such as North Saami. Later non-endangered languages were added. Italian is the first majority language to be added in a principled way. This work paves the way towards adding new languages in the future. Its purpose is threefold: it contributes to the raising of Italian from its *beta* status towards a full development stage; it formulates best practices for defining support for a new language and it serves as a documentation of what has been done, how and what remains to be done.

Grammars and linguistic resources were consulted to compile an inventory of learning constructs for Italian. Analytic and pronominal verbs, verb government with prepositions, and noun phrase agreement were implemented by designing pattern rules that match sequences of tokens with specific parts-of-speech, surfaces and morphological tags. The rules were tested with test sentences that allowed further refining and correction of the rules. Current precision of the 47 rules for analytic and pronominal verbs on 177 test sentences results in 100%. Recall is 96.4%. Both precision and recall for the 5 noun phrase agreement rules result in 96.0% in respect to the 34 test sentences. Analytic and pronominal verb, as well as noun phrase agreement patterns, were used to generate open-end clozes.
Verb government pattern rules were implemented into multiple-choice exercises where one of the four presented options is the correct preposition and the other three are prepositions that do not fit in context. The patterns were designed based on *colligations*, combinations of tokens (collocations) that are also explained by grammatical constraints. Verb government exercises were generated on a specifically collected corpus of 29074 words. The corpus included three types of text: biography sections from Wikipedia, Italian news articles and Italian language matriculation exams. The last text type generated the most exercises with a rate of 19 exercises every 10000 words, suggesting that the semi-authentic text met best the level of verb government exercises because of appropriate vocabulary frequency and sentence structure complexity.
Four native language experts, either teachers of Italian as L2 or linguists, evaluated *usability* of the generated multiple-choice clozes, which resulted in 93.55%. This result suggests that minor adjustments i.e., the exclusion of target verbs that cause multiple-admissibility, are sufficient to consider verb government patterns usable until the possibility of dealing with multiple-admissible answers is addressed.

The implementation of some of the most important learning constructs for Italian resulted feasible with current NLP tools, although quantitative evaluation of precision and recall of the designed rules is needed to evaluate the generation of exercises on authentic text. This work paves the way towards a full development stage of Italian in Revita and enables further pilot studies with actual learners, which will allow to measure learning outcomes in quantitative terms.

Avainsanat – Nyckelord – Keywords
language technology, Italian, CALL, Computer-Assisted Language Learning, language learning, cloze generation, automatic exercise generation

Säilytyspaikka – Förvaringställe – Where deposited
Helsinki University Library

Muita tietoja – Övriga uppgifter – Additional information

Contents

1 Motivation and research questions

Revita is a free to use online platform where automatic cloze generation is used to create exercises from a text. Revita was designed to provide language learning tools with the aim of revitalizing endangered languages including several Finno-Ugric languages such as Udmurt, Meadow Mari and North Saami [27]. At the beginning of the writing process of this thesis (November 2020) Revita provides language learning exercises also for non-endangered languages such as Russian and Finnish. These two languages were originally developed to account for code-switching in the endangered Finno-Ugric languages, Finnish for North Saami, and Russian for the rest. While both Finnish and Russian were added to Revita in an *ad hoc* manner for special reasons and evolved as "byproducts" they are now at the most developed phase among all available learning languages in Revita. All the other languages are in early stages of development. The languages at *Beta* phase are Erzya, Komi-Zyrian, Meadow-Mar, North-Saami, Sakha, Tatar, Turkish, Catalan, Chinese, French, German, Italian, Kazakh, Portuguese, Russian, Spanish, Swedish and Udmurt.

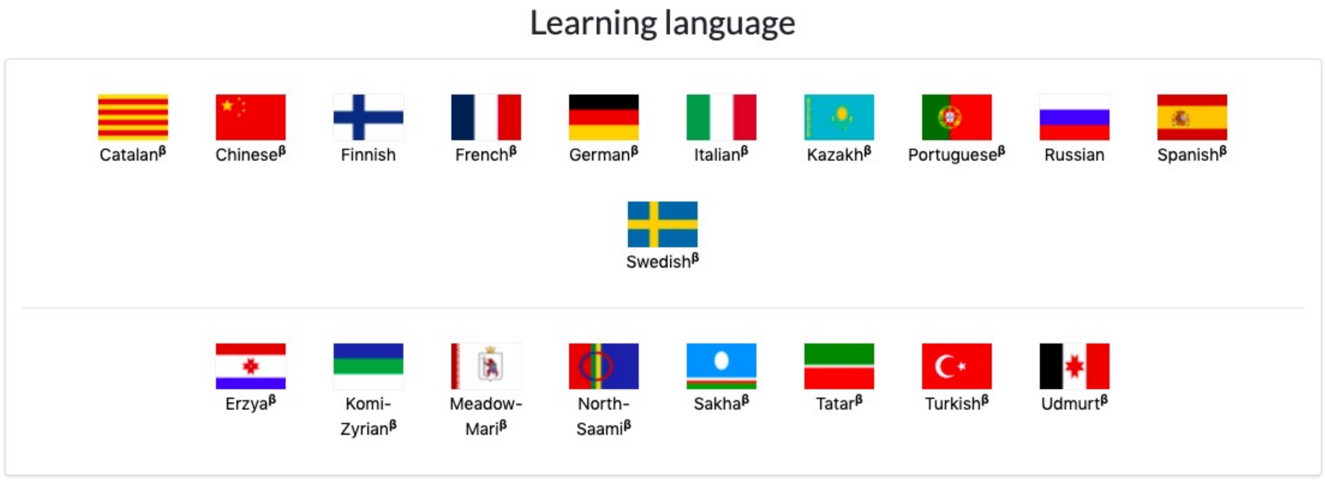


Figure 1: Available learning languages in Revita. Finnish and Russian are languages developed to the most advanced stage.

Revita's system is constantly evolving as several people are actively working on its development. Italian is the first majority language which is being added to Revita in a principled way. This work paves the way toward adding new languages in the future. Its purpose is threefold:

- contribute in the raising of Italian in Revita from its *beta* status to a full development stage
- formulate best practices for defining support for a new language in the Revita Framework

- serve as a documentation of what has been done, how and and what remains to be done

Specifically, the research questions of this thesis are:

RQ1: What are the most important Italian morphosyntactic constructs for language learning?

RQ2: Is it feasible to implement some of the identified morphosyntactic constructs as automatic exercises in Revita with the currently available NLP tools?

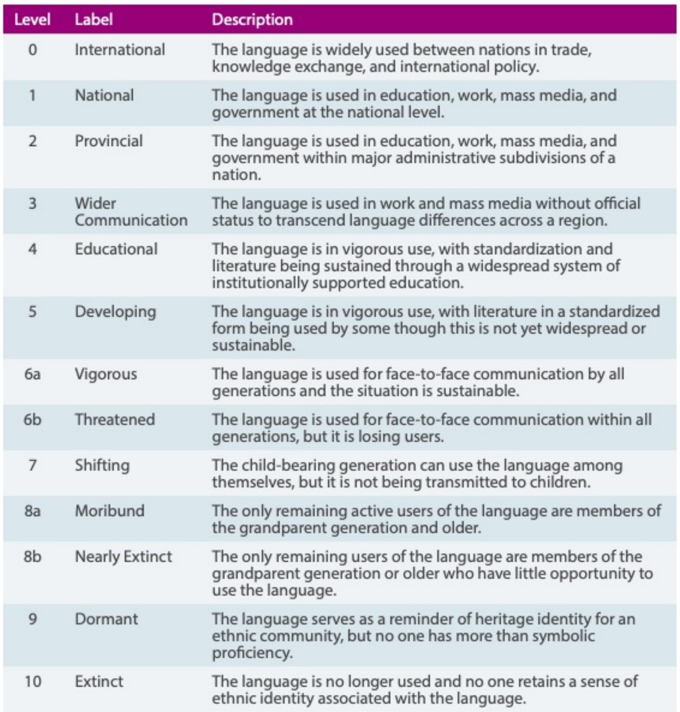
RQ3: How the implemented constructs can be evaluated?

Answering the research questions will enable further pilot studies with actual learners and teachers, which will allow to measure in rigorous and quantitative terms the usefulness of Italian in Revita.

Italian was chosen as first majority language to be fully implemented because it had potential for pilot studies with real users — the main motivator for development of new languages in Revita. These users are in the South Tyrol region of Italy, where population is about evenly divided between Italian and German speakers, and Italian is a required language at schools and universities. The choice of language to work with was also motivated by the authors's personal involvement in the Italian community in Finland. The possibilities offered by Revita for the language learners sparked particular interested because of the possibility to use authentic texts for exercise creation, an attractive opportunity to encourage active independent use of the target language not only in language learners but also in heritage speakers.

According to the Ethnologue [16] the vitality status of Italian is estimated at level 1 on the Expanded Graded Intergenerational Disruption Scale (EGIDS), meaning that "The language is used in education, work, mass media, and government at the national level." The EGIDS consists of 13 levels from 0 to 10 with some intermediate levels. The greater the number the greater the level of disruption to the intergenerational transmission of the language.

Italian is not an endangered language. Nevertheless, Revita's platform can be considered a valid tool to support learning and activation of Italian, as well as any other language, as a heritage language since the platform is aimed at people "who already possess some competence in the target language — intermediate to advanced students (i.e., not for the very beginners)" [27]. This is often the competence level of many heritage languages speakers.



Level	Label	Description
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.
1	National	The language is used in education, work, mass media, and government at the national level.
2	Provincial	The language is used in education, work, mass media, and government within major administrative subdivisions of a nation.
3	Wider Communication	The language is used in work and mass media without official status to transcend language differences across a region.
4	Educational	The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.
5	Developing	The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
6a	Vigorous	The language is used for face-to-face communication by all generations and the situation is sustainable.
6b	Threatened	The language is used for face-to-face communication within all generations, but it is losing users.
7	Shifting	The child-bearing generation can use the language among themselves, but it is not being transmitted to children.
8a	Moribund	The only remaining active users of the language are members of the grandparent generation and older.
8b	Nearly Extinct	The only remaining users of the language are members of the grandparent generation or older who have little opportunity to use the language.
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.
10	Extinct	The language is no longer used and no one retains a sense of ethnic identity associated with the language.

Figure 2: Expanded Graded Intergenerational Disruption Scale (EGIDS) [16]
.

In this sense Revita can support learning of a heritage language that is at risk of not being transmitted at a family level, meaning that it is not the language itself in the position of being endangered. Instead the proficiency of its heritage speakers in a specific setting is at risk and in some cases the language is not transmitted from the child-bearing generation to the children, reflecting EGIDS' scale level 7 i.e., "shifting". According to Austin and Sallabank (2011), language shift and language attrition take place when "speakers of a language make a decision to stop speaking their ancestral tongue or not to speak it to their children and to use another language instead. In such cases of more gradual shift and attrition, speakers abandon their language in favour of a more dominant and 'useful' language over the course of one or more generations. This other language is almost always the language of a majority culture, usually in terms of population..." [5]. Linguists and sociolinguists generally agree on the fact that usually it takes three generations for *language shift* to happen [15] in a migrants community. Fishman (1970) describes language shift as the progressive loss of *domains* in which the language is used. In the first phase the majority language is only used instead of the heritage language in settings outside family context and for work purposes. In the second and third phase the overlapping of domains is stronger. In the fourth phase the heritage language is used in very few private domains, such as in a family context. Revita's exercises are created from any text chosen by the users themselves and can therefore support language learning and activation in any domain.

Heritage speakers have often a so called passive competence of the language. Here the term *passive competence* refers to the ability to understand a language either in spoken or written form while *active competence* is the ability to product it [17]. Revita provides an opportunity to improve competence “through *active* language use” [27].
Additionally Revita is a useful tool for foreign language learning at an intermediate and advanced level. From a pedagogical point of view, the revolutionary aspects of Revita are:

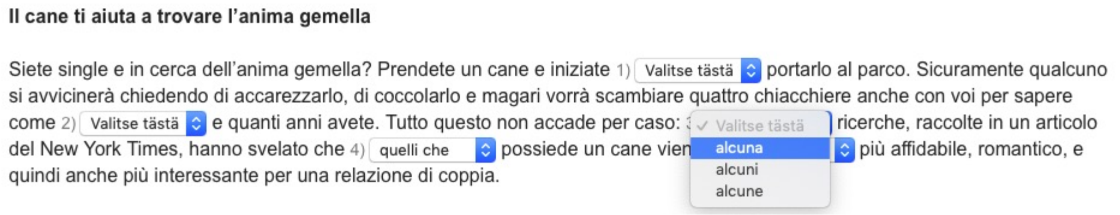
- the ability to adapt to the competence level of the learner
- the possibility to use any text of choice to automatically create the exercises

According to statistics of the Public Register of Italian Residents Abroad at date 31.12.2018, there were 5 288 281 Italians living abroad, of which 3 094 366 were families [1]. Almost one million of the Italians residents abroad were under the age of 20. A large-scale emigration wave, caused by socio-economic difficulties resulting from the financial crisis, is reported at present. This phenomenon has especially been observed amongst the age group 21-40. Figures of Italians abroad rose from 3,106,251 in 2006 to 4,636,647 in 2015, growing by 49.3% in just ten years. With already a conspicuous number of young people living abroad and more young adults and families emigrating, it becomes even more important to transmit the heritage language to the next generation and activate it to avoid the phenomenon of shifting and cultural loss. Aaltio’s ethnographic study [2] confirms that in some cases, speakers in the Italian-Finnish community have only a passive competence in the heritage language. Automatic exercise generation from a text of one’s own interest provides a good motivation for activating and developing the heritage language in young adults and even children.

Automated exercise generation can additionally support language learning on intermediate and advanced levels. In so doing it can provide learning tools for on of the most studied language in the world.

According to Bruner (1971) “What is taught should be self-rewarding by some existential criterion of being ‘real’, or ‘exciting’ or ‘meaningful’.” Being able to activate language through a text that is relevant and meaningful to the learner is of inestimable value from the point of view of learning outcome. The process of learning is meaningful when it is perceived as relevant on a personal level and the learner has not only a cognitive but also an emotional connection with the object of their learning [18].

Moreover, Revita can be used to collect learner corpora of Italian (and any other implemented language) and to further improve teaching and understanding of language learning mechanisms [26], [33]. Finally, automated generation of exercises could be applied also in the generation process of standardized exams such as the Finnish Matriculation Examination [44], not only by reducing human effort but also providing the students a tool for language learning practice.



Il cane ti aiuta a trovare l'anima gemella

Siete single e in cerca dell'anima gemella? Prendete un cane e iniziate 1) Valitse tästä portarlo al parco. Sicuramente qualcuno si avvicinerà chiedendo di accarezzarlo, di coccolarlo e magari vorrà scambiare quattro chiacchiere anche con voi per sapere come 2) Valitse tästä e quanti anni avete. Tutto questo non accade per caso: 3 ricerche, raccolte in un articolo del New York Times, hanno svelato che 4) quelli che possiede un cane vien più affidabile, romantico, e quindi anche più interessante per una relazione di coppia.

Figure 3: An example of the Finnish Matriculation Exam for Italian, cloze exercise.

2 Prior work

In the following subsections a short overview of the interdisciplinarity involved in automatic cloze generation will be illustrated. In subsection 2.1 the main characteristic of Italian as a language to learn and to activate is briefly summarized. Section 2.2 is an overview of clozes in Computer-Assisted Language Leaning. In section 2.3 the basic idea of exercise creation in Revita is presented.

2.1 Learning Italian

Italian has a simpler nominal morphology compared to other languages implemented in Revita such as Finnish and Russian. However, Italian shows a highly complex verb morphology, including analytic verb forms and other challenging linguistic phenomena. In the following paragraphs some of the most salient learning constructs of Italian will be broadly illustrated. In section 3 these constructs and their implementations into exercises will be presented in more detail.

Italian **verbal morphology** encodes tense, modality, voice and — in some cases — aspect through suffixes attached to the verb root and through auxiliaries combined with a participle. Finite moods encode also person and number of the subject. In analytic verbal forms the auxiliary encode person and number, and the participle encodes number and gender. For example the finite analytic verb *è stata mangiata* carries information on mood Indicativo, tense Passato Prossimo, passive voice, third person, singular feminine. The auxiliary *è* encodes 3rd singular person and the past participles *stata mangiata* encodes singular number and feminine gender [39].

(1)	*è*	*stat-a*	*mangiat-a*
	be.IND.PRS.3SG	be.PTCP-F.SG	eat.PTCP-F.SG

‘was eaten’

where:

Indicative Present + Participle + Participle = Passive Indicative Passato Prossimo

Verb inflection represents a challenge for language learners not only because of a substantial number of moods and respective tenses, but also due to the fact that in analytic verb formation the choice of the auxiliary (either *essere* or *avere*) is not straightforward.

Transitive verbs like *mangiare* (to eat) have auxiliary *avere* (have). Roughly 50% of intransitive verbs like *andare* (to go) have auxiliary *essere* (be) and the other 50%, like *parlare* (to speak), have auxiliary *avere.* There’s no straightforward criteria to know whether an intransitive verb is accompanied by auxiliary *essere* or *avere*, only some tendencies: usually auxiliary is *avere* for verb that express an action actually conducted by the subject like *dormire* (to sleep). For verbs who’s subject undergoes the action *nascere* (to be born) and for verbs of motion and *andare* (to go)[1] the auxiliary is *essere.* In addition, some verbs such as *cambiare* (to change) can be both transitive (auxiliary *avere*) and intransitive (auxiliary *essere*) and have different meanings depending on transitivity. In other languages the contrast in transitivity is expressed by two different verbs, in Finnish for example by “muuttaa” vs. “muuttua”. Finally, some verbs can rely either on auxiliary *essere* or *avere* interchangeably. For example both *è piovuto* (it rained) and *ha piovuto* (it rained) are valid verbal forms [41].

Another stumbling block in learning Italian consists in prepositions. The use of prepositions with nouns, verbs, adjectives and other expressions is hard to generalize exhaustively. Teaching the correct usage of prepositions represents a considerable challenge [37]. In many cases a token can be followed by a range of prepositions depending on the meaning we want to convey:

(2)	a.	*Quest-a*	*è*	*un-a*	*torta*	***di***	*mel-e.*
		this.DEM-F.SG	be.PRS.3SG	a.DET.INDF-F.SG	cake.F.SG	of.INS	apple.F-PL

‘This is an apple pie’

(3)	*Quest-a*	*è*	*un-a*	*torta*	***per***	*Maria*
	this.DEM-F.SG	be.PRS.3SG	a.DET.INDF-F.SG	cake.F.SG	for.DAT	Maria

‘This is a cake for Maria’

[1] *camminare* (to walk) is an exception since it requires auxiliary *avere*

(4)	*Quest-a*	*è*	*un-a*	*torta*	**da**	*Maria*
	this.DEM-F.SG	be.PRS.3SG	a.DET.INDF-F.SG	cake.F.SG	from.A	Maria
	‘This is a cake from Maria’					

For the above examples the student must learn the correct usage of different prepositions based on the semantic relationship with the token they point to. So the token *torta* (apple) is followed by the preposition *di* when we want to express the material or instrument it is made of. It is instead followed by preposition *per* when we want to express that *cake* has a receiver, and by preposition *da* to say that *cake* has a sender or an agent. In some cases, instead, there is a specific preposition to be used with a particular token, whether before or after it. This learning construct is **government** (*reggenza* in Italian). Although there are no case inflection in Italian (except for a minimal case paradigm for personal pronouns) and therefore no proper case government, there is a number of verbs, nouns and adjectives with a specific rection, meaning that a particular word can or must appear with its governee word. Traditional case government is common, for example, in Finnish. For example the relation between the verb *rakastaa* and its argument is that the case of the governee must be partitive: *Pekka rakastaa Merjaa* (Pekka loves Merja). Similarly but in a broader sense, many Italian verbs, nouns and adjectives govern a specific preposition. The combination of tokens, a *collocation*, that is also explained by a grammatical constraint is called *colligation* by some linguists [45, 6]. This type of government, or rection, implies a dependency relation between governor and argument. For instance the adjective *sensibile* (sensitive) can only be followed by preposition *a*, whether alone or concatenated and possibly contracted together with an article like in the following example:

(5)	*Sono*	*sensibile*	**alle**	*lusinghe*
	be.PRS.1SG	sensitive.F.SG	to.ART.F.PL	flattery.F.PL
	‘I am sensitive to flatteries’			

In example (5) the preposition *a* is concatenated and contracted with the feminine plural article *le*. Together they form the articulated preposition (*preposizione articolata*) *alle* [14]. The choice of the article (*le*) is bind to the features of the following token, *lusinghe*, which determines gender (feminine), number (plural) and presence or absence of elision depending on the initial character of the token (no elision because initial letter of token *lusinghe* is a consonant). The choice of the preposition, instead, is bind to the adjective *sensibile* which governs it. Selecting the correct governee preposition is one of the most difficult learning tasks both for L1 and L2 learners [3]. It requires repetition and exposure to authentic language use.

Other linguistic phenomena typical of Italian that represent a challenge in language learning are the above mentioned contracted forms of articulated prepositions (*preposizioni articolate*) and clitics. Articulated prepositions are prepositions combined and contracted with articles such as in *dello*, where the preposition *di* is conactenated and contracted with the article *lo*:

(6) *dello*
of.ART.M.SG
'of the'

Clitics can forms complex contracted forms such as *gliele* (them to him/her) and even attach to verbs clitisicing them such as in *dargliele*:

(7) *dar-glie-le*
give.INF-he.DAT.MF.SG-them.ACC.F.PL
'give them to him/her'

Personal pronouns are the only lexical category with a (minimal) inflection in cases, a reminiscence from Latin. Otherwise the **nominal morphology** has two main morphological features, namely gender and number. These features are compulsorily marked both on the head of the noun phrase (NP) and on the other constituents by inflectional *portmanteau* morphemes that encode simultaneously gender and number:

(8) *gatt-e* *pelos-e*
cat.NOUN-F.PL furry.ADJ-F.PL
'furry cats'

In the noun phrase *gatte pelose* the morpheme -*e* encodes both gender feminine and number plural [39]. Constituents of the noun phrase must agree in gender and number.

Nouns referring to inanimate objects have an arbitrary *grammatical* gender. The speaker must know whether a noun has *masculine* or *feminine* grammatical gender. Most frequently nouns with feminine grammatical gender end in "a" while masculine grammatical gender is denoted by ending "o". Exceptions to this general tendency do exist and they represent a challenge for language learners. For instance *sistema* is masculine, *mano* is feminine. Nouns ending in "e" are either feminine or masculine. Their gender cannot be inferred from the ending: *fiore* (flower, masculine), *luce* (light, feminine). Adjectives ending in "e" can be both feminine or masculine: *gentile* (kind, feminine and masculine).

Plural ending is generally “i” for masculine and “e” for feminine. Adjectives inflect following the same paradigm. On the other hand some nouns are invariable in plural: *auto* is both singular and plural [39]. Nouns referring to animate entities can have *natural* gender and inflect in both feminine and masculine: *gatta* (cat, feminine) vs. *gatto* (cat, masculine), *scultrice* (sculptor, feminine) vs. *scultore* (scultor, masculine).

2.2 Clozes in Computer-Assisted Language Learning

Computer-Assisted Language Learning (CALL) is one of the sub-disciplines within the area of Second Language Acquisition (SLA) [8]. While academics seem to agree on the interdisciplinary nature of CALL and although CALL is now the accepted acronym, other terminology such as CASLA (Computer-Assisted Language Learning in Second Language Acquisition), used by [9] refers to the same research area.

The PLATO project, started at the University of Illinois in 1960, can be considered the birth of CALL. “The first system, PLATO I, was simply a teletype terminal attached to a mainframe, scarcely more than a typewriter that could occasionally talk back” [21].

From these early experiments of computer as tools, CALL-systems evolved into adaptive intelligent tutoring systems providing data “which encourage the student to develop and confirm or refute hypotheses, rather than playing a passive role” [8]. Intelligent tutoring system (ITS) can dynamically provide learning content which is appropriate for the student’s understanding. “Many traditional ITSs, however, have static predefined contents without regard to individual preferences. It causes that students lose their motivation” [40]. Being able to choose the input text i.e., content can impact positively motivation and thus learning outcomes.

In Intelligent Computer-Assisted Language Learning (ICALL) NLP techniques are developed to analyse learner language by tutoring systems. According to Meureurs (2012) in language learning there’s a need to expose learners “to native language and its properties”, and to automatically generate activities from authentic texts. NLP used to process native language in Authentic Texts is referred to by the acronym ATICALL.

Clozes, also called fill-in-the-blank, are largely used exercises in language learning. A cloze is an exercise where the learner is asked to provide a fragment that has been removed from the text. The cloze procedure was initially developed in 1953 to measure readability of a text. “The method consisted of simply deleting every nth word from a passage and replacing it with a blank of standard length” [36]. Later its use was extended to measure reading skills, vocabulary usage and reading comprehension in native and non-native speakers. For non-native speakers the requirement of having to fill the gap with the exact word missing was shown to lower test results and a method which allows any contextually acceptable responses was adopted [36].

For language learning purposes, in addition to **open-end clozes** where one or more words are removed and the student must fill the gap, another type of cloze is commonly used, namely **multiple-choice clozes**. In a multiple-choice cloze a fragment or word is removed from the text but the learner is provided different answers to choose from. Multiple-choice exercises are probably "the most common way of testing grammatical knowledge" [30] because they are easily graded and they can cover many different grammatical constructs. According to [29] "it is much easier to make cloze than multiple choice exercises".

Open-end clozes have two components i.e., a suitable fragment to gap in the sentence and the answer to the gap. A multiple-choice cloze has traditionally three components: the sentence with a gap, the correct choice to the gap as the key and the other incorrect choices as the *distractors* [10]. The design of clozes involves different challenges depending of the type of cloze. Distractors must not leave space for ambiguity of ambivalence: there must be only one valid answer. At the same time they must be sufficiently similar to the correct answer. Lee and Seneff state in other words that "a good distractor must satisfy two requirements. First and foremost, it must result in an incorrect sentence. Secondly, it must be similar enough to the key to be a viable alternative". To mirror these two requirements, Lee and Seneff consider the evaluation metrics of *usability* and *difficulty* (or *facility index*). A multiple-choice cloze is *usable* when there's only one correct answer to it. *Difficulty* in turn measures how tricky or obviously wrong a distractor is.

Lee and Seneff propose methods to generate distractors for English prepositions and claim that the quality of a multiple-choice cloze depends on the choice of distractors [31]. They propose two methods in addition to the commonly used baseline. The first is based on collocations and the second on non-native corpora. Both were found to be "more successful in attracting users than a baseline that relies only on word frequency, a common criterion in past research". They report a usability of 96.3% of generated clozes, although they don't report which distractors caused the unusability of the clozes. Since they generate the distractors with three different methods, it remains unclear what is the method that effects usability the most.

The first method to generate a distractor is the baseline. It consists in ignoring the context tokens and return a token with the same part-of-speech with a frequency in a chosen corpus close to the frequency of the key. This method poses issues of usability since multiple valid answers are easily generated.

The second method, based on collocations, takes into account the context of the key. In a trigram $<A,p,B>$, p is the key (preposition), A is the previous token and B is the following. Considering only one adjacent token, either A or B, results in a more difficult distractor while taking into account both adjacent tokens results in lower usability, since the distractor could be correct in context. The method extracts trigrams where the preposition appears frequently with A or B but not with both. Lee and Seneff state that distractors in language learning application may be of different kind from distractors of clozes used as a proficiency assessment tool. While in learning application difficulty is stressed since “An easy cloze test, on which the user scores perfectly, would not be very educational; arguably, the user learns most when his/her mistake is corrected”, in proficiency assessment, difficulty is less crucial. On the contrary a less difficult cloze is needed to discriminate between proficient and less proficient students.

The last method proposed by Lee and Seneff [31] consists in harvesting the most frequent mistakes in a non-native corpus. This method requires correction of the erroneous sentences. So trigrams with prepositions are extracted from the corrected non-native corpus:

Corrected: He **studies at** the **university**

Original: He **studies in** the **university**

Target trigram: <study, at, university>

Extracted distractor: <study, in, university>

While it is a solid way to collect good quality distractors, large non-native corpora annotated with corrections are expensive to produce and they usually are restricted to speakers of specific languages, as the authors point out.

Some other methods differ in the aim of cloze generation. Malafeev [32] aims at reproducing open-end clozes similar to the ones used in the Cambridge certificate exams by using a static list of 146 target word forms. While their results are promising, the focus of the methods is in simulating the generation of the keys regardless of the possible answers. There’s no unique correct answer for the cloze, and the answers need to be manually evaluated. Usability for automatic grading is not taken into consideration.

As mentioned earlier, other methods for automatic generation of text-based clozes exist. However these methods are applicable to other types of cloze generation, such as reading comprehension, vocabulary checking (semantics) and factual knowledge testing, which are out of the scope of this work.

2.3 Revita, an ATICALL language learning platform

As a language learning platform Revita is placed at the intersection of ITS and ICALL [27], or, since it processes native language in authentic texts, ATICALL.

Authenticity is underlined by the Finnish National Agency for Education in their report about challenges emerged in language learning outcomes in Finnish schools [22]. In language learning *authenticity* usually refers to the teaching material's origins i.e., whether the material was created for teaching purposes (not authentic), to what extent it has been modified for teaching purposes and whether it was produced by native speakers. On the other hand authenticity requires also the learner's own agency. A student's learning activities are authentic when she experiences them as meaningful to herself and when she is able to take an active part in them [22]. Being able to choose an authentic text of own choosing enables the sense of agency of the student and makes learning activities personally meaningful. Some of the most relevant inadequacies identified in learning and teaching methods were the scarce use of technology, the lack of authentic material and language use, and the rarity of learning situations were the student's autonomy is fostered [22]. Revita addresses all these inadequacies. In addition, unlike similar automatic generation systems for grammar exercises from authentic text [10], it provides freedom in the selection of the material, making the learner an active agent in the learning process.

Revita can be accessed via browser at `https://revita.cs.helsinki.fi`. This version of the platform is the production side where the finalized and tested exercises are implemented. A separate version of the platform[2] is used for development purposes in order not to disrupt users' activity. The following is an overview of the process of exercise creation.

Revita can be used by registered or unregistered users. In the latter case no information on personal progress will be available. To create exercises to practice with, the user can choose one of the stories in the publicly available library or upload a text of their own choosing. This option is crucial to achieve authenticity not only in learning material but also in the sense of experiencing the material as personally meaningful, as previously stated by Hildén and Härmälä.

There are several options to add self-selected texts. The user can provide a web address such as a Wikipedia URL or any other address to extract text from. Alternatively the user can upload a text file, type in or copy-paste a text. The first line of the text is chosen as the title for later retrieval.

[2] https://mobvita.cs.helsinki.fi/home

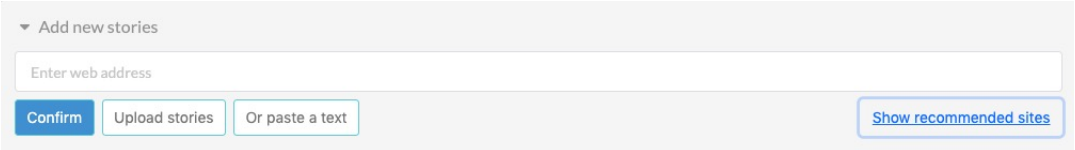
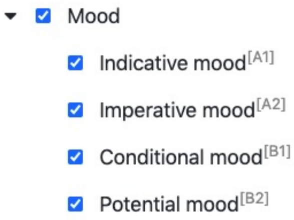


Figure 4: Revita allows the user to practice with a text of his choice.

Once a text has been provided or selected from the library, Revita processes it and automatically creates exercises. Input text is tokenized, analyzed for PoS- and morphological tags, run through a chunker and assigned features. Then single-tokens and chunks are mapped into *learning construct* candidates. *Learning constructs* are units identified to represent language knowledge. Some constructs are grouped by lexical categories. For instance a construct can represent knowledge of the conjugation of a specific verb type in a given tense. Other constructs are more complex and involve multiple lexical categories such as in *noun phrase agreement*. A construct can have different sub-constructs. For instance the construct *verb* has sub-constructs *mood* and *tense*. Therefore a student can, for example, practice the conjugation of verb for *Indicativo Passato Prossimo* where Indicativo refers to the verb mood and Passato Prossimo refers to the tense.

Depending on what constructs the user has activated, a text snippet with the cloze exercises are presented. Revita's constructs are ranked on six levels defined by the The Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) from A1 to C2, which can be regrouped into three broad levels: Basic User (A1, A2), Independent User (B1, B2) and Proficient User (C1, C2) [12]. The user can set a maximum level allowed for exercises or choose the constructs to practice:

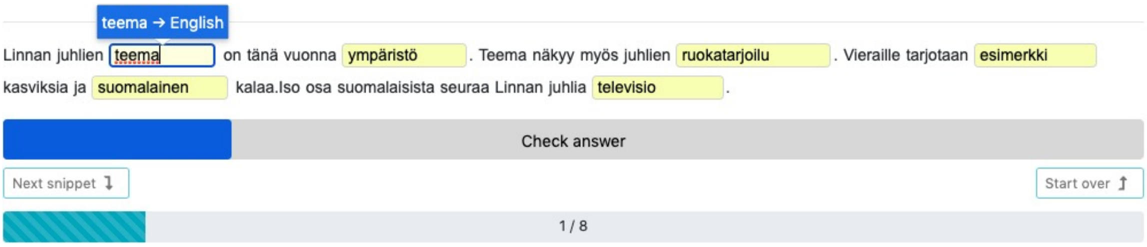


☑ Mood
- ☑ Indicative mood[A1]
- ☑ Imperative mood[A2]
- ☑ Conditional mood[B1]
- ☑ Potential mood[B2]

Figure 5: CEFR levels (A1, A2, B1, B2) for verb mood constructs in Finnish.

Revita's exercise types vary depending on the language and may consist of open-ended clozes, multiple choice clozes and clozes based on listening (dictation of tokens). Revita's clozes test language competencies such as inflection of words, agreement (in case, gender, number...) among different parts-of-speech belonging to the same phrase, verb conjugation, government and orthography.

In Revita, in the case of open-end clozes the user is given a hint, usually the base form or lemma of the word, to avoid multiple contextually acceptable answers. Unknown words can be clicked to be translated in one of the 17 available languages (Chinese, English, Finnish, French, German, Spanish, Norwegian, Russian, Swedish, Turkish, Japanese, Kazakh, Polish, Czech, Portuguese, Hindi).

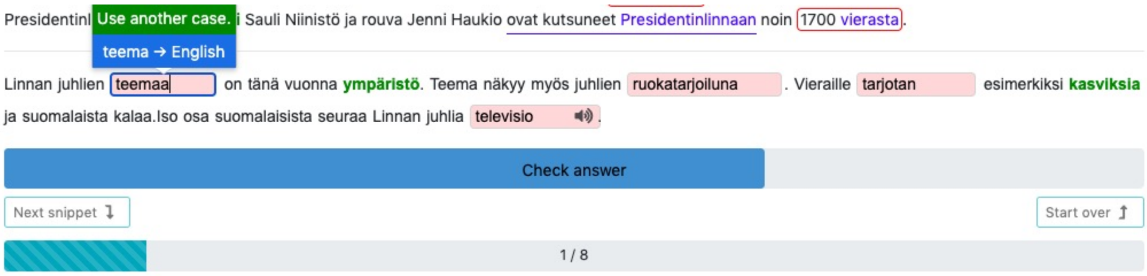


Linnan juhlien teema on tänä vuonna ympäristö . Teema näkyy myös juhlien ruokatarjoilu . Vieraille tarjotaan esimerkki kasviksia ja suomalainen kalaa.Iso osa suomalaisista seuraa Linnan juhlia televisio .

Figure 6: An example snippet from a text in Finnish with open-end clozes. Revita can also provide translations of unknown words.

The learner fills in the gaps and/or chooses one of the alternatives of the multiple-choice exercises. Once all the answers in the text snippet are correct the next text snippet is shown. When one or more answers are wrong i.e., they do not match with the original blanked single-tokens or MWE, the system may give feedback meaning that it gives a hint to produce the correct answers and the user can try to answer again. Feedback varies depending on the language, the type of exercise and the learning construct tested. In order to give meaningful feedback, the incorrect answer is run through the analyzer and the chunker to extract features, that are in turn compared to the features of the target answer. Based on this comparison, feedback is provided. So in an a verb tense exercise, if the user's answer has matching features for *mood*, *tense*, *voice* and *person* but a non-matching feature for *number*, specific feedback about *number* will be generated. Depending on the learning construct and on the type of wrong answer, feedback can be multi-level and hierarchical, shallow or, in case of non-recognized features, a general *try again* prompt. It may be direct such as *use abessive case* where the student is explicitly guided on what feature to use, or indirect such as *use another case* where the student is expected to infer herself the feature required in the given cloze.

A bare construct is usually represented by a single-token and feedback is generated using features extracted by the morphological analyzer.



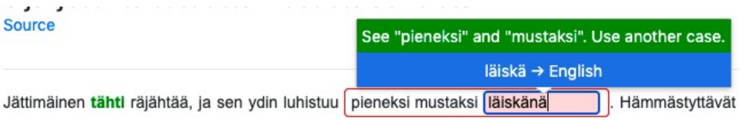
Presidentinl Use another case. i Sauli Niinistö ja rouva Jenni Haukio ovat kutsuneet Presidentinlinnaan noin 1700 vierasta.

teema → English

Linnan juhlien teemaa on tänä vuonna ympäristö. Teema näkyy myös juhlien ruokatarjoiluna . Vieraille tarjotan esimerkiksi kasviksia ja suomalaista kalaa.Iso osa suomalaisista seuraa Linnan juhlia televisio .

Figure 7: An example snippet of feedback on a bare construct in a text in Finnish.

Anchored constructs are, instead, chunks composed of multiple tokens where some features of the tokens must agree or have a specific type of relationship. For example in a noun phrase chunk, the adjective agrees in *case* with its head noun. If the blanked token is a noun, the student is suggested to infer correct features i.e., the correct *case*, from the adjective.



See "pieneksi" and "mustaksi". Use another case.

läiskä → English

Jättimäinen tähti räjähtää, ja sen ydin luhistuu pieneksi mustaksi läiskänä . Hämmästyttävät

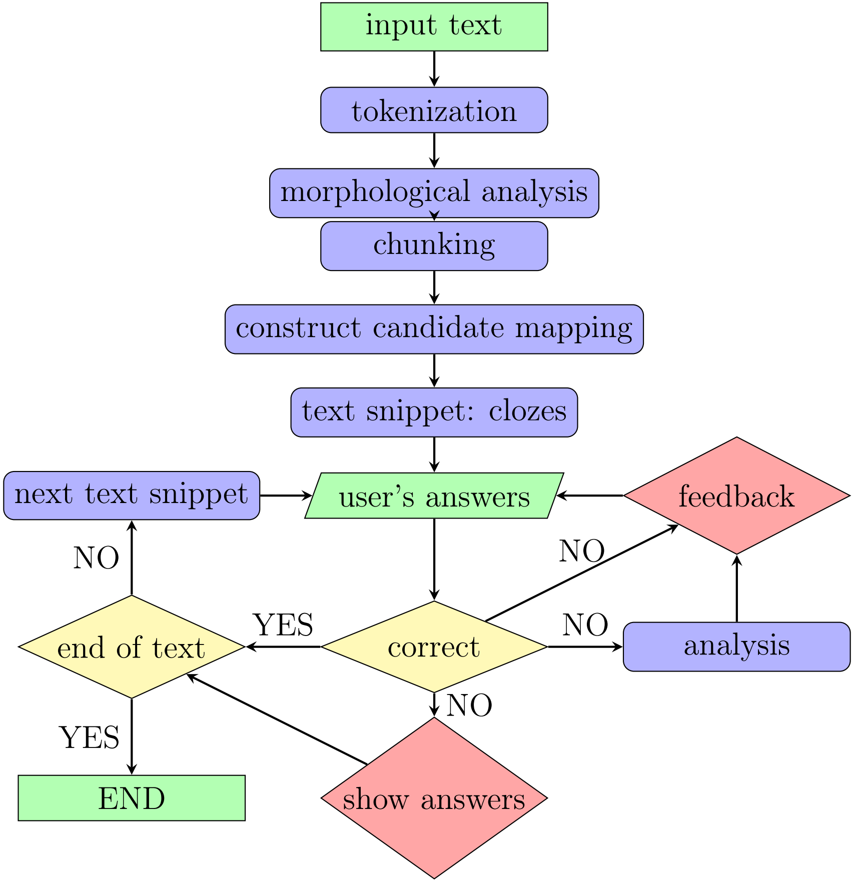
Figure 8: An example snippet of anchored feedback in a text in Finnish.

Revita also gathers words from the text to create out of context exercises such as crosswords and flashcards. Those types of exercises are out of the scope of this work.

Finally, Revita keeps track of learning progress and adapts future exercises to the student's skills. Exercises that are too easy or too difficult for the user are presented less frequently [28].

The following is a simplified visualization of the processes that occur from providing an input text to the conclusion of last text snippet with exercises. Many processes, like intelligent tutoring, are omitted because out of the scope of this work:

Figure 9: Process pipeline in Revita



The focus of this work is in chunking, mapping candidate constructs and feedback.

2.3.1 Language processing tools

Revita builds upon many tools and resources depending on the language. For Italian, at current, Apertium's morphological analyzer is used. Apertium[3] is a "free/open source framework for creating rule-based machine translation systems" [43]. It contains freely available tools such as finite-state morphologies, bilingual transfer lexica and probabilistic part-of-speech taggers.

Apertium's analyzers consist in finite-state transducers (FST), a "type of finite-state automata, which may be used as one-pass morphological analysers and generators and may be very efficiently implemented" [20]. Apertium's Italian morphological analyser delivers one or more lexical forms consisting of lemma, lexical category (noun, verb, preposition, etc.) and morphological inflection information (number, gender, person, tense...). The surface form *gatti* is analyzed as:

[3] https://github.com/apertium

(9) `gatti/gatto<n><m><pl>`

where *gatti* is the surface form to be analyzed, *gatto* is the lemma, *<n>* refers to the lexical category *noun*, *<m>* and *<pl>* to the morphological inflection information of gender (*masculine*) and number (*plural*). "In the case of contractions, the system reads a single surface form and gives as output a sequence of two or more lexical forms" [20].

The contracted form *delle* is analyzed as follows:

(10) `delle/un<det><ind><f><pl>/di<pr>+il<det><def><f><pl>.`

Apertium's analyzer is able to recognize some multiwords. "Lexical units made of more than one word (multiwords) are treated as single lexical forms and processed specifically according to its type." [20]. Although the Italian analyzer recognizes a few MWEs, for example some adverbial constructions such as *verso sud*, analytic verbs are not recognized. In Revita, at the moment, the analyzer is fed one token at a time and does not take into account MWEs.

The morphological analyzer recognizes 462 319 surface forms [43] and covers about 88% of a representative corpus (EuroParl).

To modify and augment the analyzer's output according to various needs, a wrapper is used. The wrapper also adds information not provided by the analyzer. For instance some verb moods, like the conditional mood have no tenses in the analyzer. This is mostly likely because the analyzer provides analyses of single-token verbs and it is therefore able to only give information on *Conditional Present* since in Italian *Conditional Past* is formed as a multi-word lexeme and the analyzer has no knowledge about it. Information on tense of single-tokens is still necessary to provide appropriate feedback to the user so the analyzer's information has been augmented through the wrapper.

Another type of information added by the wrapper is auxiliary of verbs. As illustrated in section 2.1 there are two main auxiliaries for Italian verbs, namely *essere* (be) and *avere* (have). The Italian auxiliary system is quite complex in the sense that choice of auxiliary is affected by intrinsic transitivity or intransitivity of the verb, transitive or intransitive use of the verb, voice (in case of transitive verbs) and finally by a, to some extent arbitrary, feature of the verb [14]. Nevertheless, knowledge of the right auxiliary of a verb is crucial for the correct conjugation of analytic verbal forms.

The auxiliary features of verbs were crawled from an online dictionary [35]. It is worth noting that absolute accuracy of information about verbs' auxiliaries is unlikely if not impossible to achieve, because of the complexity of Italian auxiliary verb system mentioned above.

Another challenge for the analyzer is ambiguity. In example (11) *delle* is ambiguous in terms of lemma. Possible lemmata are *un*[4] and *di+le*.

According to Forcada et al. "a sizeable fraction of surface forms (in Romance languages, for instance, around one out of every three words) are ambiguous, that is, they can be analysed into more than one lemma, more than one part-of-speech or have more than one inflection analysis".

The risk of imprecise lemma- and PoS-disambiguation in language learning is to teach a wrong or non-existing grammatical construct and to give erroneous feedback. It is fundamental not to misguide the learner and not to create an exercise on a wrongly assigned part-of-speech (or lemma).

Apertium has a part-of-speech tagger that "chooses, using a statistical model (hidden Markov model), one of the analyses of an ambiguous word according to its context" [20]. A *constraint grammar* reduces or removes PoS-ambiguity before the statistical PoS-tagger by applying a *forbid* rule that that removes two sequences (in the first-order models, these sequences can only include two parts of speech) of tags and an *enforce* rule that defines what tags are allowed after a specific tag. "These rules are applied to the HMM parameters by introducing quasi-zeroes in the state transition probabilities of forbidden sequences"[43]. Even state-of-the-art PoS-tagger (98.01% for Stanza, [38]) is not enough to be applicable to automatic cloze generation because of the risk of generating exercises with wrongly assigned parts-of-speech tags.

For this reason, only unambiguous surface forms are good candidates for single-token cloze exercises. Chunks are, instead, disambiguated by matching rules. A rule defines a sequence of features such as *PoS*, lemma, morphological tags and *surface* features. For instance a chunk like *non ho ancora mangiato* (I haven't eaten yet) is matched by a rule where four consecutive tokens have the following features:

token$_1$: surface is *non*

token$_2$: verb indicative present

token$_3$: adverb

token$_4$: verb past participle

In the chunk *non ho ancora mangiato* the third token *ancora* has four different possible readings returned by the morphological analyzer:

[4] formally *un* is not the lemma of a plural indefinite determiner. *Delle* is a partitive construction that *functions* as the plural of an indefinite determiner.

- ancora (yet), *adverb*
- ancora (anchor), *noun*, singular, feminine
- ancorare (to anchor), *verb*, indicative, third person, singular
- ancorare (to anchor), *verb*, imperative, second person, singular

Non ho ancora mangiato will match the rule and will therefore disambiguate the reading for token *ancora* as *adverb*. Failing in disambiguation may occur, although rarely. An evaluation of accuracy of disambiguation will be conducted in the future.

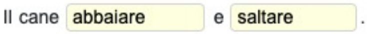
2.3.2 Italian *beta* version

Exercises in the *beta* version of Italian include open-end clozes for verbs, adjectives and pronouns where single-tokens are chosen as keys based on their part-of-speech.

Multiple-choice clozes are instead created for prepositions, conjunctions, adverbs and determiners. Distractors are generated by picking at random tokens in the input text with the same part-of-speech as the key.

While already these implementations offer a great deal of variety of exercises, further development is necessary to increase variety even more and to effect *usability* and *difficulty* of clozes. It must be noted that, in Revita, *usability* as depicted by Lee [31] (see 2.2) does not determine whether an exercise can be legitimately used or not, since the main purpose of the platform is to help students in learning languages. The main purpose of Revita is not evaluating language proficiency. For this reason *multiple admissibility* [25], that is, when a cloze can be resolved by multiple correct answers, can be seen, when properly addressed, as a further learning possibility for the student rather then a obstacle or a confusing element. It provides a broader view on a particular learning construct, enabling a deeper learning experience.

In open-end clozes *multiple admissibility* is addressed by providing feedback that further guides the student towards the correct answer in the specific context.



Il cane abbaiare e saltare .

Figure 10: An example of an open-end cloze on verbs where multiple correct answers are possible in context.

In figure 10 various moods and tenses are possible but only the target answer is considered correct. While the required mood and tense can be inferred in most cases by further context, feedback is necessary to guide the learner on the type of answer she is expected to give.

Introducing new and more specific learning constructs will contribute in providing a greater variety of exercises and in raising Revita from its *beta* status.

Next section illustrates the applications into Revita of the learning constructs defined in 2.1. Evaluation of the implementations of the learning constructs is also conducted to determine the next steps needed for further development. In the same section, an inventory of additional learning constructs are suggested to raise Italian from its *beta* status towards a fully developed language in Revita. Several grammars and linguistic resources were consulted to make an inventory of learning constructs for Italian. The main linguistic references are *Dardano*[14], *Accademia della Crusca* [13], *Istituto Treccani*'s grammar [42] and Imperato's grammar [23] and exercise book [24].
Learning constructs that are considered not feasible with the current tools are still taken into account for future work when additional tools will be implemented. For a full list of learning constructs see appendix D.

3 Learning constructs: applications to Italian

3.1 Analytic verbs

The learning constructs illustrated in this subsection correspond to constructs 1-19 in appendix D.

In the *beta* version of Revita only single-token verbs are used to generate clozes since the morphological analyzer only recognizes verbs in their simple one-token forms. Tokens that are unambiguously analyzed as verbs are selected as keys of open-end clozes.
Compound tenses (*tempi composti*) are multi-token (analytic verbs). Italian verb conjugation consists of 21 tenses for 7 different moods. In the active form of a verb conjugation, 9 of these tenses are analytic verbs. In the passive conjugation of a transitive verbs all tenses are in analytic forms. For a complete conjugation of verbs see appendices A and B.

To match chunks of multi-word verbs, rules were designed and chunks were assigned to learning constructs. Constructs, in this case analytic verb tenses, are assigned to a CEFR level. It is worth noting that assigning a specific level to a construct is, in some cases, problematic since the level of the exercises depends also on the difficulty of the text provided by the user. In general, the broader the tested linguistic construct is, the harder it is to accurately assess the level of it. It also must be kept in mind that CEFR levels measure language skills rather then specific grammatical constructs [11].

To assign the construct of *Passato-Prossimo* (present perfect) a rule will identify two consecutive tokens analyzed as *verb*, respectively a verb in Indicative Present *(Indicativo*

Presente) and a verb in Past Participle *(Participio Passato)*:

ho giocato = Passato Prossimo
[ho = Indicativo Presente] + [giocato = Participio Passato]

When the target verb is *ho giocato* (I played), the single-token form *ho* must not be assigned to the construct of Indicative Present, since in this context it is part of an analytic verb. Assigning the right construct is fundamental for exercise creation and to provide appropriate feedback. Blanking *ho* and asking the learner to use Indicative Present Tense would be erroneous and confusing for the learner.

To assign an analytic verb chunk to the correct learning construct, in this case to an active or passive tense, several features of the verb must be taken into account. As seen in 2.1 the Italian verb conjugation is complex not only because the great number of different moods and tenses but also because of the ambivalence of the two auxiliaries, *essere* and *avere.* Some verbs are constructed in their active analytic forms with *essere*, some with *avere*, others with either one or the other depending on transitivity or even both interchangeably. The passive analytic forms are always constructed with *essere.* This cause some ambiguity, i.e., same surface chunks for intransitive verbs with auxiliary *essere* and passive of transitive verbs (which always have auxiliary *avere*).

CAMBIARE (to change)				
	Presente	Passato-Prossimo	Imperfetto	Trapassato-Prossimo
TRANSITIVE ACTIVE	cambio	ho cambiato	cambiavo	avevo cambiato
INTRANSITIVE (ACTIVE)	cambio	**sono cambiato/a**	cambiavo	**ero cambiato/a**
TRANSITIVE PASSIVE	**sono cambiato/a**	sono stato/a cambiato/a	**ero cambiato/a**	ero stato/a cambiato/a

Table 1: Some examples of ambiguity between intransitive and passive of analytic verbs.

To assign the correct chunks to the learning constructs and resolve ambiguity, the auxiliary feature must be taken into account and separate rules are designed. A total of 35 rules cover for analytic verbs with the following features:

- Active voice of verbs with auxiliary *avere* as surface, including transitive (*ho amato*, I loved) and intransitive (*ho abbaiato*, I barked) verbs. Participle is in masculine singular. All analytic moods and tenses are covered:
  - Indicativo: Passato-Prossimo, Trapassato-Prossimo, Trapassato-Remoto, Futuro-Anteriore
  - Congiuntivo: Passato, Trapassato
  - Condizionale: Passato
  - Infinito: Passato
  - Gerundio: Passato

- Intransitive (active) verbs that can only take auxiliary *essere*, for instance *sono andato* (I went). Participle agrees with subject in gender and number. All analytic moods and tenses (as above).
- Passive voice of verbs that can only take auxiliary *avere*, as in *sono stato amato*. The following moods and tenses:
  - Indicativo: Presente, Passato-Prossimo, Imperfetto, Trapassato-Prossimo, Passato-Remoto, Trapassato-Remoto, Futuro-Semplice, Futuro-Anteriore
  - Congiuntivo: Presente, Passato, Imperfetto, Trapassato
  - Condizionale: Presente, Passato
  - Infinito: Presente, Passato
  - Gerundio: Presente, Passato
- Following unambiguous moods and tenses for verbs that can take both auxiliary *essere* and *avere*
  - Indicativo: Passato-Prossimo, Trapassato-Prossimo, Trapassato-Remoto, Futuro-Anteriore
  - Congiuntivo: Passato, Trapassato
  - Condizionale: Passato
  - Infinito: Passato
  - Gerundio: Passato

In addition to the above illustrated analytic verbs, 9 rules for chunks of pronominal verbs were designed. The need to implement this kind of chunks emerged while working at the implementation of another learning construct, namely governments of verbs. The auxiliary for this kind of verbs is *essere* even when the non-pronominal active form has auxiliary *avere*:

(11) *mi* *sono* *lavat-o*

I.REFL be.AUX.PRS.1SG wash.PTCP.PST-SG.M

'I washed myself'

(12) *ho* *lavat-o*

have.AUX.1SG wash.PTCP.PST-SG.M

'I washed'

In example (11) the pronominal reflexive verb *lavarsi* (to wash onself) requires auxiliary *essere*, while the non-pronominal verb *lavare* (to wash) requires auxiliary *avere* in its active inflection. The rules that identify chunks of pronominal verbs match sequences

where reciprocal and reflexive personal pronouns *mi, ti, ci, vi, si* agree in number and person with a following auxiliary (*essere*) and a past participle in the case of analytic verbs or a single verb the case of simple form verbs. This allows to catch pronominal verbs like in *mi allontano* (I am moving away), *ti chiami Piero* (your name is Piero), *si è accertata che fossero usciti tutti* (she made sure that everyone left).
These additional rules, that are separated from their non-pronominal counterparts, also allow to correctly assign passive forms of analytical verbs to their learning constructs:

(13) *sono* *lavat-o*
be.AUX.PRS.1SG wash.PTCP.PST-SG.M
'I am being washed'

Example (13) shows that the sequence of tokens in the passive non-pronominal analytical verb *sono lavato* (I am being washed) is contained in the sequence of tokens in the pronominal verb *mi sono lavato* illustrated in example (9). For this reason separate rules for pronominal verbs are needed.

The rules for pronominal verbs cover the following modes and tenses for analytic forms:

- Indicativo Passato-Prossimo, *mi sono lavato*
- Indicativo Trapassato-Prossimo, *mi ero lavato*
- Indicativo Trapassato-Remoto, *mi fui lavato*
- Indicativo Futuro-Anteriore, *mi sarò lavato*
- Congiuntivo Passato, *mi sia lavato*
- Congiuntivo Trapassato, *mi fossi lavato*
- Condizionale Passato, *mi sarei lavato*

In addition, a single rule catches sequences of a personal pronominal pronoun and a single-token verb, covering all the finite single-token modes and tenses: Indicativo Presente (*mi lavo*), Imperfetto (*mi lavavo*), Passato-Remoto (*mi lavai*), Futuro-Semplice (*mi laverò*), Congiuntivo Presente (*mi lavi*) and Imperfetto (*mi lavassi*), Condizionale Presente (*mi laverei*).

In the next subsection the excluded verbal forms that are not covered by the rules are presented. Cases of possible ambiguity are illustrated and alternative solutions to solve them are suggested.

3.1.1 Evaluation

The designed rules cover a great deal of analytic verb forms. Nevertheless coverage is not full. Some surfaces of auxiliaries are excluded because they present ambiguity that propagates into analytic verb forms. The excluded auxiliary surfaces and their ambiguous constructs are:

- *abbiamo* and *siamo*, 1st person plural, Indicativo Presente vs. Congiuntivo Presente
- *aveste* and *foste*, 2nd persone plural, Indicativo Passato-Remoto ambiguous with Congiuntivo Imperfetto

Verbs that can take both auxiliaries *essere* and *avere* are not covered in the following cases because of ambiguity between passive and intransitive surfaces:

- Active analytical verb forms with *essere* as auxiliary surface, as in *sono cambiato* (I changed), Indicativo Passato Prossimo.
- Passive voice as in *sono cambiato* (I'm being changed), Indicativo Presente.

For a complete illustration of ambiguity between intransitive verbs that take auxiliary *essere* and passive verb forms see appendix C.

For pronominal verbs the moods and tenses not covered are the cliticised forms that were not enabled at the moment of testing and that are still under development:

- Infinito Presente *lavarsi* and Passato, *essersi lavato*
- Gerundio Presente *lavandosi* and Passato, *essendomi lavato*

The excluded analytic and the pronominal verb forms that are not covered by the rules have some minor impact in recall. This mean that not all analytic verbs are caught by the rules and therefore exercises will not be created from all verbs or all moods and tenses. Recall is not crucial, since it only effect the number of fragments to be assigned as keys in clozes. It does not consist in a danger in the sense that a small decrease in recall will not cause to generate a wrong exercise or to present to the student misleading feedback. Precision is, instead, fundamental because it is not acceptable to generate and present to the student clozes with wrongly assigned learning constructs. For this reason, priority was given to precision over recall.

To evaluate rules for the patterns, a chunker test was used. Over a total of 47 rules for analytic and pronominal verbs, 177 sentences were evaluated. Test sentences were annotated with indices that indicate tokens that are expected to match and tokens that are not expected to match. This kind of testing supported debugging and further development of the rules.

total sentences	177
correct match	108
correct no_match	65
no rule found	4
wrong rule	0
wrong token	0
wrong choice	0
unexpected match	0

Table 2: Chunker test results for analytic and pronominal verbs.

Table 2 shows the results of the chunker test for analytic and pronominal verbs rules. A *correct match* represents a test sentence that correctly matches a specific rule at a given token index. A correct *no_ match* consists in a test sentence that correctly does not match a given rule. A *no rule found* is a test sentence that is supposed to match a rule but does not. A *wrong rule* is a test sentence that matches a different rule from the one that it is supposed to. A *wrong token* represents a sentence that correctly matches a pattern rule but with a different token index. *Wrong choice* consists in a sentence that matches multiple rules including the one it is tested for, but it is assigned the wrong rule and not the one it is tested for. Finally, *unexpected match* is a unexpectedly matched pattern.

(14) `{'test': 'ti chiami Piero', 'match': [(0, 2)]}`.

(15) `{'test': 'ti chiamo domani', 'nomatch': [(0, 2)]}`.

In example (14) the test sentence is supposed to match the rule starting at index 0 including 2 consecutive tokens (*ti chiami*). The rule catches chunks of single-token pronominal verbs. In example (15) the test sentence is not supposed to match the same rule since *ti chiamo* is a transitive non-pronominal verb with a direct object (*ti*).

Results from the chunker test were actively used to further refine and correct the rules. Table 3 below shows the final confusion matrix for analytic and pronominal verbs. On a total of 177 test sentences, the true positives (TP) i.e., correctly matched patterns, are 108. True negatives (TN) i.e., correctly not matched, are 65. False negatives (FN), i.e., patterns that were not matched when they should have, are 2. The non matched test sentences include clitics, that are not implemented at the time of writing and remain an area of further development. There are no false positives.

Actual values/predicted vales	Positive	Negative
Positive	108 (TP)	4 (FN)
Negative	0 (FP)	65 (TN)

Table 3: Confusion matrix for analytic and pronominal verbs.

In respect to the specific test sentences, precision of the rules for analytic and pronominal verbs is 100% while recall is 96.4%. However, it must be kept in mind that the test sentences were designed for development purposes and do not provide a reliable evaluation on exercise generation on authentic text. Quantitative evaluation of precision and recall of analytic and pronominal verbs on authentic text remains to be conducted in the future, for instance by manually annotating a representative test corpus for analytic and pronominal verb modes and tenses and comparing it with the features assigned by the chunker. This kind of quantitative evaluation is, however, humanly costly and perhaps not relevant in the near future since recall is not crucial and the qualitative estimated precision based on the categorical exclusion of known ambiguous forms is 100%.

Finally, possible ambiguity of passive analytic verbs with copula must be considered. As an example *era dipinto* (it was painted vs. it was being painted) can be either a passive analytic verb where *dipinto* is a verb (past participle) or it can be a copula where *dipinto* is an adjective. In Italian, participles can often function as adjectives. The only way to surely state that it is a verb and not an adjective, is to have an explicit agent in the sentence. Alternatively disambiguation is possible based on semantics, although in some cases the question remains unsolvable also by human judgement. This ambiguity is not considered an issue in implementation since the verbal moods and tenses, i.e., the learning constructs, are the same whether we consider *dipinto* a verb or an adjective. It is still worth noting that to avoid most cases of ambiguity it is possible to disable rules for chunks of passive Indicative Present, as most cases of ambiguity fall into Present tense as in *è dipinto* (it is painted vs. it is being painted).

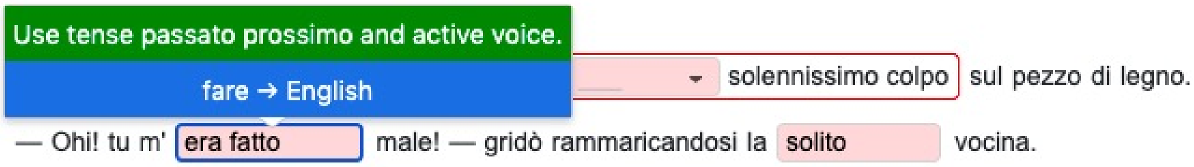
After chunks of analytic and pronominal verbs are matched, they are then assigned to learning constructs (moods and respective tenses) and open end-clozes are generated by blanking the whole analytic verb. In the next subsection implementation of feedback for verbs is illustrated.

Overall, the rules for the identification of analytic and pronominal verb chunks are functioning as expected and can be considered highly satisfactory in the purpose of creating new exercises and raising Italian from its *beta* status in Revita.

3.1.2 Feedback for verbs

When the user's answer to the cloze is wrong, i.e., when it does not match with the key, it is run through the analyzer and the chunker to get values for its features: mood, tense, voice, gender, person and number. Feedback is generated by comparing these features with the features of the key.

Feedback is implemented hierarchically. At the first level the user's answer is compared to the key by feature *mood*. If the feature's value is wrong, feedback on mood is presented, otherwise the comparison of features moves to the second level where values for features *tense* and *voice* are compared. If one or both values are incorrect, feedback is presented, otherwise comparison of features moves to the third level where values for feature *gender* is checked. In the fourth and last level, values for features *person* and *number* are compared.



Use tense passato prossimo and active voice.

fare → English

solennissimo colpo sul pezzo di legno.

— Ohi! tu m' era fatto male! — gridò rammaricandosi la solito vocina.

Figure 11: Feedback of level 2 for an analytic verb

Figure 11 shows feedback generated for a cloze on analytic verbs. The user provided *era fatto* as the answer. At level two of the feedback hierarchy, features *voice* and *tense* do not match with the target's values and direct feedback is generated. The user is explicitly told what values of the features to use i.e., tense *passato prossimo* and *active* voice.

Additionally, the feature *verb inflection* is used to correctly assign verbs to the pronominal or the non-pronominal categories, adding a level to feedback but excluding comparison of feature *voice* for pronominal verbs. Feedback on verb inflection is meant to provide a hint to the user on when to use a pronominal verb instead of a non-pronominal one. Feature values can be *Regular*, *Reflexive* and *Irregular*. The last value (*Irregular*) is not used for the time being but be will be useful in future implementation of irregular verb conjugations. Without the verb inflection feature, feedback was, in some cases, confusing for the user:

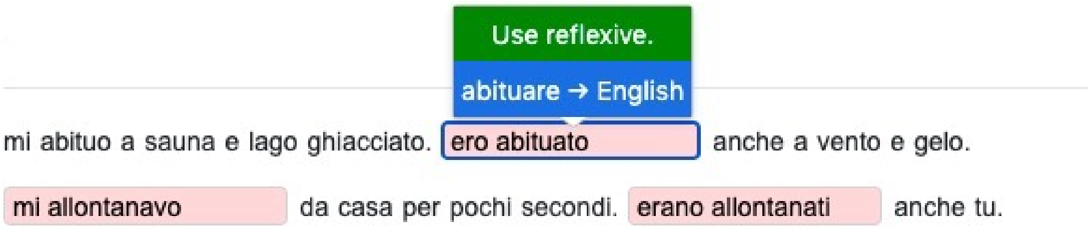
(16) TARGET:	*mi*	*ero*	*allontanat-o*
	I.REFL	be.AUX.PROG.1SG	move.PTCP.PST-SG.M
	'I moved away'		

(17) ANSWER: *ero* *allontanat-o*
be.AUX.PROG.1SG move.PTCP.PST-SG.M
'I was moved away'

Example (16) is the target answer, the fragment blanked from the text, *mi ero allontanato*. When the user supplied as an answer *ero allontanato* (example 17), feedback suggested to use *active voice* because the chunker correctly recognized *ero allontanato* as a passive analytic verb. This kind of feedback was confusing for the user because *voice* does not provide any information on the use of a pronominal or non-pronominal verb. When asked to use *active* voice the user would easily change the auxiliary to *avere*:

(18) ANSWER: *avevo* *allontanat-o*
have.AUX.PROG.1SG move.PTCP.PST-SG.M
'I moved away (something)'

In example (18) the user follows feedback and provides a new answer in *active voice.* The result is still wrong. The correct answer requires not to change the auxiliary but to add the pronominal particle *mi.* For this reason additional feedback on feature *verb inflection* was added.



Use reflexive.

abituare → English

mi abituo a sauna e lago ghiacciato. ero abituato anche a vento e gelo.

mi allontanavo da casa per pochi secondi. erano allontanati anche tu.

Figure 12: Feedback for a pronominal verb.

The hierarchy of feedback was altered so that feature *verb inflection* was checked before feature *voice* in order to not cause confusion in the user for the above illustrated reason. It is worth noting that the feature value *Reflexive* refers on a general level to pronominal verb forms and not only to proper reflexive. The term "reflexive" was chose to keep feedback metalanguage as accessible as possible from the user's point of view.

Figure 13: Levels of feedback for verbs at indicative mood

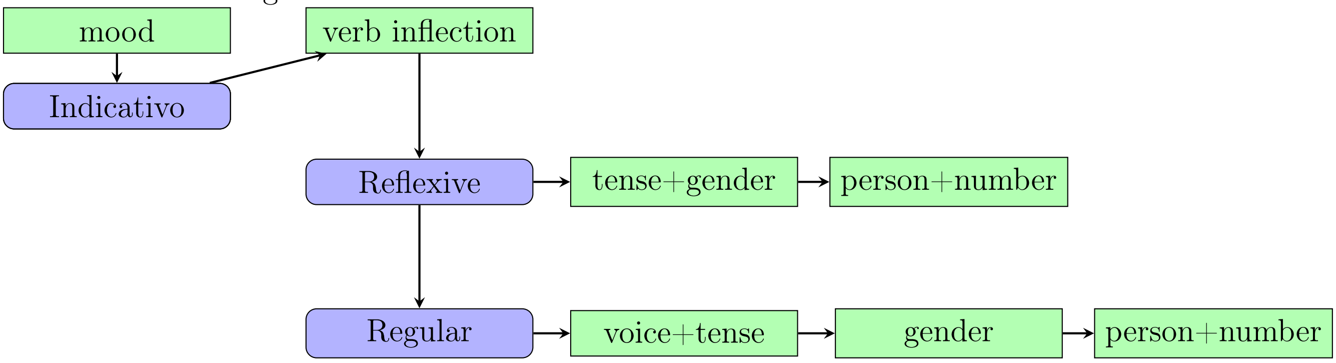


Figure 13 shows feedback for a user's provided answer that is analyzed as a *verb* in mood *Indicativo*. The value for feature *verb inflection* is compared to the target's value. In the case of non-matching values and if the target verb is a pronominal verb i.e., a verb accompanied by a specific personal pronoun, the user is asked to use a *reflexive*. Only then feature *tense, gender, person* and *number* are compared. In the case of a regular verb, i.e., a non-pronominal verb, also feature *voice* is compared and the user is possibly provided feedback on using *active* or *passive voice*.

Requiring meta-grammatical knowledge from the user has been criticized for example by Antonsen. While illustrating a system for learning Swahili he states that "the feedback put high demands upon the meta-grammatical knowledge of the users" [4]. This can be true for feedback on Italian moods and tenses, although it is assumed that language learners have this kind of meta-grammatical knowledge. A way to lower expectations on grammatical knowledge could consist in the reformulation of feedback. Instead of presenting the value of a feature as in "use *passato prossimo*" (present perfect), *simplified* feedback could be provided in form of an example with a similar verb. Simplified feedback could be an option that the user could separately activate.

Implemented feedback addresses user's errors that represent existing and recognized single-tokens and multi-words (chunks). Oller and John note that "native speakers rarely used completely ungrammatical responses, however, non-native speakers made this type of response fairly often" [36]. Feedback for non-words or invalid chunks remains an open question for future work.

3.2 Verb government of prepositions

As discussed in section 2.1 governee prepositions of Italian nouns, adjectives and verbs are some of the most challenging language learning constructs.

This subsection illustrates the learning construct of verb government with preposition and corresponds to construct 28 in appendix D. The key idea is to blank a preposition that is governed by a verb and to create a multiple-choice exercise with only one valid preposition.

Lee and Seneff [31] suggested two methods to find distractors for prepositions: non-native corpora and collocations. Using non-native corpora is problematic because of the bias of language learners that are native in different languages. Native language interferes with L2 in its own way and produces different kind of mistakes. In addition, learner corpora annotated with corrections are expensive to produce.

With the collocation method proposed by Lee and Seneff a maximum usability of 96,3% was achieved for English. It means that 3,7% of clozes were not usable because one of the three distractors resulted in a correct sentence. The authors claim that "among the unusable distractors, more than half are collocation distractors". A similar achievement for Italian is not guaranteed. Also, as pointed out in section 2.3, a usability of nearly 100% is the target for Revita. It is crucial not to create exercises that mislead the student.

A way to address the problem is to exercise only tokens that have a strict morpho-syntactic relation with their adjacent token, i.e., to consider as keys only colligations. A method to grant control of generated clozes and their usability could consist in explicit information of government of verbs, nouns and adjectives. While it can effect the number of generable clozes since it would be dependent on specific linguistic information, it should raise usability to nearly 100 percent.

Information on the most frequent colligations were gathered from grammars and other linguistic sources. The main references were Imperato's Italian grammar manual [23] and Treccani's grammar [42]. A total of 298 verbs that govern specific patterns were collected. A rule for patterns for each verb was designed. Some of the collected verbs are not recognized by the morphological analyzer at the moment but were left in the list to provide a possibility of usage in the future. For the same reason, also a few idiomatic high-frequency multi-word expressions such as *dare fastidio* (to annoy) were included in the list of patterns even if, at the moment, multi-token verbs of this kind are not implemented.

A typical rule to catch a verb government pattern states that a given verb requires either a specific preposition, a conjunction or no particle followed by either a verb, a noun phrase or a pronoun. The following example illustrates a rule that for a government pattern for the verb *piacere* (to like):

(19) `piacere, POS:Preposition+Base:a&-1 / POS:Pronoun&-1`

In example (19) the verb *piacere* (to like), governs a chunk (indicated by *POS)* i.e., a specific sequence of tokens, that consists in the preposition *a* and its components, at position *-1* meaning that the chunk precedes the verb. Chunks are defined by separate agreement rules.

An alternative pattern is indicated by a slash symbol (/). In the alternative case the verb *piacere* governs a token that has *Pronoun* as feature and that is at position *-1* meaning that it precedes the verb.

The following sentence includes the verb *piacere* and its governee *a moltissimi lettori* at preceding position:

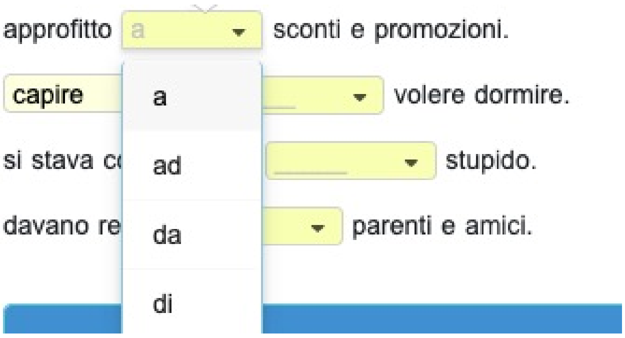
(20) *a moltissim-i lettor-i piace ancora di più*
to.DAT many-PL.M reader-PL.M like.PRS.3SG even of more
'many readers like it even more'

The designed verb government rules allow to create exercises where the user is asked to choose the correct particle, or the absence of it, in a multiple-choice cloze. Here the focus is on exercises of verb government with prepositions. Nevertheless, verb government rules allow to create also other kind of exercises such as multiple-choice exercises where the user is asked to choose between verb modes. Some verbs indeed govern a specific verb mood. More specifically some verbs require in the declarative clause *subordinata completiva* the verb mood *Congiuntivo* preceded by the conjunction *che.* In an example sentence like *credo che sia arrivato* (I think it came), the governor verb *credere* requires the mood in the declarative clause to be *Congiuntivo.* This information is included in verb government rules and thus allow to create multiple-choice exercises where the user is asked to choose between two verb moods, namely *Congiuntivo* and *Indicativo* as in the following examples:

(21) TARGET: *che sia arrivat-o*
that be.SBJV.3SG come.PTCP.PST-M.SG
'that it came'

(22) INCORRECT ANSWER: *che è arrivat-o*
that be.IND.3SG come.PTCP.PST-M.SG
'that it came'

The implementations of this kind of exercise by exploiting the designed verb government rules is left to be done in the future. Here, instead, the government rules are utilized to create multiple-choice exercises on prepositions. More precisely, the user is asked to choose a preposition, or the absence of it, between four choices where only one of them is the correct one.



approfitto a sconti e promozioni.

capire volere dormire.

si stava c stupido.

davano re parenti e amici.

Figure 14: A multiple-choice exercise about verb government. The user must choose the correct preposition.

In figure 14 the user must choose the correct preposition between four choices, *a, ad, da, di.* The key i.e., the preposition that appears in the original text and therefore the correct answer, is *di.* In the sentence *approfitto di sconti e promozioni* (I take advantage of discounts and cuts) only one of the four presented choices is valid since verb government patterns aim at allowing to pick only verbs that have a unique syntactic dependency relation with their governees. Because only one valid answer is supposed to be allowed, the generation of distractors can be conducted by randomly choosing from a designed list of distractors. More specifically the distractors for prepositions used for the generation of verb government exercises are:

(23) `["di", "a", "da", "in", "con", "per", "su", " "]`

Italian proper simple prepositions include also *tra* and *fra* but these two prepositions were left out because they are not governed by any verb taken into consideration and therefore would have only increased the facility index of the generated exercised. Instead, an empty string, i.e., the absence of preposition, was introduced among the distractors since verbs can also govern other verbs or chunk directly without prepositions.

In the next subsection an evaluation on the generation of exercises on verb government is exposed. In particular, an evaluation on usability of the generated exercises on verb government will be conducted and some of the limitations and areas of future improvement will be illustrated.

3.2.1 Evaluation

The generation of exercises on verb government with preposition can be considered successful based on the example sentences used during development. Nevertheless, the design of the rules for each verb that is considered a governor, presented some limitations. First of all, the patterns identified by the governments rules can only include unambiguous target verbs. This means that if the governor verb occurs is the text in an ambiguous form, the pattern is discarded and no exercise can be generated.

(24)	*non*	*accetta*	*di*	*portare*	*la*	*maglia*
	not.NEG	consent.IND.PRS.3SG	of	wear.INF	the.F.SG	shirt.F.SG

‘he/she does not consent to wear the shirt’

In example (24) *accetta* is the target verb for a government pattern where the verb *accettare* governs an infinitive preceded by the preposition *di*. The pattern is discarded because the surface *accetta* is ambiguous. It can be a verb in indicative mood, present tense, third singular person with lemma *accettare* (to accept) but also a noun with lemma *accetta* (hachet), presenting ambiguity in lemma. Additionaly, *accetta* can be a verb with lemma *accettare* (to accept) in mood *Imperativo* (imperative), second person singular, posing additional ambiguity in morphological tags regarding *mood* and *person*.

As mentioned in 2.3.1, roughly one third of surface forms in Romance languages are ambiguous. This means that many occurrences in an authentic text will be discarded because of ambiguity of the target verb.

To test verb government rules and the generation of multiple-choice exercises with simple preposition, a corpus of 29074 words in total was collected. The corpus consists of three different text types. The purpose of collecting different text types was to test whether generation of verb government exercises is more frequent in a particular type of text and to identify the type of text that allow the greatest deal of generated exercises.

The texts investigated are:

1. biography sections from Italian Wikipedia articles

2. Italian news articles

3. Italian language matriculation exams

More specifically the biography sections were extracted from Italian Wikipedia articles about late public figures Stephen Hawking and Raffaella Carrà. This type of texts was predicted to allow the generation of greatest deal of exercises on verb government. In Italian, biography style of deceased people consists in verb use at mood *Indicativo*, tense *Passato Remoto* (preterite). This specific tense presents minimal ambiguity in lemma and with other moods and tenses, allowing to match verb government patterns. Tense *Presente* (present) of *Indicativo* (indicative) mood instead, as seen in 3.1.1, is subject to ambiguity with mood *Imperativo* and *Congiuntivo* and with other parts-of-speech.

The news texts were collected from popular Italian online newspapers such as *Il messaggero*[5], *Corriere della Sera*[6] and *Metro*[7]. The style of news writing varies from using present tense, also as historic present *presente storico*, to preterite *Passato Remoto* and other tenses. The estimate for this kind of text was a moderate deal of generated exercises for verb government.

The texts extracted from the Finnish Matriculation Examinations consist in different types of text, from news articles, dialogues, interviews, to fictional stories. The Finnish matriculation exam is a test organized twice a year, at the end of high school where Italian is one of the optional exams for foreign languages. Past exams are publicly available online for practice[8]. The texts used in the matriculation exams are most often modified to some degree from authentic texts in order to adjust to the CEFR level A2. Texts collected for the corpus were extracted from exams ranging from autumn 2018 to spring 2021. Some minor alterations to the texts were conducted to minimize noise in the text. For instance, in the case of dialogues, the name of the speakers were omitted. The estimate for this type of text was a minimal number of generated exercises for verb government because of a low occurrence of tense *Passato Remoto* (preterite) and high frequency of ambiguous verb forms.

The collected text corpus was used to generate exercises on verb government. Results on exercises generation are illustrated below:

[5] ilmessaggero.it

[6] https://www.corriere.it/

[7] https://metronews.it/

[8] https://yle.fi/aihe/abitreenit/italia

text type	word count	exercises created	rate
Wikipedia, biography	9954	4	0.0004
news	9483	9	0.0009
matriculation exam	9637	18	0.0019

Table 4: Number of exercises for verb government generated on different text type.

As shown in table 4, the three different types of text have roughly the same number of words, ranging from 9483 to 9954 tokens. The number of verb government exercises generated ranges from only 4 in the biography type, to 18 in the matriculation exams. The news type text generated 9 exercises on verb government. These results diverge from the hypotheses made above. The biography text was estimated to generate the greatest number of exercises while in fact it generated the least (4). The matriculation exam text was predicted to generate the least number of exercises on verb government. Instead, it generated the greatest number of exercises (19). The divergence is likely explained by the curated quality of texts in the matriculation exams: the texts are either created from scratch for testing purposes or heavily modified authentic texts. They are designed to be understandable for language learners of level A2 and therefore their vocabulary matches the most frequent verbs implemented for the generation of verb government exercises. The higher rate of generated exercises in semi-authentic text suggests that frequency of vocabulary and complexity of sentence structure, have a relevant impact on automatic creation of exercises, at least for the verb government exercises with prepositions. This, in turn, suggests that assessing an authentic text's level, affects the number of exercises generated. From a pedagogical point of view Kitao and Kamiya confirm that, when generating cloze exercises from authentic text, it is crucial to choose "a text that is appropriate for students. The content should not be too difficult, and the text should not contain too many technical terms". Assessing a grammatical construct's level according to the CEFR a priori regardless the type of text is problematic. Addressing automatic rating of text level could be an area of future work.

It must be noted that generation of exercises for verb government remains relatively low also for texts extracted from the matriculation exams because of ambiguity of target verbs. Ambiguity is a non-trivial issue in the generation of exercises for Italian. Disambiguation of ambiguous lemmata, parts-of-speech and morphological tags remains an area of future work.

Another aspect that contributes to the low number of generated exercises is the lack of implementation of articulated prepositions (*preposizioni articolate*). These type of contracted forms are, at the moment, not implemented in the creation of government clozes for Italian.

(25)

Il	*compito*	*consisteva*	*nel*	*curare*	*le*
the.ART.SG.M	task.SG.M	consist.PROG.3SG	in.ART.M.SG	take_care.INF	the.ART.F.PL
piante					
plant.F.PL					

'The task consisted in taking care of the plants'

In example (25) the verb *consistere* (to consist) governs preposition *in* followed by a verb in infinitive mood. In this case *nel* is a contracted form of *in* + *il* where *in* is a preposition and *il* is a definite article. This patter is discarded at the moment because contracted forms are not yet implemented. The challenge in contracted forms consists in single surface forms that can yield multiple POS-tags. Implementation of contracted forms in future work will increase the number of generated exercises, allowing to match government pattern like the one shown in example (25).

To measure *usability* of the generated exercises on verb government, the multiple-choice clozes created on the collected corpus were evaluated by four native speakers. All the evaluators were either teachers of Italian as L2 or linguists. They were asked to evaluate a total of 31 sentences where a preposition was blanked and a multiple-choice exercise on verb government was generated. Specifically, they were ask to "choose the preposition that is correct in context. If there is more than one valid preposition, please choose all the valid options". Figure 15 shows that all four evaluators chose only one valid option in the multiple-choice exercise generated from the sentence:

> Molte famiglie italiane partecipano [con, -, su, a] questa iniziativa e mettono a disposizione la loro casa per ospitare i bambini.

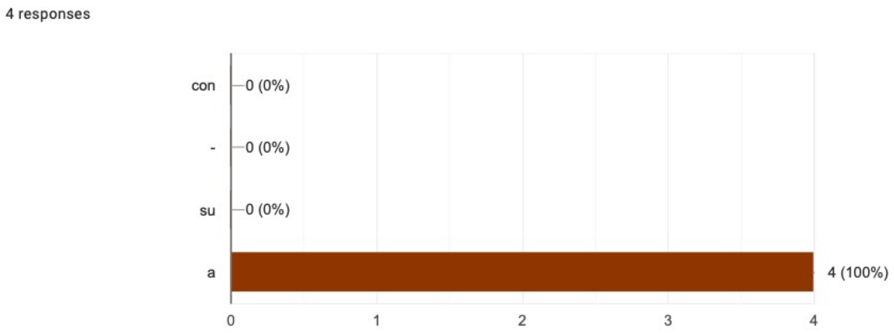


Figure 15: Full inter-annotator agreement on evaluation of a multiple-choice cloze exercise.

In two cases, annotators disagreed on evaluating whether there is only one valid preposition in context. The following sentence caused inter-annotator disagreement:

> Sono professore di educazione ambientale all'università di Padova, ma da un anno sono impegnato [per, con, in, a] un progetto con un centro di ricerca vicino a Roma, quindi viaggio tutte le settimane e passo moltissimo tempo sul treno.

Figure 16 shows that while all the evaluators chose the target preposition i.e., the preposition (*in*) blanked from the original text, two annotators considered valid also the preposition *con* and one picked also *a*.

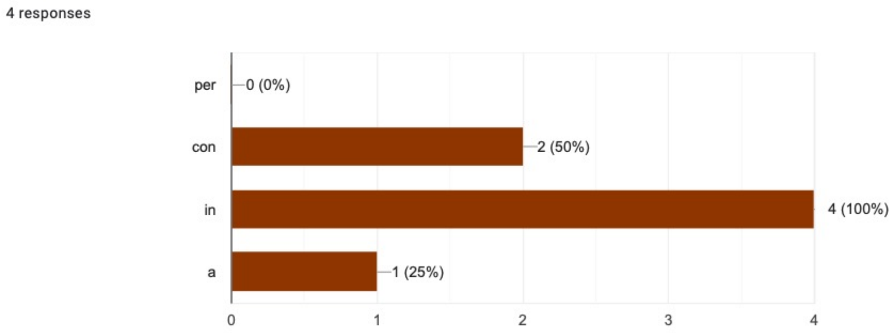


Figure 16: Disagreement on evaluation of a multiple-choice cloze exercise.

Krippendorff's alpha coefficient resulted in 0.96, showing an overall highly reliable inter-annotator agreement. On a total of 31 items, all four evaluators agreed on 29 instances by choosing only one valid preposition. The chosen preposition always corresponded with the key preposition i.e., with the original preposition blanked from the text. *Usability* resulted in 93.55% showing that the implemented patterns for verb government exercises generate — in a selected small-scale corpus — a substantial deal of exercises where there is only valid answer. The usability result achieved is slightly lower than the one obtained by Lee and Seneff (96.3%). They used a combined method of a baseline, collocations and learner corpora exploitation to generate distractors as seen in 2.2. It is worth noting that the method used in this thesis differs in the sense that it does not focus on the generation of distractors per se. It aims at identifying target verbs that govern a specific pattern and in particular a specific preposition. The evaluation conducted here pointed out that especially the verb *collaborare* (to collaborate) is problematic since two annotators evaluated the generated exercise with more than one valid answer. Excluding the verb *collaborare* from the target verbs for generating verb government exercises is an adjustment to take into consideration. It will notably increase usability with respect to the collected corpus. As noted in 2.3.2 *usability* in Revita does not imply whether a generated exercises should be used or not since the purpose of Revita is to support language learning rather then testing language competence. Nevertheless usability indicates where an additional level of feedback is necessary to instruct the user about multiple admissibility. Dealing with multiple valid answer remains an area of future investigation as well as measuring facility index. Implementation of articulated prepositions will allow a more sizable generation of exercises for verb government from authentic texts and the evaluation of facility index by testing multiple-choice clozes on actual users.

3.2.2 Feedback for prepositions

Feedback about prepositions is provided to the user by suggesting to pay attention to the governor verb and by pointing it out. It is worth noting that in future implementations of articulated prepositions (*nel, del...*) feedback needs also to address the governee chunk since in case of a governee consisting in a noun phrase, the articulated preposition must agree in gender and number with the head of the chunk.

(26)

mi	*sono*	*abituat-a*	*alla*	*sauna*
I.REFL	be.IND.PRS.1SG	use.PTCP.PST-F.SG	to.ART.F.SG	sauna.F.SG

'I got used to sauna'

In example (26) the articulated preposition *alla* is dependent on its governor verb *mi sono abituata*. It must also agree in gender and number with the noun *sauna*. In this case it must be *feminine singular*. For this reason feedback must take into consideration not only the governor verb but also features of the governee chunk.

Also, the token immediately after the articulated preposition effects the choice of the articulated preposition for phonetic reasons, behaving in the same way as articles do. A token beginning with a specific clusters of consonants requires a specific articulated preposition before it:

(27)

Il	*lavoro*	*consisteva*	*nel*	*curare*	*le*	*piante*
the.ART.SG.M	job.SG.M	consist.PROG.3SG	in.ART.M.SG	take_care.INF	the.ART.F.PL	plant.F.PL

'The job consisted in taking care of the plants'

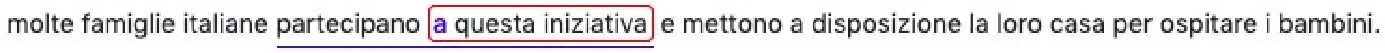
(28)

Il	*lavoro*	*consisteva*	*nello*	*scrivere*	*notizie*
the.ART.SG.M	job.SG.M	consist.PROG.3SG	in.ART.M.SG	take_care.INF	the.ART.F.PL

'The job consisted in writing news'

Example (27) shows that before the token *curare*, the articulated preposition needed is *nel* while in example (28), with the same governor verb, the articulated preposition *nello* is used. This is explained by the consonant cluster *sc* in *scrivere* that requires the article *lo* instead of *il*. For these reasons further levels of feedback need to be add in future implementations of articulated prepositions.

At the moment, only exercises with proper simple prepositions can be generated and feedback addresses the governor verb since there is no need to address the governee chunk with simple prepositions. The governee chuck is nevertheless highlighted in the last phase of feedback, giving the user already a valuable hint that can be used for implementation of articulated prepositions in the future.



molte famiglie italiane partecipano a questa iniziativa e mettono a disposizione la loro casa per ospitare i bambini.

Figure 17: Feedback on verb government.

Figure 17 shows an example of verb government with the last phase of feedback: the whole government pattern is underlined in blue and the governee chunk (*a questa iniziativa)* is circled in red.

3.3 Noun phrase agreement

The learning construct illustrated in this subsection correspond to constructs 25 in appendix D. The noun phrase agreement construct tests knowledge on agreement between the head of the noun phrase and its dependents. Inflection paradigm of nouns, adjectives and pronouns (*number* and *gender*) are prerequisite.

Five rules were designed to catch different types of noun phrase chunks. A question mark after a given part-of-speech indicates an optional token:

1. Determiner Adjective? Adjective? Noun Adjective?
2. Pronoun Adjective Noun
3. Pronoun Noun Adjective?
4. Noun Adjective
5. Adjective Noun

Rule number 1. catches patterns of a determiner, two optional adjectives, a noun and an optional adjective. Such a pattern is matched, for example, by the following token sequence:

(29) *la* *mia* *bella* *casa*
the.ART.F.SG my.ADJ.F.SG beautiful.ADJ.F.SG home.F.SG
'my beautiful home'

or:

(30) *una* *mia* *giacca* *vecchia*
a.ART.F.SG my.ADJ.F.SG old.ADJ.F.SG coat.F.SG
'an old coat of mine'

Rule number 2. catches patterns of a pronoun, an adjective and a noun, as in *queste*[9] *giovani donne* (these young women). Rule number 3. catches a pronoun, a noun and an optional adjective, *questo percorso panoramico* (this scenic route). Rule 4. and 5. catch respectively a noun and an adjective and vice versa. The rules are applied so that the longest matched pattern results in a chunk. Sequence of tokens that match multiple rules of the same length are discarded to avoid ambiguous patterns.

The exercises created from the noun phrase agreement rules are clozes that address either articles, adjectives or pronouns. In other words when a pattern is matched by a rule, a cloze exercise is generated by blanking randomly one of the allowed tokens in the pattern.

3.3.1 Evaluation

A chunker test was used to evaluate the rules with test sentences. Over a total of 5 rules for noun phrase agreement, 34 sentences were tested. Some cases of ambiguity emerged during designing and testing of the rules. Tokens that caused ambiguous patterns were excluded to avoid generating erroneous exercises. For example *anche* (hips) as a noun was excluded in rule number 5. (adjective noun) because of its ambiguity in part-of-speech. Indeed *anche* (also) can be also an adverb or conjunction and resulted in ambiguous chunks when preceded by a token that is ambiguous too. For example the sequence of tokens *presenti anche* was matched by rule 5. that looks for a pattern of adjective and noun. Since *presenti* (present) can be an adjective and *anche* (hips) can be a noun, the sequence was matched by the rule. Nevertheless most of the time *anche* (also) occurs

[9] *queste* is, in fact, a demonstrative *adjective* in this particular case. The morphological analyzer refer to all demonstrative adjectives as *pronouns*.

to be an adverb or conjunction, as it did in the test sequence. For this reason it was excluded from the possible surfaces for the given rule. Excluded tokens emerged during development and do not eliminate ambiguous chunks systematically. A consistent evaluation of potentially ambiguous chunks is left to be performed in future, for example by confronting an annotated corpus against the rules.

total sentences	34
correct match	24
correct no_match	8
no rule found	1
wrong rule	1
wrong token	0
wrong choice	0
unexpected match	0

Table 5: Chunker test results for noun phrase agreement rules.

Table 5 above and 6, below, show that over the 34 tested sentences, correct matches (true positives) were 24, correct no matches (true negatives) 8. One sentence did not match any rule (false negative). Investigation of this error indicated that the pattern is not matched because the analyzer is not able to identify the elided adjective *grand'* in the test sentence *un grand'uomo* (a great man). Dealing with elided tokens is an area of future improvement. Finally, one sentence matched the wrong rule (false positive). Inspection of this kind of error revealed that the test sentence *questo lungo percorso* (this long journey) is matched by a shorter rule that catches *lungo percorso* leaving the demonstrative adjective[10] out of the pattern. This kind of error does not actually affect precision since the head of the chunk is still identified correctly, even if by another, shorter rule.

Actual values/predicted vales	Positive	Negative
Positive	24 (TP)	1 (FN)
Negative	1 (FP)	8 (TN)

Table 6: Confusion matrix for noun phrase agreement rules.

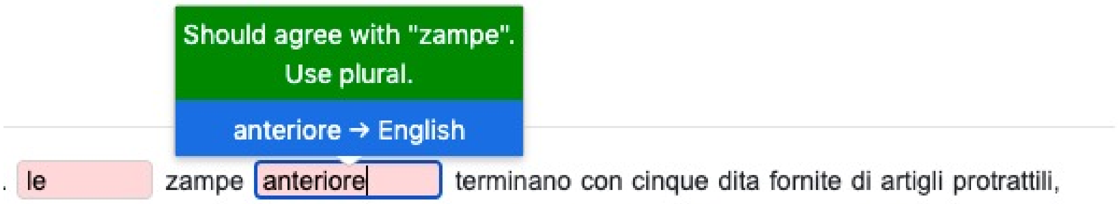
On the designed test sentences, both precision and recall of the rules for noun phrase agreement is 96.0%. Since the test sentences are modest in quantity, a slight fluctuation in test result affects massively precision. As pointed out in 3.1.1 the test sentences were designed for development purposes and do not provide a reliable evaluation on exercises generation on authentic text. Quantitative evaluation of precision and recall of noun

[10] *pronoun* in the system

phrase agreement on authentic text remains to be conducted in the future.

3.3.2 Feedback for NP

Feedback for noun phrase agreement is designed to point out to the user the word that the blanked token is supposed to agree with i.e., the head of the chunk. Since the blanked token can be a determiner, an adjective or a pronoun, also direct feedback on a possibly wrong value for *number* and *gender* is provided.



Should agree with "zampe". Use plural.

anteriore → English

. le zampe anteriore terminano con cinque dita fornite di artigli protrattili,

Figure 18: Feedback for noun phrase agreement.

In figure 18 the target chunk for noun phrase agreement is *le zampe anteriori* (the front paws). Two open-end clozes were created form the chunk, one by blanking the article *le* (the) and the other by blanking the adjective *anteriori* (front, in plural form). The user's answer to the second cloze is wrong since he typed *anteriore* (front) in singular form. The value for feature *number* does not match with the value of the target answer and feedback is generated. Feedback also suggests that the token in question should agree with its head, *zampe* (paws) in this case, providing a deeper kind of guidance that improve the learning experience.

3.4 Other constructs

In this subsection more learning constructs to be implemented in the future are suggested. While it is not an exhaustive list of constructs, see list see appendix D for a full list.

3.4.1 Auxiliary

One of the most challenging learning constructs of Italian is the choice of the auxiliary in analytic verbs. This learning construct is partially covered by the analytic and pronominal verb constructs (see 3.1). In particular, transitive verbs generally use the auxiliary *avere* in their active voice and *essere* in the passive voice. These verbs are already covered by analytic and pronominal verb constructs with respective feedback on voice. A separate construct that allows multiple-choice exercise generation for auxiliary is suggested for a limited number of intransitive verbs, i.e., for verbs that can have only one possible auxiliary and voice and therefore do not present multiple-admissibility. In the example *[sono, ho] riuscito* (I managed), only auxiliary *sono* is valid.

3.4.2 Conjugation

Italian verb conjugation paradigms consist in three regular classes. The end of the infinitive form of the verb determines the class: *giocare* belongs to the first class (-ARE) (to play), *correre* (to run) belong to the second (-ERE) and *dormire* (to sleep) belongs to the third (-IRE). In addition, a considerable number of verbs belong to the irregular class where inflection do not follow any specific pattern. While the conjugation of all classes of verbs is intrinsically included in the analytic and pronominal verb constructs illustrated in 3.1, a separate construct for verb class is suggested to allow the user to practice a specific verb class only.

3.4.3 Difficult gender and number of nouns

The learning construct of noun phrase agreement (3.3) allows to generate exercises for articles, adjectives and pronouns based on their chunk head. A separate learning construct is nevertheless suggested to allow only practicing of specific, demanding classes of words such as masculine nouns ending in -*a* and feminine noun ending in -*o*. A separate exercise on difficult gender of nouns could be created not by targeting nouns themselves but by targeting adjectives in the same chunk instead. So in the chunk *un programma costoso* (an expensive software) the blanked token would be the adjective *costoso* (expensive) and not the noun *programma* (software), requiring the user to first recognized gender (masculine) and number (singular) of the noun, and then to inflect the adjective accordingly. Nouns whose gender and number cannot be revealed by the ending of the token itself are not many in number but some of them are rather frequent and cause even highly proficient speakers to make errors. Examples of feminine nouns ending in -*o* are: *auto, moto, pallacanestro, pallavolo, radio, mano, libido, metro, foto...*. Examples of masculine nouns ending in -*a* are: *eremita, monarca, pirata, profeta, sosia, pilota, papa, pianeta, sistema, trauma, schema, poeta, pigiama, pianeta, parassita, gorgonzola, enigma, lemma, cinema, problema, clima, panorama, programma, fantasma...*. Extraction from the analyzer of all the nouns that have a gender or number surface that contradict the general rule, is suggested.

3.4.4 Verb agreement with preceding pronoun

Past participle of analytic verbs in their active voice is usually in masculine gender and singular number as in *hanno mangiato* (they ate) where the past participle *mangiato* is masculine singular even if the analytic verb in its whole has *plural* as value of feature *number*. Past particle, instead, agrees in gender and number with the personal pronoun that functions as direct object preceding the analytic verb. In *li hai visti* (you saw them) *li* is a pronoun that functions as direct object and precedes the analytic verb. The participle *visti* agrees in gender (plural) and number (masculine) with the pronoun. This learning construct could allow to generate either open-end clozes similar to the analytic and pronominal verbs for mood and tenses or multiple-choice clozes with the key participle and three participle distractors inflected in incorrect gender and number as in *li hai [visto, vista, viste,* ***visti****]*. This learning construct appears as Agreement-PronParticle in appendix D as construct 26.

3.4.5 Government

As for verb government (3.2), also nouns and adjectives can govern specific prepositions or other type of tokens. In the examples *allergico a* (allergic to), *curioso di* (curious about) and *diverso da* (different from), the adjectives govern the following prepositions, making them colligations rather than simple collocations. The same apply for the nouns in the examples *paura di* (fear of), *coraggio di* (dare to) and *fretta di* (hurry to), where the co-occurrence of the noun and the preposition is explained by a grammatical constraint. These learning constructs are listed in appendix D 29-31.

4 Conclusions

The purpose of this thesis was threefold: to contribute in the raising of Italian in Revita from its *beta* status to a full development stage, to formulate best practices for defining support for a new language in the Revita Framework and to serve as a documentation of what has been done, how and and what remains to be done.

In section 2.1 the most important Italian morphosyntactic constructs for language learning were identified, addressing Research Question 1. Learning constructs for analytic and pronominal verbs, verb government and noun phrase agreement were applied to the generation of automatic exercises. Additionally, in subsection 3.4, more constructs were suggested for future implementation.

In section 3, implementations of the most salient morphosyntactic construct were illustrated in detail. Generation of automatic exercises for analytic and pronominal verbs, verb government and noun phrase agreement was implemented by designing rules for matching patterns. Deep feedback was structured in order to provided to the user a valuable learning experience. The most important learning constructs were successfully implemented with the currently available NLP tools as automatic exercises in Revita, answering affirmatively to Research Question 2.

To answer Research Question 3, evaluation of the implemented constructs was pursued. Analytic and pronominal verb rules were tested on a total of 177 sentences. Precision on the given set resulted in 100% and recall 96.4%. Noun phrase agreement rules were tested on a total of 34 test sentences. Both precision and recall resulted in 96.0%. Investigation of the errors showed that the errors generated do not actually affect the automatic creation of exercises by generating wrong clozes. It has been shown that testing pattern rules on specifically designed sentences is a valuable method for further development and correction of the rules. Results from this kind of evaluation provide a useful indication to advance the implementation of the exercises to a stage where precision and recall can be measured on authentic text at a greater extent.

Usability of exercises created on verb government resulted in 93.55%. Therefore the exercises can be used for learning purposes but an additional layer of feedback need to be implemented to deal with multiple-admissibility. Alternatively, usability can be raised by excluding the verbs that cause multiple-admissibility in multiple-choice exercises.

The generation of verb government exercises on authentic text still achieved low numbers of generated multiple-choice clozes, ranging from about 4 to 19 exercises every 10000 words depending on the type of text. Among the investigated text types, semi-authentic text resulted in the highest generation rate, suggesting that frequency of vocabulary and complexity of sentence structure have a relevant impact on automatic creation of exercises, at least for the verb government exercises with prepositions. The number of generated exercises for verb government was estimated to increase considerably once articulated prepositions are implemented.

This thesis contributes to the raising of Italian in Revita from its *beta* status towards an advanced development stage. It proposes best practices for defining support for a new language in the Revita Framework by identifying learning constructs and by suggesting evaluation methods, paving the way toward adding new languages in the future. Additionally, it serves as a documentation of what has been done, how and and what remains to be done.

5 Future work

While this work contributes to the raising of Italian from its *beta* status towards a fully developed language in Revita, many areas of improvement and further advancement remain to be explored.

Disambiguation of lemmata, parts-of-speech and morphological tags is one of the major challenges in the acurrent generation of automatic exercises in Italian. As pointed out in 2.3.1, failing in disambiguation during chunk creation may occur, although rarely. A rigorous evaluation of accuracy of disambiguation, for instance against annotated data, needs to be conducted to assess reliability of the designed pattern rules. A suitable annotated corpus will also allow to quantitatively measure precision and recall of patterns for analytic and pronominal verbs, verb government and noun phrase agreement on authentic text.

Implementation of articulated prepositions and, more generally, of contracted forms like cliticised tokens, will improve recall of patterns of analytic and pronominal verbs, as well as verb government, as mentioned in 3.1.1 and 3.2.1. Minor adjustments to detect elided tokens will slightly improve recall of patterns of noun phrase agreement, as seen in 3.3.1.

The implementations of more learning constructs amongst the ones suggested in 3.4 will provide a greater variety of exercises, improving and widening the learning experience of the learner of Italian. Also, the results achieved in this work will enable further pilot studies with actual learners, which will allow to measure in rigorous and quantitative terms the usefulness of Italian in Revita. In particular evaluating *facility index* of multiple-choice clozes for prepositions, as seen in 2.2, will test the difficulty of distractors. By gathering data from actual users, it will be possible to assess how often learners pick the wrong choice. If the language competence of the students fits with the level of the exercise, and if the cloze does not allow multiple-admissibility of answers, a rarely picked distractor suggests that the distractor is too obviously wrong. On the opposite, a distractor that is often chosen indicates that the distractor is relevant and holds high value in the learning setting. Evaluation on real users, both language learners and heritage speakers, will also allow to assess overall learning outcomes and to collect valuable learner corpora for future use.

While conclusions and suggested future work exposed in this thesis arouse from the implementation of exercises for Italian, in many cases they are valid language-independently.

For multiple-choice clozes, addressing multiple-admissibility remains an area of future investigation that will allow to enhance the learning experience as seen in 2.3.2.

For open-end clozes, dealing with ungrammatical answers, that cannot be recognized by the analyzer or chunker, remains also a relevant area of future study, since this kind of answers are common in non-native speakers as seen in 3.1.2.

Finally, a field of future study could be the application of automatic generation of exercises in standardized exams such as the Finnish Matriculation Examination, if not to substitute human effort which is still unattainable in the case of exercises aimed at testing language competence, at least to provide an inestimable practising tool for language learning.

References

[1] A.I.R.E. *Annuario delle statistiche ufficiali del Ministero dell' Interno*. 2019. URL: %5Csmall%7B%22http://ucs.interno.gov.it/FILES/AllegatiPag/1263/INT00041%5C_ANAGRAFE%5C%5C%5C_DEGLI%5C_ITALIANI%5C_RESIDENTI%5C_ALL%5C_ESTERO%5C_-A.I.R.E.-%5C_ed%5C_2019.pdf%22%7D.

[2] Tuula Aaltio. "LASCIATEMI PARLARE: Bilinguismo di bambini italo-finlandesi". In: (2019).

[3] Cecilia Andorno. *Reggenza*. 2011. URL: %5Csmall%7Bhttps://www.treccani.it/enciclopedia/reggenza%5C_(Enciclopedia-dell%5C%27Italiano)/%7D. (accessed: 14.4.2021).

[4] Lene Antonsen et al. "Generating modular grammar exercises with finite-state transducers". In: *Proceedings of the second workshop on NLP for computer-assisted language learning at NODALIDA*. 17. 2013, pp. 27–38.

[5] Peter K Austin and Julia Sallabank. *The Cambridge handbook of endangered languages*. Cambridge University Press, 2011.

[6] Morton Benson, Evelyn Benson, and Robert F Ilson. *The BBI Combinatory Dictionary of English: Your guide to collocations and grammar. revised by Robert Ilson*. John Benjamins Publishing, 2010.

[7] J. Bruner. *The Relevance of Education*. W. W. Norton, 1971. ISBN: 9780393240931.

[8] Angela Chambers. "Computer-assisted language learning: Mapping the territory". In: *Language teaching, 2010-01, Vol.43 (1)* (2009).

[9] Carol A Chapelle. *Computer applications in second language acquisition*. Cambridge University Press, 2001.

[10] Chia-Yin Chen, Hsien-Chin Liou, and Jason S Chang. "FAST–An Automatic Generation System for Grammar Tests". In: *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*. 2006, pp. 1–4.

[11] Council of Europe. *Common European Framework of Reference for Languages*. 2020. URL: https://europa.eu/europass/system/files/2020-05/CEFR%5C%20self-assessment%5C%20grid%5C%20IT.pdf.

[12] Council of Europe. *Common European Framework of Reference for Languages*. 2021. URL: https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale.

[13] Accademia della Crusca. *Crusca*. URL: https://accademiadellacrusca.it.

[14] Maurizio Dardano. *Nuovo manualetto di linguistica italiana.* Zanichelli, 2005.

[15] M Di Salvo. *Migrazioni, famiglie, generazioni: la trasmissione della lingua in alcune comunità italiane d'Inghilterra contesto inglese.* 2013.

[16] David M. Eberhard, F. Simons Gary, and Charles D. Fennig (eds.) *Ethnologue: Languages of the World. Twenty-second edition. Dallas, Texas: SIL International.* 2019.

[17] John Edwards. *Multilingualism.* Routledge, 1994.

[18] Yrjö Engeström. *Mielekäs Oppiminen Ja Opetus.* Helsinki: Valtion painatuskeskus, 1998.

[19] Joshua A Fishman. "Sociolinguistics: A brief introduction." In: (1970).

[20] Mikel L Forcada et al. "Documentation of the open-source shallow-transfer machine translation platform Apertium". In: *[Online] Departament de Llenguatges i Sistemes Informatics Universitat d Alacant, Available: http://xixona. dlsi. ua. es/˜ fran/apertium2-documentation. pdf* (2007).

[21] Robert Hart. "Language study and the PLATO system." In: *Studies in Language Learning* 3.1 (1981), pp. 1–24.

[22] Raili Hildén and Marita Härmälä. "Hyvästä paremmaksi - Kehittämisideoita kielten oppimistulosten arviointien osoittamiin haasteisiin". In: *Kansallinen koulutuksen arviointikeskus* (2015).

[23] Ciro Imperato. *Italian kielioppi.* suomi. Suomi: Finn Lectura, 2013. ISBN: 978-951-792-524-2.

[24] Ciro Imperato. *Italian kielioppi. Harjoituskirja.* Finn Lectura, 2017.

[25] Anisia Katinskaia, Sardana Ivanova, Roman Yangarber, et al. "Multiple Admissibility in Language Learning: Judging Grammaticality Using Unlabeled Data". In: *The 7th Workshop on Balto-Slavic Natural Language Processing Proceedings of the Workshop.* The Association for Computational Linguistics. 2019.

[26] Anisia Katinskaia, Javad Nouri, and Roman Yangarber. "Revita: a Language-learning Platform at the Intersection of ITS and CALL". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* European Language Resources Association (ELRA). 2018.

[27] Anisia Katinskaia, Javad Nouri, and Roman Yangarber. “Revita: a system for language learning and supporting endangered languages”. In: *Proceedings of the Joint 6th Workshop on NLP for Computer Assisted Language Learning and 2nd Workshop on NLP for Research on Language Acquisition at NoDaLiDa, Gothenburg, 22nd May 2017*. 134. Linköping University Electronic Press. 2017.

[28] Anisia Katinskaia, Roman Yangarber, et al. “Digital cultural heritage and revitalization of endangered Finno-Ugric languages”. In: *CEUR Workshop Proceedings*. 2018.

[29] Kenji Kitao and Kenichi Kamiya. “Using cloze generator to make cloze exercises”. In: *International Journal of Pedagogies and Learning* 5.2 (2010), pp. 67–79.

[30] S Kathleen Kitao and Kenji Kitao. “Testing Grammar.” In: (1996).

[31] John Lee and Stephanie Seneff. “Automatic generation of cloze items for prepositions”. In: *Eighth Annual Conference of the International Speech Communication Association*. 2007.

[32] Alexey Malafeev. “Automatic Generation of Text-Based Open Cloze Exercises”. In: *International Conference on Analysis of Images, Social Networks and Texts*. Springer. 2014, pp. 140–151.

[33] Michael McCarthy. “Putting the CEFR to good use: Designing grammars based on learner-corpus evidence”. In: *Language Teaching* 49.1 (2016), p. 99.

[34] Detmar Meurers. “Natural language processing and language learning”. In: *The encyclopedia of applied linguistics* (2012).

[35] Enrico Olivetti. *Dizionario Italiano Olivetti*. URL: `https://www.dizionario-italiano.it/`.

[36] Jr Oller and W John. “Cloze tests of second language proficiency and what they measure 1”. In: *Language learning* 23.1 (1973), pp. 105–118.

[37] Luca Pavan. “Preposizioni italiane e articoli: difficoltà e strategie nell’apprendimento ai livelli più bassi del QCER”. In: *Verbum* 7 (2016), pp. 252–260.

[38] Peng Qi et al. “Stanza: A Python Natural Language Processing Toolkit for Many Human Languages”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 2020.

[39] Alberto A Sobrero and Paola Benincà. *Introduzione all’italiano contemporaneo. Le strutture*. 1993.

[40] Ayse Saliha Sunar, Yuki Hayashi, and Toyohide Watanabe. "Multiple–choice cloze exercise generation through English grammar learning support". In: *International Journal of Knowledge and Web Intelligence* 4.1 (2013), pp. 79–92.

[41] Istituto Treccani. *Grammatica.* URL: `https://www.treccani.it/enciclopedia/avere-o-essere_%5C%28La-grammatica-italiana%5C%29`. (accessed: 21.4.2021).

[42] Istituto Treccani. *Grammatica.* URL: `https://www.treccani.it/enciclopedia/arte_lingua_e_letteratura/lingua/grammatica/`.

[43] Francis M Tyers et al. "Free/open-source resources in the Apertium platform for machine translation research and development". In: (2010).

[44] YLE Ylioppilastutkintolautakunta. *Abitreenit.* URL: `https://yle.fi/aihe/abitreenit/harjoittele`. (accessed: 06.04.2020), select the subject: *aine> Italia, lyhyt oppimäärä* and then specify the type of exercise: *Rajaa hakua> Kielioppi ja sanasto.*

[45] Kamariah Yunus, Radzuwan Ab Rashid, et al. "Colligations of prepositions: Essential properties of legal phraseology". In: *International Journal of Applied Linguistics and English Literature* 5.6 (2016), pp. 199–208.

Appendices

Appendix A Active conjugation of a verb

Conjugation of transitive verb *mangiare*, active voice			
SINGLE-TOKEN	ANALYTIC VERB	SINGLE-TOKEN	ANALYTIC VERB
INDICATIVO			
Presente	**Passato Prossimo**	**Passato Remoto**	**Trapassato Remoto**
mangio	ho mangiato	mangiai	ebbi mangiato
mangi	hai mangiato	mangiasti	avesti mangiato
mangia	ha mangiato	mangiò	ebbe mangiato
mangiamo	abbiamo mangiato	mangiammo	avemmo mangiato
mangiate	avete mangiato	**mangiaste**	aveste mangiato
mangiano	hanno mangiato	mangiarono	ebbero mangiato
Imperfetto	**Trapassato Prossimo**	**Futuro Semplice**	**Futuro Anteriore**
mangiavo	avevo mangiato	mangerò	avrò mangiato
mangiavi	avevi mangiato	mangerai	avrai mangiato
mangiava	aveva mangiato	mangerà	avrà mangiato
mangiavamo	avevamo mangiato	mangeremo	avremo mangiato
mangiavate	avevate mangiato	mangerete	avrete mangiato
mangiavano	avevano mangiato	mangeranno	avranno mangiato
CONGIUNTIVO		**CONDIZIONALE**	
Presente	**Passato**	**Presente**	**Passato**
mangi	abbia mangiato	mangerei	avrei mangiato
mangi	abbia mangiato	mangeresti	avresti mangiato
mangi	abbia mangiato	mangerebbe	avrebbe mangiato
mangiamo	abbiamo mangiato	mageremmo	avremmo mangiato
mangiate	abbiate mangiato	mangereste	avreste mangiato
mangino	abbiano mangiato	mangerebbero	avrebbero mangiato
Imperfetto	**Trapassato**	**IMPERATIVO**	
mangiassi	avessi mangiato	**Presente**	
mangiassi	avessi mangiato	**mangia**	
mangiasse	avesse mangiato	**mangiate**	
mangiassimo	avessimo mangiato	**INFINITO**	
mangiaste	aveste mangiato	**Presente**	**Passato**
mangiassero	avessero mangiato	mangiare	avere mangiato
GERUNDIO		**PARTICIPIO**	
Presente	**Passato**	**Presente**	
mangiando	avendo mangiato	mangiante	
		Passato	
		mangiato	

Appendix B **Passive conjugation of a verb**

Conjugation of a transitive verb, passive voice			
ANALYTIC VERB	ANALYTIC VERB	ANALYTIC VERB	ANALYTIC VERB
INDICATIVO			
Presente	**Passato Prossimo**	**Passato Remoto**	**Trapassato Remoto**
sono mangiato/a	sono stato/a mangiato/a	fui mangiato/a	fui stato/a mangiato/a
sei mangiato/a	sei stato/a mangiato/a	fosti mangiato/a	fosti stato/a mangiato/a
è mangiato/a	è stato/a mangiato/a	fu mangiato/a	fu stato/a mangiato/a
siamo mangiati/e	siamo stati/e mangiati/e	fummo mangiati/e	fummo stati/e mangiati/e
siete mangiati/e	siete stati/e mangiati/e	foste mangiati/e	foste stati/e mangiati/e
sono mangiati/e	sono stati/e mangiati/e	furono mangiati/e	furono stati/e mangiati/e
Imperfetto	**Trapassato Prossimo**	**Futuro Semplice**	**Futuro Anteriore**
ero mangiato/a	ero stato/a mangiato/a	sarò mangiato/a	sarò stato/a mangiato/a
eri mangiato/a	eri stato/a mangiato/a	sarai mangiato/a	sarai stato/a mangiato/a
era mangiato/a	era stato/a mangiato/a	sarà mangiato/a	sarà stato/a mangiato/a
eravamo mangiati/e	eravamo stati/e mangiati/e	saremo mangiati/e	saremo stati/e mangiati/e
eravate mangiati/e	eravate stati/e mangiati/e	sarete mangiati/e	sarete stati/e mangiati/e
eravano mangiati/e	eravano stati/e mangiati/e	saranno mangiati/e	saranno stati/e mangiati/e
CONGIUNTIVO		**CONDIZIONALE**	
Presente	**Passato**	**Presente**	**Passato**
sia mangiato/a	sia stato/a mangiato/a	sarei mangiato/a	sarei stato/a mangiato/a
sia mangiato/a	sia stato/a mangiato/a	saresti mangiato/a	saresti stato/a mangiato/a
sia mangiato/a	sia stato/a mangiato/a	sarebbe mangiato/a	sarebbe stato/a mangiato/a
siamo mangiati/e	siamo stati/e mangiati/e	saremmo mangiati/e	saremmo stati/e mangiati/e
siate mangiati/e	siate stati/e mangiati/e	sareste mangiati/e	sareste stati/e mangiati/e
siano mangiati/e	siano stati/e mangiati/e	sarebbero mangiati/e	sarebbero stati/e mangiati/e
Imperfetto	**Trapassato**	**IMPERATIVO**	
fossi mangiato/a	fossi stato/a mangiato/a		**Presente**
fossi mangiato/a	fossi stato/a mangiato/a		sii mangiato/a
fosse mangiato/a	fosse stato/a mangiato/a		siate mangiati/e
fossimo mangiati/e	fossimo stati/e mangiati/e	**INFINITO**	
foste mangiati/e	foste stati/e mangiati/e		**Presente**
fossero mangiati/e	foste stati/e mangiati/e		essere mangiato/a/i/e
GERUNDIO			**Passato**
	Passato		essere stato/a/i/e mangiato/a/i/e
	essendo mangiato/a/i/e		
	Passato		
	essendo stato/a/i/e mangiato/a/i/e		

Appendix C Ambiguity between intransitive and passive verb forms

INDICATIVO				
	Presente	Passato-Prossimo	Imperfetto	Trapassato-Prossimo
TRANSITIVE ACTIVE	cambio	ho cambiato	cambiavo	avevo cambiato
INTRANSITIVE (ACTIVE)	cambio	**sono cambiato/a**	cambiavo	**ero cambiato/a**
TRANSITIVE PASSIVE	**sono cambiato/a**	sono stato/a cambiato/a	**ero cambiato/a**	ero stato/a cambiato/a
	Passato-Remoto	Trapassato-Remoto	Futuro-semplice	Futuro-Anteriore
TRANSITIVE ACTIVE	cambiai	ebbi cambiato	cambierò	avrò cambiato
INTRANSITIVE (ACTIVE)	cambiai	**fui cambiato/a**	cambierò	**sarò cambiato/a**
TRANSITIVE PASSIVE	**fui cambiato/a**	fui stato/a cambiato/a	**sarò cambiato/a**	sarò stato/a cambiato/a
CONGIUNTIVO				
	Presente	Passato	Imperfetto	Trapassato
TRANSITIVE Active	cambi	abbia cambiato	cambiassi	avessi cambiato
INTRANSITIVE (ACTIVE)	cambi	**sia cambiato/a**	cambiassi	**fossi cambiato/a**
TRANSITIVE PASSIVE	**sia cambiato/a**	sia stato/a cambiato/a	**fossi cambiato/a**	fossi stato/a cambiato/a
CONDIZIONALE				
	Presente	Passato		
TRANSITIVE ACTIVE	cambierei	avrei cambiato		
INTRANSITIVE (ACTIVE)	cambierei	**sarei cambiato/a**		
TRANSITIVE PASSIVE	**sarei cambiato/a**	sarei stato/a cambiato/a		
INFINITO				
	Presente	Passato		
TRANSITIVE ACTIVE	cambiare	avere cambiato		
INTRANSITIVE (ACTIVE)	cambiare	**essere cambiato/a**		
TRANSITIVE PASSIVE	**essere cambiato/a**	essere stato/a cambiato/a		
GERUNDIO				
	Presente	Passato		
TRANSITIVE ACTIVE	cambiando	avendo cambiato		
INTRANSITIVE (ACTIVE)	cambiando	**essendo cambiato/a**		
TRANSITIVE PASSIVE	**essendo cambiato/a**	essendo stato/a cambiato/a		

Appendix D Learning constructs

	Learning concept	**CEFR**
1	Indicativo-Presente	A1
2	Indicativo-Imperfetto	A1
3	Indicativo-Passato-Prossimo	A2
4	Indicativo-Trapassato-Prossimo	A2
5	Indicativo-Futuro-Semplice	B1
6	Indicativo-Futuro-Anteriore	B1
7	Indicativo-Passato-Remoto	B2
8	Indicativo-Trapassato-Remoto	B2
9	Congiuntivo-Presente	C1
10	Congiuntivo-Imperfetto	C1
11	Congiuntivo-Passato	C1
12	Congiuntivo-Trapassato	C1
13	Condizionale-Presente	B1
14	Condizionale-Passato	B1
15	Infinito-Presente	A2
16	Infinito-Passato	A2
17	Gerundio-Presente	A2
18	Gerundio-Passato	A2
19	Imperativo	A2
20	Auxiliary	A2, B1, B2
21	Conjugation-ARE	A1,A2,B1,B2
22	Conjugation-ERE	A1,A2,B1,B2
23	Conjugation-IRE	A1,A2,B1,B2
24	Conjugation-Irregular	A2,B1,B2,C1
25	Agreement-NP	B1, A2
26	Agreement-PronParticiple	B2
27	Agreement-DifficultNoun	B1, B2, C1
28	Government-Verb-prepositions	A2, B1, B2, C1
29	Government-Verb-conjunctions	A2, B1, B2, C1
30	Government-Adjective	A2, B1, B2, C1
31	Government-Noun	A2, B1, B2, C1