Engineering *Ingenium*: Improving Engagement and Accuracy with the Visualization of Latin for Language Learning

A thesis presented by

Sharon Zhou

to

The Department of the Classics and the Department of Computer Science in partial fulfillment of the requirements for the degree with Honors of Bachelor of

Arts

in the subjects of Classical Languages and Literatures and of Computer Science

Directed by

Associate Professor of Computer Science Krzysztof Gajos

Preceptor in the Classics Ivy Livingston

Professor of the Classics Mark Schiefsky

Harvard University

Cambridge, Massachusetts

March 2015

Table of Contents

Abstract	ii
Acknowledgements	iv
Introduction	1
1 Related Work	5
2 Formative Design Research	19
3 System Design	22
4 Experiment	35
5 Discussion	64
6 Future Work	67
Conclusion	69
Appendix	71
Bibliography	197

Abstract

The goal of Ingenium is to prompt beginning Latin students to think consciously and critically on Latin grammar prior to translating a sentence, while engaging them with the grammar in an intuitive and hands-on way. Learners commonly make errors in reading Latin, because they do not fully understand the impact of Latin's grammatical structure—its morphology and syntax—on a sentence's meaning. Synthesizing instructional methods used for Latin and artificial programming languages, Ingenium visualizes the logical structure of grammar by making each word into a puzzle block, whose shape and color reflect the word's morphological forms and roles. Ingenium is designed so that students do not focus on words in isolation, but make logical connections between words and group words together, so that the number of elements involved in the translation, or the cognitive load, is instantly reduced. For this reason, puzzle blocks only fit together if there is sound grammatical logic, preventing students from making syntactic errors and allowing them to experiment in a mistake-free environment. The blocks also serve to abstract out the grammatical terminology in favor of visual representation, making it easy for Ingenium to supplement current methods of Latin instruction and to maximize its adoption potential.



The audience of Ingenium is novice Latin students. When students' experience and confidence are at their lowest, they will benefit most from an interface that is not exclusively text-based, but engaging, hands-on, familiar, and intuitive. To evaluate Ingenium, I conducted an experiment with 67 beginning Latin students, measuring objective engagement, emotional and cognitive engagement, learning, change in self-efficacy, cognitive load, and accuracy in sentence translation with Ingenium over a traditional text-based interface as the baseline control. From the study, students, when using Ingenium, reported considerably higher levels of engagement and learning, opted to perform more optional problems, and completed translation exercises with substantially greater accuracy than when using the traditional interface. Ingenium has demonstrated its potential to be a powerful tool in improving students' engagement, learning, and accuracy in reading Latin. Thus, Ingenium supports Latin pedagogy by introducing the following research-based innovations:

- An approach to learning a natural language through hands-on interactions with the logic of its grammar
- 2. A method of visually and intuitively representing technical grammatical concepts
- 3. A tool that can effectively engage early learners of Latin with the grammar of the language

Acknowledgements

First and foremost, I would like to thank my thesis advisers. Thank you, Ivy Livingston, for giving me pointed direction on how to craft a more Ciceronian argument and for nourishing me both figuratively and literally. Thank you, Krzysztof Gajos, for encouraging me to think critically about my design problems and for teaching me the art of data analysis. Thank you, Mark Schiefsky, for brainstorming through the realm of possible intersections with me. Thank you, all three of you, for inspiring me, for criticizing me, for dedicating inordinate amounts of time to and with me, for compelling me to think beyond the realm of the possible, and for distilling and elucidating my jungle of intermingling thoughts.

I would also like to thank Professor Richard Thomas for always believing in me and never losing confidence in what I can do, even when I myself should waver. The magnitude of my appreciation has always been inexpressible.

I would like to thank my parents and my sister for constantly providing me with nuanced insight and a more holistic perspective on life.

I would like to thank Brittany Liebenow for always being on my side, for being the captain of team Sharon, for giving me the healing hand of love, and for fueling my dreams and ambitions.

I would like to thank both the Classics and Computer Science departments at Harvard University for not only allowing me, but trusting in my *ingenium* (my ability), to pioneer the first joint concentration of this kind in university history. I appreciate your willingness to endeavor with me and to embark with me on this unprecedented odyssey of invention and synthesis.

Finally, I would like to thank my friends for helping me recruit participants for my study. I could not have done nearly as much alone and appreciate the above and beyond effort put into spreading the word on my work.

Introduction

Defining the Problem

Beginning Latin students, particularly those who have never learned a second language, commonly make errors in reading Latin, because they do not fully understand the impact of Latin's grammatical structure on meaning. Often, this is a result of falsely assuming that the process of deriving meaning from a sentence in Latin mirrors the process in their native language. For example, native English speakers expect the first word in a sentence to be the subject, which certainly appears in Latin, but not consistently as in English. The other key source of difficulty lies in the numerous areas where sense is communicated differently by Latin and the student's native language.¹ For example, the subject of a sentence in English is typically identified by its place before the verb, whereas in Latin the ending of a word in the nominative case indicates its function as the subject. Similarly, an adjective that does not agree with a noun in its case, number, and gender, cannot modify that noun, however closely they appear in the sentence. In Latin, the forms of words tell much more about their function than their position in a sentence.

Latin instruction has regularly given students their grammatical training with repetitive, yet rigorous, form drills to address this problem.² However, mastery

¹ These common errors in reading Latin are discussed in Markus and Ross's paper, "Reading Proficiency in Latin through Expectations and Visualization." More anecdotally, I have also borne witness to these errors from my experience in Latin classes in both high school and college.

² These exercises are common to most widely used Latin textbooks, including *Wheelock's Latin, Jenney's First Year Latin, Oxford Latin Course*, and *Cambridge Latin Course*.

of isolated forms fails to transfer to the reading of a sentence. For example, when asked the case of *omnem*, meaning "every," a student may correctly identify it as an adjective in the accusative singular (either masculine or feminine). Yet, they may still translate it as the subject when it appears first in the sentence or fail to recognize that it could also modify *virum* (masculine accusative singular) in addition to *mulierem* (feminine accusative singular), with which it seems to "match" more obviously by their similar endings. In effect, students do not understand the connection between the form and its function in context, and still rely on the familiar grammatical structure of their native tongue to interpret a Latin sentence.

Furthermore, the meanings of words can vary substantially depending on their interactions with other words in the sentence. Traditional vocabulary drills practice words in isolation and provide students with poor preparation for applying that knowledge to full sentences. As a result, students fixate on a single meaning of a word without considering the impact of other words in the sentence. This presents particular problems with idioms, such as with the verb *gerere*, meaning "to carry" in isolation, in the idiom *bellum gerere*, which is translated idiomatically as "to wage war" and not "to carry a war."

Often, students will latch onto a single function of a word without analyzing the influence that their decision may have over the entire sentence. Because more than one form of a Latin word may be spelled the same, students must learn to determine which is appropriate to the context. Exercises, such as completing paradigm tables or parsing decontextualized forms, may train students to learn the

possible forms and functions, but not how to decide which form and function are correct in a specific context. For example, upon seeing *puellae*, an ambiguous form of "girl," in a sentence, students may correctly recognize its four possible forms (genitive singular, dative singular, nominative plural, vocative plural), but they struggle to determine which of the four this particular *puellae* should be. The correct form and function of *puellae* thus depend on the constraints restricted, and the possibilities allowed, by the rest of sentence. Ultimately, grammatical connections made between words, not their meanings or functions in isolation, dictate meaning in Latin; nevertheless, traditional drills train students to diligently memorize and master the latter.

Below, I present a brief summary of the two issues at hand:

1: Influences of Language Preconceptions and Traditional Instruction

- a. Students apply grammatical concepts of their native languages to Latin, generating misconceptions about Latin grammar
- b. Students do not internalize the concepts of Latin grammar with traditional drills

2: Nature of Latin Grammar

- a. The structure of Latin requires considering the interactions between words and the grammatical context in which they appear
- b. The lexical meaning of a word is context-dependent
- c. The function of a word is context-dependent

Establishing the Goals

To address these problems, I have identified and established the following objectives for Latin instruction moving forward:

- A. Place students' attention more on form and less on word order (addressing problems 1a and 2c above)
- B. Encourage students to reflect on the morphology and syntax first and vocabulary afterward (2a and 2b)
- C. Prompt students to think in terms of connected units, not isolated words (1b, 2a, 2b, and 2c)
- D. Teach students with a more intuitive approach that does not rely exclusively on formal grammatical terminology (1b)
- E. Actively engage students with the study of grammar (1b)

The Michigan Latin approach, developed at the University of Michigan, is an experimental methodology for teaching Latin that places a heavy emphasis on grammar, and has effectively taken steps to address objectives A, B, and C. In computer science, new instructional methods have surfaced to address goals C, D and E for beginners to learn the logical concepts, or "grammar," of text-based programming languages like Java and C++. Scratch, for example, is a visual programming language that has effectively engaged students with the basics of programming by abstracting out the precise syntax of text-based languages and

focusing on the interactions between fundamental concepts in the "grammar," or programming logic.

I aim to combine the advantages of both approaches and to expand upon them in a new system: Ingenium.³ Ingenium is an interactive visual representation of grammatical concepts in Latin (Goals D and E) that focuses on grammatical forms over word order (Goal A), redirecting students' attention from the meanings and functions of words in isolation to the grammatical interactions among all words in a sentence (Goals B and C). Ingenium is a piece of digital software that can be accessed from any online web browser for students and self-learners of Latin to interact with at home or in the classroom on their computer screens.

1 Related Work

Shortcomings of Traditional Drills

Latin instruction among English speakers has remained largely text-based since at least the nineteenth century. Traditional styles of activities include completion of declension and conjugation tables, English-to-Latin and Latin-to-English translations, and identification of forms and their roles (see Figure 1).⁴ It is customary for students to work individually

³ On the naming of Ingenium, the word *ingenium* in Latin means "innate quality," "invention," and "genius," from which we derive the English words "ingenuity" and "engineer."

⁴ These exercises are common in *Wheelock's Latin, Jenney's First Year Latin, Oxford Latin Course, Cambridge Latin Course,* among others.

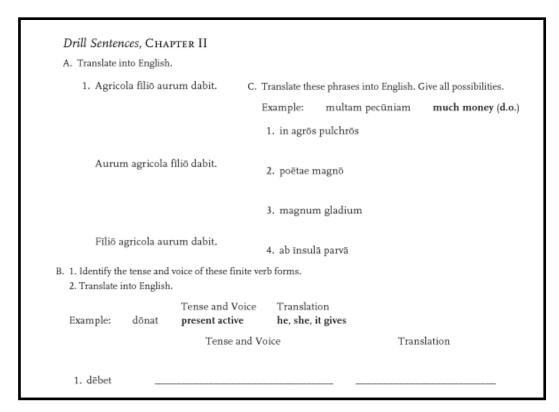


Figure 1: These standard, traditional exercises are taken from Keller and Russell's *Learn to Read Latin*, and can be found in many other Latin textbooks.⁵ On the top left, students are asked to translate sentences from Latin-to-English. On the top right, the translation exercises ask that students identify all possible interpretations of a sentence, if the grammar is ambiguous. On the bottom, students must parse the verb to show their understanding of the its form and role, giving its tense, voice, and meaning.

on rigorous drills of the same exercise for each lesson. While some students reach mastery of the language through this approach, many others struggle to fully grasp the broader concepts behind the drills and rely on rote memorization to complete them. Ultimately, these students fail to apply the concepts beyond the scope of these isolated exercises. Without internalizing the grammatical concepts, students can recognize that the word, *puellae*, as mentioned before, has four different functions, but cannot discern which function a specific *puellae* pertains to in a given context.

⁵ Equivalent exercises can also be found in *Wheelock's Latin, Jenney's First Year Latin, Oxford Latin Course, Cambridge Latin Course,* and other beginning Latin textbooks.

Michigan Latin Approach

Experimentation with instructional design in the early 1950s at the University of Michigan addressed the problem presented by isolated drills. In the Michigan Latin approach, students apply the exercise of "metaphrasing" to reading Latin sentences in order to narrow down the number of possible roles that each word in a sentence can assume. As the main method in the approach, "metaphrasing" is defined as "a beginner's technique for translating and reading Latin by graphically representing in English word order the morpho-syntactic expectations raised by the Latin words" (see Figure 2).6

adjectives are metaphrased before the expectation of a nounhead - $bonus \text{ good } \underline{S} \underline{Vs} \underline{+} \underline{DO} \underline{.}$

Figure 2: An example of metaphrasing on a worksheet,⁷ this exercise shows how the adjective *bonus* in nominative case, meaning "good" and often acting as a subject, prompts the student to consider what other must be necessary to complete the sentence, specifically a subject (S), a verb in the singular (Vs), and possibly, though not certainly (±), a direct object (DO).

The exercise has students analyze and deconstruct each word in a sentence into its possible functions and to consider how each of these functions might predict or affect the functions of other words in the sentence. Through "metaphrasing," students are trained to anticipate a constrained set of feasible roles that the remaining words can take. These anticipated predictions are called "unfulfilled syntactical expectations" and, like blanks in a fill-in-the-blank exercise, form "gaps"

⁶ From a worksheet on the Michigan Latin website (Ross, "Teaching Materials").

⁷ Found on the Michigan Latin website (Ross, "Teaching Materials").

in the sentence.⁸ With terminology of its own, the Michigan Latin approach attempts to address the problem of teaching with traditional grammatical meta-language that fail to explain concepts in an easy and intuitive manner (Goal D). However, instead of replacing formal grammatical terms with more easily comprehensible explanations, the approach introduces a new set of words, such as "metaphrasing" and "gaps," that ultimately prevents its success.

The Michigan Latin approach has not become widely adopted because its distinctive terminology does not integrate well with the existing and established lexicon for Latin instruction. It has nevertheless impacted the way of teaching Latin today in the 21st century. Some of its concepts have been incorporated into *Introduction to Latin* by Susan Shelmerdine, first published in 2005 and again in 2013, and now used at the University of Michigan after several years of using *Latin for Reading*, the latest textbook for the Michigan Latin approach. Shelmerdine's text scaffolds the process of transferring students' knowledge of forms to reading comprehension. The exercises in this text, however, are more traditional than the "metaphrasing" activities in the Michigan Latin approach and use the terminology of

⁸ On the concept and discussion of "gaps" and "gapping," Markus and Ross note in their paper, "Reading Proficiency in Latin through Expectations and Visualization," that students should develop a "fill in the gaps instinct" when approaching a Latin sentence, and that they should form gaps that correspond to "unfulfilled syntactical expectations." These concepts are part of the Michigan Latin approach.

⁹ From the University of Michigan's Classics department site: "After many years of using *Latin For Reading* by G. Knusdsvig, S. Craig and G. Seligson, we recently started experimenting with the second improved edition of S. Schelmerdein's [*sic*] *Introduction to Latin*" (Ross).

¹⁰ Shelmerdine writes about "expectations" in much the same way the Michigan Latin approach explains "unfulfilled syntactical expectations": "you should also be able to predict what form is coming next when you see a phrase such as 'puella et ______' (another nominative). If you pay close attention to the endings on words and practice the art of expecting what is likely to come next in a sentence you will soon find yourself reading Latin successfully" (18).

standard reference grammars. Her text has been adopted by several other programs across America, including—on a trial basis—Harvard University, and has become accepted and respected by Classics instructors and researchers.

Ultimately, the Michigan Latin approach addresses the common mistake among beginning and advanced students alike of spending too little time reflecting on the grammatical construction of a sentence and relying too heavily on word order and isolated dictionary definitions (Goals A and B). The approach trains students to think consciously about a Latin sentence's grammar before translating, and to use the grammar of the language, not their preconceptions such as word-order dependence, to guide their interpretations. The effectiveness of this approach rests in breaking down the grammatical construction of a sentence into more manageable parts and establishing a structured and dependable approach to understanding each part. Nevertheless, the approach still has a rigidity that is remnant of traditional methods. One particular limitation of the approach is the linearity of "gapping," as cited by University of Michigan Professor Gerda M.

Seligson, because it fails to successfully capture the flexibility of Latin's natural word order.

Teaching with Comprehensible Input

With the Michigan Latin approach failing to gain adoption and popularity among instructors and students, another method, Teaching with Comprehensible

ligson was one of the founders of the Michigan Latin approach, and

¹¹ Seligson was one of the founders of the Michigan Latin approach, and he cites it in his journal article, "Latin at Michigan."

Input (TCI), has prioritized generating greater enthusiasm for learning Latin over teaching formal grammar. TCI exposes students to Latin through the use of content that is "compelling," or so engaging that students are no longer aware of learning it. Yet, the material is still "understandable" to the student, so that the instructor may facilitate the "natural" and "unconscious" acquisition of the language, in lieu of formal grammatical instruction. TCI effectively shifts the focus away from the discussion of grammar, and particularly away from the use of grammatical terminology (Goal D). TCI has been gaining traction in Latin classes since the second half of the 20th century, and the primary TCI textbook, *Latin for Beginners*, has been published and adopted by Latin classes across the country. The strength of this method has been largely credited to its effectiveness in increasing pupil enthusiasm and interest, due to decreased formal instruction (Goal E). From the popularity of this approach, it is evident that a student's enthusiasm and intrinsic motivation are essential to effective Latin learning in the classroom.

¹² As explained by Robert Harrell in 2013 in his PDF publication "TCI FAQs Sheet."

¹³ The approach has even been cited as a viable and recommended method by members of the Classical Association of New England (CANE) and the Classical Association of the Middle West and South (CAMWS). TJ Howell (CANE) cites Dr. Robert Patrick (CAMWS) in his blog post on CANE's website in February 2014. Dr. Patrick, who draws inspiration from linguist Stephen Krashen, discusses the benefits of TCI for Latin specifically in his conference paper for the SALVI Board Summit in 2011. Robert Harrell discusses the application of TCI to Latin in his "TCI FAQs Sheet" in 2013. Stephen Kashen on TCI on all languages in his 2013 paper in the *Journal of Bilingual Education Research & Instruction*, entitled "The Case for Non-Targeted, Comprehensible Input."

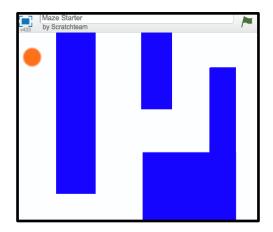
¹⁴ This TCI textbook has been used to replace or supplement traditional approaches to teaching the language in beginning Latin courses.

¹⁵ A study published in the science journal *System* in support of TCI strongly cites the increased enthusiasm and interest as factors in increased effectiveness of TCI as a teaching approach (Rodrigo, Krashen, Gribbons "The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level").

Scratch

There exists a more recent instructional design that engages students and removes the metalanguage of specialized terminology, which TCI effectively achieves, yet still emphasizes broader concepts, which the Michigan Latin approach has accomplished. In computer science, visual programming languages that move away from text-based languages like Java have surfaced to address the questions of keeping novices engaged with computer programming and of reducing early intimidation of foreign programming concepts. Scratch, a notable visual language born out of the Lifelong Kindergarten Group at MIT, emphasizes focusing on the broad and simple logical concepts and removing the precisions of a language's syntax and diction. Scratch uses puzzle blocks to represent the concepts in an intuitive and familiar, visual and hands-on interface for students to "play with," as they would with a physical jigsaw. Blocks assume shapes and colors that represent their concepts, with brief descriptors to help students quickly identify and associate keywords with concepts (see Figure 3).

¹⁶ When MIT Scratch was first tested in CS1, an introductory computer science course at Harvard College, professors David J. Malan and Henry H. Leitner found that "not only did Scratch excite students at a critical time (i.e., their first foray into computer science), it also familiarized the inexperienced among them with fundamentals of programming without the distraction of syntax" ("Scratch for Budding Computer Scientists" 2007).



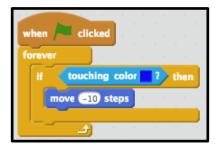


Figure 3: The top image is a screen displaying the Scratch program, in which blue walls construct a maze for the orange ball on the top left corner to move through. The bottom image is a block of code in Scratch for the ball. What does the code mean? When the green flag at the top right corner of the screen (top image) is clicked, the following "forever" block will execute forever or until the user decides to terminate the program. What will the ball do forever? If the ball is touching the color blue, then it will move backwards (negative) 10 steps. In action, this means that when the ball touches a blue wall, it will bounce back a distance of 10 steps, a measure of length like that is defined in Scratch. The function of the "forever" block illustrates the concept of a "for-loop" in programming, in which a block of code is performed forever as long as the program is run. The function of the "if" block depicts the conditional "if-statement" in programming which runs the code nested inside of it (here, moving -10 steps), if the condition that follows it (here, touching the color blue) is true. The function of the "order in programming which runs the code nested inside of it (here, moving -10 steps), if the condition that follows it (here, touching the color blue) is true.

With these shapes, colors, and brief textual components, students do not need to memorize a concept or how it works, but can instead focus on the concept's application and interaction with other concepts. Students then have the opportunity to interact with the logic behind basic computer programming by the familiar activity of moving around puzzle blocks. By using a visual approach that focuses on

 17 Screenshots from Scratch sample game program: Maze Runner (Lifelong Kindergarten Group at the MIT Media Lab).

the logic and meaning behind the concepts and that removes the intimidating diction of a text-based language such as Java, Scratch excites and engages novice learners. 18

Another key advantage of Scratch is that students also cannot make logical mistakes, because blocks only fit together if their concepts match logically.

Removing the discouraging error-prone environment of most text-based programming languages and encouraging students to experiment at their own pace, Scratch allows for a more welcoming foray into computer science, whose long-term effect is increasing students' engagement with the program itself and thus with using and understanding computer programming concepts in general. Ultimately, its hands-on and intuitive environment, the elimination of syntax precision and diction particulars, the focus on broad concepts, and the removal of student-made errors make Scratch succeed in engaging novice programmers.

Cognitive Load Theory

Both Scratch and the Michigan Latin approach effectively reduce the number of elements that students need to remember and mentally manage at any given moment as they program and translate Latin, respectively. In Scratch, students can move concepts around without memorizing a language's syntax and diction to construct that concept. Students are also able to join blocks so that multiple blocks become a larger block that they can handle as a single element. In the Michigan Latin

¹⁸ This finding is supported by research published by Harvard professors David J. Malan and Henry H. Leitner in their SIGCSE '07 conference paper, "Scratch for Budding Computer Scientists."

approach, students are taught to focus on one word at a time instead of trying to make sense of a sentence all at once. Students also form predictive constraints for the rest of the sentence, reducing the number of possible options that they must handle in their memory as they make decisions on the interpretation of the sentence. This method implies that Latin grammar is intrinsically difficult because students must hold in their working memories multiple combinations of meaning and function for every word before reaching a final interpretation that logically suits all of the words in the sentence.

In psychology, Cognitive Load Theory (CLT) dictates that people can only effectively hold a fixed number of seven elements in working memory, and that high interactivity among elements held in memory, such as relationships among words in a sentence, requires greater cognitive load and memory to manage. The core insight of CLT on teaching is that an effective instructional method is one that minimizes the burden of instructional materials on cognitive load, or the "extrinsic cognitive load," and that maximizes the "germane cognitive load," which is measured by the students' effort concentrated towards building schemas. Schemas are organized patterns of thought or behavior that turn into long-term strategies and techniques for similar problems.¹⁹

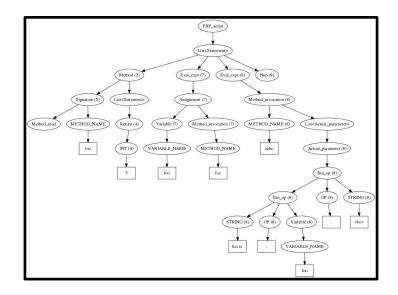
The visual aid provided by Scratch clears space in students' working memories from the granularities of syntax and diction, so that they can hold onto, and thus focus on, the broader conceptual relationships. According to CLT, the

¹⁹ Cognitive load theory is discussed at length in Sweller, van Merrienboer, and Paas's 1998 paper in the *Educational Psychology Review*, entitled "Cognitive Architecture and Instructional Design."

integration of text on the visual blocks of Scratch is helpful and necessary in effectively decreasing extrinsic cognitive load because it removes the need to recall what a visual representation symbolizes. In the Michigan Latin approach, the individual attention placed on each word initially moves students away from the high interactivity of the whole sentence, effectively reducing the cognitive load of translation.

Abstract Syntax Trees

In the theory of computer science, abstract syntax trees (ASTs) are diagrams that depict the conceptual breakdown and structure of a block of code. They abstract out the detailed syntax of the chosen language, focusing on the concepts and how they relate to each other in the selected section of code (see Figure 4).



```
<?php
   function foo()
   {
      return 5;
   }
   $foo = foo();
   echo "foo is $foo<br>";
?>
```

Figure 4: The AST is the image on top and illustrates the breakdown of a script in a tree-like form. On the bottom is its corresponding block of code in the programming language, PHP.²⁰ It deconstructs the more rigid structure of textual code into another more flexible configuration.

Scratch allows students to create visual ASTs with interactive blocks that represent the concepts found in programming. One distinct advantage of ASTs is their nonlinear, two-dimensional structure that makes them particularly versatile in illustrating hierarchies of concepts and the dependence of certain structures on others.

²⁰ These AST and code images were taken from an open source php program documentation, called *Phc documentation*.

Introducing Ingenium

Ingenium synthesizes these previous approaches to learning Latin and learning programming languages into a novel interface for Latin instruction that visualizes the logical structure of grammar by making each word into a puzzle block whose shape and color reflect the word's valid morphological forms and roles.

Below, I reiterate the five primary instructional goals that I address in my system:

- A. Place students' attention more on form and less on word order
- B. Encourage students to reflect on the morphology and syntax first and vocabulary afterward
- C. Prompt students to think in terms of connected units, not isolated words
- D. Teach students with a more intuitive approach that does not rely exclusively on formal grammatical terminology
- E. Actively engage students with the study of grammar

Similar to the Michigan Latin approach, Ingenium is designed so that students do not exclusively focus on words in isolation, but make logical connections between words and group words together, so that the number of elements involved in the translation, or the cognitive load, is instantly reduced (Goal C). For this reason, the puzzle blocks only fit together if there is sound grammatical logic, similar to those in Scratch, preventing students from making logical errors and allowing them to experiment in a mistake-free environment. The blocks also serve to abstract out the grammatical terminology in favor of visual representation (Goal D), making it easy

for Ingenium to supplement current methods of Latin instruction and to maximize its adoption potential.

Like "metaphrasing," the purpose of emphasizing the grammatical structures and connections in my program is to prompt students to reflect on grammatical roles and interactions within a Latin sentence. I have designed Ingenium to depart from linear word order and to place emphasis on the function of words over their position in a sentence by adding vertical dimensionality to the blocks as inspired by ASTs (Goal A). This in turn will prompt students to think consciously and critically on the grammatical relationships among words prior to translating a sentence (Goal B). Finally, like Scratch, the audience of Ingenium is composed of novice learners, and for the same reason. When students' experience and confidence are at their lowest, they will benefit most from an interface that is not exclusively text-based, but engaging, hands-on, familiar, and intuitive (Goal E).

Thus, Ingenium supports Latin pedagogy by introducing the following research-based innovations:

- An approach to learning a natural language through hands-on interactions with the logic of its grammar
- 2. A method of visually and intuitively representing technical grammatical concepts
- 3. A tool that can effectively engage early learners of Latin with the grammar of the language

2 Formative Design Research

In the earliest stages of the development of Ingenium, I created physical paper prototypes of puzzle blocks, cut from colorful 5 x 8-inch index cards. With them, I experimented with the visualization of the following grammatical concepts:

- Different valid cases of nouns with the same morphology
- Nominative subject and verb agreement
- Accusative case as the direct object
- Adjective and noun agreement
- Prepositional phrases
- Verb transitivity
- Use of linking verbs
- Subordination of relative clauses

In the first design (see Figure 5), the color of the block indicated its part of speech and its shape reflected its function. Interactions between prepositions and nouns, adjectives and nouns, verbs and nouns, and subordinate clauses and main clauses were all demonstrated. There were two sub-designs that I explored to depict the agreement of adjectives with nouns: adjectives could become embedded into a noun or they could attach themselves adjacent to an agreeing noun.





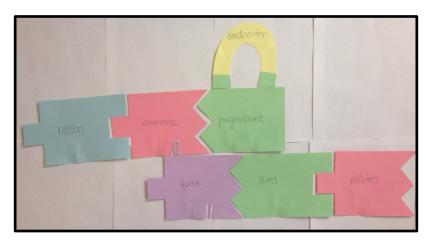


Figure 5: This is the first paper prototype design, where word order is subject-verb-object and parts of speech are color-coded as follows: red for nouns, green for verbs, purple for relative pronouns, blue for adjectives, and yellow for adverbs. In the top left image, there are two sentences: "the son calls the farmer" on top and "the sailor is a boy" on bottom. The top sentence illustrates the different shapes for a transitive verb (*vocat*) and a linking verb (*est*). Notice how *nauta* and *puer* in the bottom sentence can be exchanged and still make grammatical sense with *est*, a conjugated form of the verb "to be." The bottom image shows how the shape of an intransitive verb that does not take any noun thereafter might look like with *pugnabunt*. The full sentence in the bottom image is "our animals which are soldiers shall fight bravely." The relative pronoun *quae*, agreeing in number and gender with *animalia*, fits with *animalia* as it stems down to begin the relative clause: "which are soldiers." The magnet-like piece is the adverb, meaning "bravely," whose properties stick and are attracted to the green on verbs. Finally, the top right image demonstrate how a prepositional phrase would look: *agricolam*, meaning "farmer" in the accusative, would shrink to fit inside *ad*, the preposition meaning "towards" and which takes the accusative.

The second design (see Figure 6) showed how nouns of the same morphological form could assume multiple functions and gave users the ability to choose among the possible functions. In this design, color represented function.





Figure 6: This is the second paper prototype design. In the top left image, the blue *animal*, meaning "animal" is in the nominative and is the subject, while pressing the yellow button would cause it change into the yellow *animal* in the accusative, acting then as the direct object. The sentences are arranged so that the verb is in the middle with extending arms suggesting its subject to the left in blue and its object to the right yellow (subject-verb-object word order). In the bottom image, the full sentence is "a small boy calls the farmer," where *parvus*, meaning "small" modifies *puer*, meaning boy, showing how adjectives can be added to nouns.

Both designs enforced a set subject-verb-object word order. Not unlike "gapping" in the Michigan Latin approach, they constructed sentences in a horizontal and linear structure.

I asked instructors at the Montessori School in Lexington, MA, who were familiar with color-coded blocks representing parts of speech in English,²¹ and

²¹ The Montessori schools use color-coded pieces to depict various parts of speech in English. The relationships among words are not illustrated, though that may suffice for English sentences, whose meanings depend largely on word order. Because my approach focuses on Latin sentences, which depend much less on word order, I emphasize the logical connections and dependencies between words, as opposed to their independent roles.

Harvard students with high-school level Latin to play freely with 30 puzzle blocks and the two separate designs. Participants constructed both sensical and nonsensical sentences, demonstrating that our prototype enforced grammatical constraints, but allowed participants semantic freedom. This result was desirable, as the key objective was to shift the emphasis of reading Latin from vocabulary and definitions to concepts pertaining to grammar.

3 System Design

I have designed Ingenium with the two core objectives of making the logic of Latin grammar apparent with an intuitive interface (Goal D) and of creating an engaging experience (Goal E). Addressing the former objective, Ingenium visualizes Latin grammar by representing inflected Latin words with puzzle blocks whose shapes and colors indicate the words' grammatical roles in a sentence. The decision to have blocks represent an inflected form, as opposed to a dictionary lemma, or dictionary form, is to highlight the role of the word in a sentence, and to explore its constraints and possibilities within the context of the sentence.

A word's role is chiefly determined by its part of speech. I have thus designed the main shape of a word to reflect its part of speech. In the system, the following five parts of speech have distinctive shapes: nouns, verbs, adjectives, prepositions, and adverbs. I chose these parts of speech because they are more dependent on one another than others, e.g. conjunctions and interjections, and because they are most common and their mastery is essential to interpreting a Latin sentence. The

different shapes allow the student literally to connect words with one another to build larger units of meaning.

Noun. The noun block has a horizontally projecting knob that signals its case and a vertical, jaw-like clamp that can clasp onto optional components that depend on the noun (see Figure 7), e.g. adjectives.

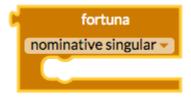


Figure 7: This is the block for *fortuna*, meaning "fortune" or "chance." The knob on the top left hand corner indicates the nominative case of *fortuna*. Its color indicates its specific case-number-gender combination: nominative singular feminine.

These optional components can be stacked up to an unlimited number and the enclosing clamp expands to accommodate them. While semantically it may not be realistic to have an infinite number of adjectives agreeing with a noun, it is not technically ungrammatical. A unique color is assigned to each of 36 possible combination of case, number, and gender.²² These same colors are used for adjectives (on which, see below), so that the agreement between noun and adjective is visually reinforced.

²² There are six main cases (nominative, genitive, dative, accusative, ablative, vocative), two numbers (singular, plural), and three genders (masculine, feminine, neuter) in Latin. I excluded the locative case, because it is used with a limited set of nouns and is generally introduced considerably later than the other cases. The 36 colors are distributed evenly on a 360-degree color wheel, so that they

can be maximally distinguished in hue from each other.

The external knob indicates only the noun's case, without taking into account its number or gender. While this might seem redundant because color also indicates case, the reasons for this decision are two-fold.

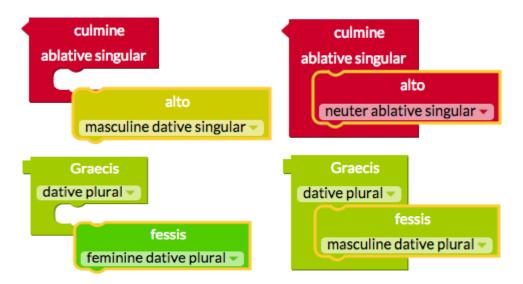


Figure 8: In the top row, *culmine*, meaning "fortress" and in the neuter ablative singular, does not agree with *alto*, meaning "high" when it is in the masculine dative singular. Their lack of correspondence in their colors signals immediately to the student that they do not fit together. In fact, if the student tried to put them together, *alto* would automatically jump away from *culmine* when dropped near it. This interaction is detailed later. When the student chooses the matching casenumber-gender combination of *alto*, however, the block immediately changes to the corresponding combination color and signal its fit with *culmine*. Similarly, the next row shows *Graecis*, the masculine dative plural of "the Greeks," and dative plural *fessis*, meaning "weary," first in the feminine, then in the masculine. Their colors are close, however, and many students would identify both as green. This is one factor that motivated the decision behind having knobs in addition to colors to signal the case of a noun.

With 36 total colors, each case can be represented by six different colors, making it difficult to determine a noun's case at a glance (see Figure 8). Since information on the noun's case is essential for the noun's primary interactions with verbs and prepositions, a clear indication of its case is particularly important, whereas these

interactions often do not require knowledge of the noun's number or gender.²³ Finally, the difference between the clamp and the knob reflects the noun's possible relationships: e.g. the knob locks into a verb or preposition, of which the noun is the object, and the clamp encloses those items (typically adjectives) whose form is dependent on that of the noun.

Because the same spelling of a noun can indicate several different combinations of case, number, and gender, I have included on noun blocks a dropdown menu of the noun's possible forms (see Figure 9).

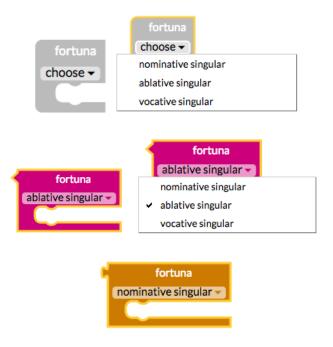


Figure 9: This is the flow of the dropdown menu for nouns with an ambiguous case-number-gender combination. There are three case-number combinations for *fortuna*: nominative singular, ablative singular, and vocative singular. Students are able to select from all of the noun's possible case-number combinations in the dropdown menu. Before a selection is made, the block does not have a knob and is a neutral gray color. Once a combination is chosen and the case-number combination is established, the corresponding knob appears dynamically in the top left hand corner and the color updates to the one assigned to that inflection combination. The dropdown menu can be used by students to change the block's function, and thus corresponding shape and color, at any time.

 23 With the notable exception of subject-verb agreement that requires information on the noun's number.

25

Students can choose the combination of case and number, which dynamically changes the noun's knob shape and color to reflect the selected combination. I designed this feature to emphasize that the same word in Latin can have multiple different forms, each with different potential roles, and thus that form in Latin is tied to function.

Adjective. The adjective block is very similar to the noun block: it is shaded the color of its inflected form (case, number, and gender) and includes a dropdown menu that allows students to change its form dynamically, if applicable (see Figure 10).

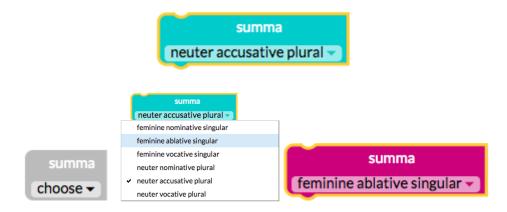


Figure 10: Above is the adjective *summa*, meaning "highest, top of" in Latin. The round teeth on the top and bottom left of the block correspond to the teeth shape in a noun block's clamp, suggesting that adjectives can be nested inside the noun clamp. Like declined nouns, declined adjectives can have an ambiguous case-number-gender combination. In the case of *summa*, there exist six such case-number-gender combinations. Like the noun block, the adjective block is initially knobless and gray to signify neutrality, until a combination is chosen.

The adjective has a concavity on its top and convexity on its bottom to indicate that it fits inside the clamps of nouns, which have mating shapes (see Figure 11).

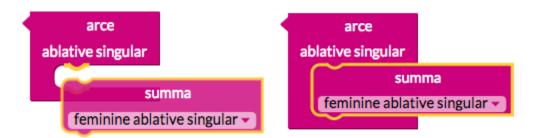


Figure 11: The adjective *summa*, meaning "highest, top of," and the noun *arce*, meaning "citadel" agree in case, number, and gender which make them the same color: ablative, singular, feminine. When *summa* is being dragged, it is highlighted with a yellow halo and becomes slightly transparent so that students can still see the shapes of other blocks that they are dragging over. When the teeth on *summa* are dragged close to the matching teeth in the clamp of *arce*, the teeth on *arce* highlight to signal a logical grammatical connection. At this point, students can drop *summa*, which will automatically fit tightly inside *arce*, as shown in the second image on the right. When this fit occurs, a snapping click sound and a yellow circular flash from the center of the connection will occur to provide positive feedback reinforcing the validity of the connection.

With their recesses above and teeth below, agreeing adjectives can be stacked infinitely on top of each other and also, as mentioned above, nested inside a noun clamp.

Verb. I created two different designs for the verb block, which I will refer to as *inline* and *external* (see Figure 12).

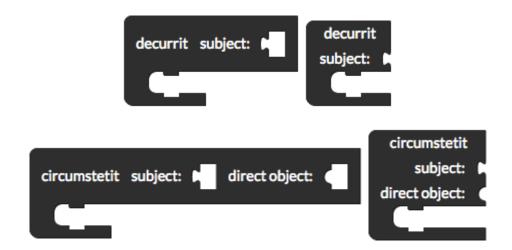




Figure 12: On the top row, the intransitive verb *decurrit*, meaning "runs down," is shown in the inline interface on the left and the external interface on the right. The knob shapes in both interfaces are identical and match with a nominative noun. In the subsequent rows are the verbs *circumstetit*, a transitive verb meaning "surrounded," and *dat*, a transitive verb that also takes an indirect object and means "gives." The inline design (verbs on the left) preserves the horizontal linear structure of a sentence in the original text, yet the external design (verbs on the right) deconstructs natural word order and focuses on function. The jaw-like clamps on both are the same, but their rectangular teeth differ in shape from the round teeth in noun clamps, signalling that they take will match with different blocks.

In both designs, verbs have two important components: a knob cutout and a verb clamp. The knob cutout allows connections between verbs and nouns. The shape and number of cutouts on a verb varies by the cases of nouns it requires or expects, e.g. nominative subject, accusative direct object, dative indirect object, etc. Inspired by the exercises with "unfulfilled syntactical expectations" outlined in the Michigan Latin approach, I designed the knob cutouts to reflect the "gaps" that a student would expect when coming across a verb. All finite, personal verbs have a gap with a nominative knob shape for the subject, and transitive verbs, for example, have an additional gap with an accusative knob shape for the direct object. Comparable to the noun clamp, the verb clamp is designed to enclose optional and infinitely stackable components that fit logically and grammatically within the verb phrase, such as prepositional phrases and adverbs. To visually distinguish what components

can be enclosed by the verb clamp as opposed to the noun clamp, I have designed the two clamps to have different teeth shapes.

In the inline design, knob shape cutouts rest inside the main body of the block, rather than protrude outwards. The motivation for this design is for the verb to enclose the entire noun block, reiterating the concept of one block enclosing another. In this case, the enclosure shows that the noun depends on the verb, or that the verb takes the noun. In the external design, the knob shape cutouts are aligned vertically beneath the body of the verb and above the verb clamp. I made this design decision to remove the horizontally linear order of words entirely. The deconstruction of linear word order further emphasizes the role that a noun plays vis-à-vis a verb. Regardless of design, I created the knob shape cutout to correspond in case to the knob shape on the noun to make the relationship between case and role intuitive and obvious in interactions between nouns and verbs.

Preposition. The preposition block also has two designs: *inline* and *external* (see Figure 13).



Figure 13: *Ab* is a preposition meaning "away from" that takes an ablative noun. Its inline design is one the left, and its external design is on the right. These designs are consistent with the inline and external verb designs. Their rectangular teeth signal to the student that these will fit inside the verb clamps, and the triangular knob shape corresponds to the ablative case on nouns.

Like the verb, the preposition also takes a noun of a certain case, usually either accusative and/or ablative. The decision to have two designs was to maintain visual

consistency with the verbs, in order to minimize confusion and to reduce the time to learn a new design. Students with inline verbs would also see inline prepositions, while those with external verbs would see external prepositions. In the inline design, the knob shape cutout is within the main body of the block, while in the external design, the knob shape is vertically oriented beneath the body, both identical to the designs for the verb. Additionally, because prepositional phrases fit inside the verb clamp, the preposition has teeth fitting the verb clamp on the top and bottom of the block to indicate both that it can be stacked and that it can fit inside the verb clamp.

Adverb. The adverb block has teeth on the top and bottom that visually indicate that it will fit inside the verb clamp and is stackable (see Figure 14).



Figure 14: The teeth on the adverb *tum*, meaning "then," signal that it can fit inside a verb clamp, and can be stacked on top of other adverbs and prepositions, all of which share the same teeth shape and can nestle inside a verb clamp.

Interactions between blocks. Students interact with these puzzle blocks by dragging and dropping them on the computer screen. When a student drags a block close to another block with which it shares a grammatically logical connection, the two blocks will automatically snap to fit together when the student stops dragging, or drops, the block. Before the student actually drops the block, there is a hint that the two blocks will fit together, through highlighting the contour of the shape, either the knob or the teeth shape, that will snap together with the dragged block. If the two blocks do not match, there is no such highlighting. In fact, if the student drops

such a block near a block with which it shares no connection, the block springs backward to indicate that the connection is not grammatically logical and therefore cannot happen. Students can experiment freely, but Ingenium will not allow them to make grammatical errors. Words simply do not snap together if there is no viable connection, which, like Scratch, makes learning the logic of Latin grammar less intimidating and more approachable. Students, however, are given ownership over the semantic decisions and thus can technically make semantic errors. I made this design decision to both highlight the emphasis placed on grammar in the goal of this project and to give students the creative license to construct grammatically sound sentences on their own. It is important for students to distinguish between grammatical and semantic sense of a sentence; too often in translation, students will draw an interpretation that may seem semantically more probable, but does not accurately reflect the grammar. For example, a student may translate the sentence canem mordet vir as "the dog bites the man," but upon closer inspection, the only grammatically logical translation with the Latin words used is "the man bites the dog." With Ingenium, they will find that they can also construct sentences that are grammatically logical, but semantically improbable or even impossible.

As students drag and drop the blocks, they are able to group individual blocks together into larger units, reducing the number of components they have to keep in working memory for each sentence—a desideratum in terms of Cognitive Load Theory. If given a Latin sentence to recreate graphically, students must demonstrate their comprehension by fitting all the blocks together so that no word is left

unattached. Alternatively, students may be allowed to experiment with any number of words to create sentences of their own; this activity demonstrates that a sentence can only make logical grammatical sense if the non-optional components, such as the knob shape cutouts on verbs, are all fully matched.

Goals Revisited. I designed the system with movable blocks centered around roles and relationships between words with the goal of encouraging students to consider the grammatical constraints and possibilities of a sentence before drawing conclusions on its interpretation (Goal C). The design of the system seeks to accomplish this by having students manipulate and restructure a sentence to make grammatical sense, thereby shifting visual attention from word order to word roles (Goal A). With Ingenium, students are able to not only see, but also visually manipulate, words with multiple roles for one form, such as nouns or adjectives with multiple combinations of case, number, and gender. Thus, they are able to dynamically visualize how the changes they make can alter and affect the configuration of the entire sentence. Students can see not only what logically works, but also what does not work, so that the feedback, through obvious visual changes in shape and color, is immediate and automatic.

Because one goal of Ingenium is to focus on the grammatical logic of a sentence (Goal B), I added the ability to gloss a word's dictionary form by right-clicking on its block. This shifts the focus from recalling vocabulary words and thus places greater emphasis on the grammar of the language. This design decision also makes Ingenium more accessible to any student who knows the basic grammar of

the words, yet not necessarily the lexicon of words used in the system. Finally, this fits into my broader objective of abstracting out the fine details and focusing on grammatical concepts over the specifics of vocabulary.

The ultimate design goal of Ingenium was to minimize the intimidation of learning Latin grammar by presenting an intuitive and immersive interface (Goals D, E). The freedom to experiment creatively with dynamically changing blocks and to make corrections based on the instantaneous feedback is intended to increase student engagement while reinforcing the point that changes in the form of words are inseparable from changes in function.

Usability tests

The level of dynamism among the shifting shapes and colors of the puzzle blocks may sound overwhelming, but it was refined after early usability testing with five volunteers (one who had no knowledge of Latin, one with knowledge of another inflected language [Ancient Greek] but not of Latin, one with novice knowledge of Latin, one with advanced knowledge of Latin, and one professor of Classics). They motivated the choice of black, charcoal gray, and light gray colors for the verb, preposition, and adverb blocks, respectively, to reflect these blocks' neutrality and independence from the colorful nouns and adjectives.

Because there were no definitions available on the blocks in my usability tests, some users expressed uncertainty and confusion over over the meanings of certain words. One user, who had been enthusiastic initially, was shaken after forgetting the meaning of a word. His embarrassment prompted him to ask to stop

using the system shortly after, while the other users asked whether there was more to test and play with. This observation motivated my decision to include dictionary glossing when right-clicking on a word, to shift the student's attention to the grammatical problems and away from vocabulary questions. I also wanted to remove the user's ability to make mistakes related to either vocabulary or grammar, because, as research on Scratch had found, an error-prone environment has a particularly negative impact on novice learners.²⁴

The clamps on nouns and verbs were initially a source of confusion. This was because users were first shown the blocks including the clamps, but without explicit direction on how to use them. This led to the conclusion that users needed to be eased into the usage of clamps. In the current design, nouns and verbs are first shown without modifiers; by removing the need for clamps to enclose the modifiers, the blocks can be presented first without clamps. Only later, when optional components were added (e.g. adjectives modifying nouns and adverbs modifying verbs), did the clamps holding them become a feature of the block. After their initial introduction, the clamps were consistently shown, regardless of the presence or absence of components that fit inside them. Finally, users' imperfect comprehension of my text-only directions for using the system motivated my decision to switch to more visual directions, combining images and text for each step.

²⁴ This finding is supported by research published by Harvard professors David J. Malan and Henry H. Leitner in their SIGCSE '07 conference paper, "Scratch for Budding Computer Scientists."

4 Experiment

I designed my main experiment to evaluate Ingenium against traditional Latin exercises. Specifically, I sought to observe the effect of using Ingenium over traditional text-based exercises on novice Latin students' cognitive and emotional engagement, self-efficacy, learning, cognitive load, and accuracy in translation.

Hypotheses

Hypothesis 0: Students will prefer using Ingenium over the traditional system.

Hypothesis 1: Students will be more cognitively engaged with Ingenium than with the traditional system.

 Students will be equally cognitively engaged with the external interface and with the inline interface of Ingenium.

Hypothesis 2: Students will be more emotionally engaged with Ingenium than with the traditional system.

 Students will be equally emotionally engaged with the external interface and with the inline interface of Ingenium.

Hypothesis 3: Students will be willing to spend more time engaged with Ingenium than on a traditional system.

• Students will be willing to spend an equal amount of time engaged with the external interface and with the inline interface of Ingenium.

Hypothesis 4: Students will report a greater positive change in self-efficacy after using Ingenium than the traditional system.

• Students will report no differences in changes in self-efficacy after using the external interface and after using the inline interface of Ingenium.

Hypothesis 5: Students will report that they learn more with Ingenium than with the traditional system.

 Students will report that they learn equally as much using the external interface as with the inline interface of Ingenium.

Hypothesis 6: Students will report lower levels of cognitive load after using Ingenium than after using the traditional system.

• Students will report equal levels of cognitive load with the external interface as with the inline interface of Ingenium.

Hypothesis 7: Students will be able to complete exercises with greater accuracy on translation exercises using Ingenium than using the traditional system.

- Students will be able to complete exercises with equal accuracy using the external interface as using the inline interface of Ingenium.
- When the use of blocks is optional in the translation exercises, students who
 use the blocks in Ingenium will outperform in accuracy those who did not
 use them.
- When the use of blocks is optional in the translation exercises, students who
 complete the blocks into full sentences will outperform in accuracy those
 who did not complete them.

Tasks

I designed two activities that are standard and common to many instructional Latin materials: translation exercises and multiple-choice fill-in-theblank exercises. In both the experimental condition using Ingenium and the control condition using the traditional system, students were asked to perform translation and fill-in-the-blank activities. In both conditions, students were presented with three rounds of exercises. Each round consisted of a set of four translation exercises and one fill-in-the-blank exercise. Each level increased in the complexity of interactions among words: level 1) interactions between nouns and verbs; level 2) interactions among nouns, verbs, and adjectives; level 3) interactions among nouns, verbs, adjectives, adverbs and prepositions. In every set of four translations, using the blocks in the experimental condition was required for the first two and optional for the last two. In all fill-in-the-blank exercises, the use of blocks was always completely optional. For both conditions, students could skip the last round of exercises, i.e. the third set of translations and the third fill-in-the-blank exercise, at any point by pressing a button stuck to the top right corner of their screen which read: "I'm finished using this tool." This button was used primarily to measure objective engagement, by determining how many optional exercises the participant completed in each condition and when the participant decided to stop engaging with the tool.

The sentences in the two conditions were isomorphic; each sentence in one condition had an equivalent sentence in the other condition. Sentences used for

translation and fill-in-the-blank exercises were largely quoted or adapted for the novice level from commonly read Latin texts, including Virgil's *Aeneid* and *Georgics*, Cicero's *First Oration Against Catiline* and *Letters to Atticus*, Catullus' *Poem 64*, and Caesar's *Commentaries on the Gallic War*, among others.²⁵

To investigate the impact of Ingenium on students' feeling of self-efficacy and confidence in their ability to perform the exercises, I asked students the same five questions evaluating self-efficacy in a pre-assessment given before using any of the tools and in a post-assessment after the use of each tool. To evaluate emotional and cognitive engagement, cognitive load, and learning, I asked the same questions pertaining to all in both post-assessments.

Procedures

In-person participants were recruited through visits to college Latin classes, and remote participants through email correspondence with Classics departments at over 50 colleges and universities across the United States, email notices to Harvard College's 12 house lists and other internal lists, and presence on social media (Facebook and Twitter).

All participants were asked to navigate to www.teachmelatin.com on their computers, where the study was hosted (See Figure 15).

 $^{\rm 25}$ See the Appendix: Flows of the Experiment for all the sentences used in exercises.

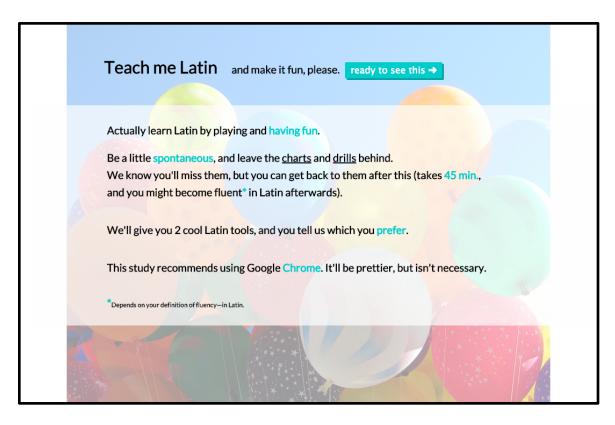


Figure 15: The landing page of www.teachmelatin.com includes information about the duration of the study and the preferred browser. It does not disclose that there is a clear control or experimental condition, but that participants will be presented with two different Latin tools.

Remote participants who were recruited through Harvard College's email lists and social media were screened through an online Google Form that determined their age, and level and years of Latin.²⁶ Only adults with ½ year to 1½ years of college or high school Latin, or whose Latin was at the novice level, were given the link to the study and eligible to participate. In this form, remote participants were notified of the \$10 Amazon gift credit as compensation upon completion of the study. Other remote participants were told of the compensation in the body of the email notice.

²⁶ See the *Remote Participant Filtering* form in the *Appendix: Form Questions*.

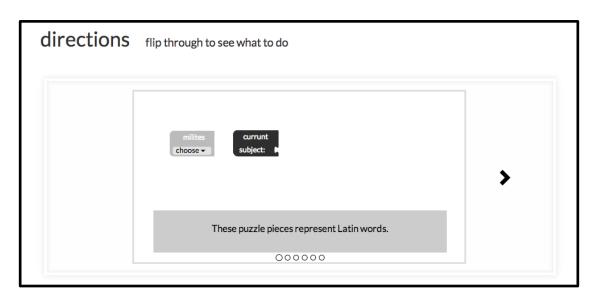
In-person participants were verbally informed of \$10 cash compensation, prior to signing up for the study.

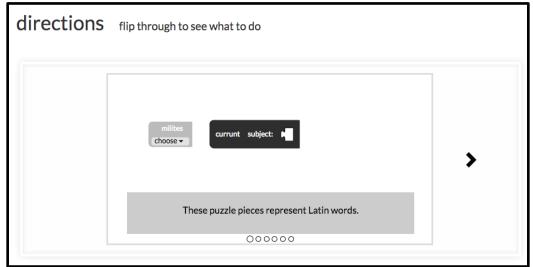
At the start of the study, in-person participants were told that they should make comments if they found anything confusing, but that the experimenter could not answer their questions or interfere unless the system faulted. All participants were then asked to fill out a questionnaire on their basic demographics and Latin experience. Following this questionnaire, they were presented with a preassessment form that asked questions on their subjective self-efficacy, or confidence, with fundamental grammatical concepts in Latin. Their answers were multiple choice on a 1–7 Likert scale, with 1 being "strongly disagree" and 7 being "strongly agree." All questions in both forms had to be answered in order to proceed.

They were then randomly assigned with a 50% chance to perform either the control or the experimental condition first. They were also randomly assigned with a 50% chance to either the inline or the external design of the experimental condition. They were not aware of the assignments, but were introduced to the first set of directions for the first set of translation exercises if the control was first and a set of directions for a warmup exercise to practice using the blocks in the experimental condition (see Figure 16).

²⁷ See the *Basic Demographics* form in the *Appendix: Form Questions*.

²⁸ See the *Pre-Assessment* form in the *Appendix: Form Questions*.





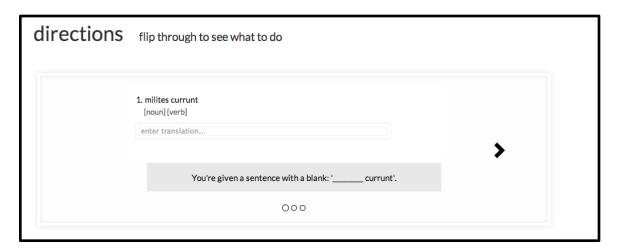
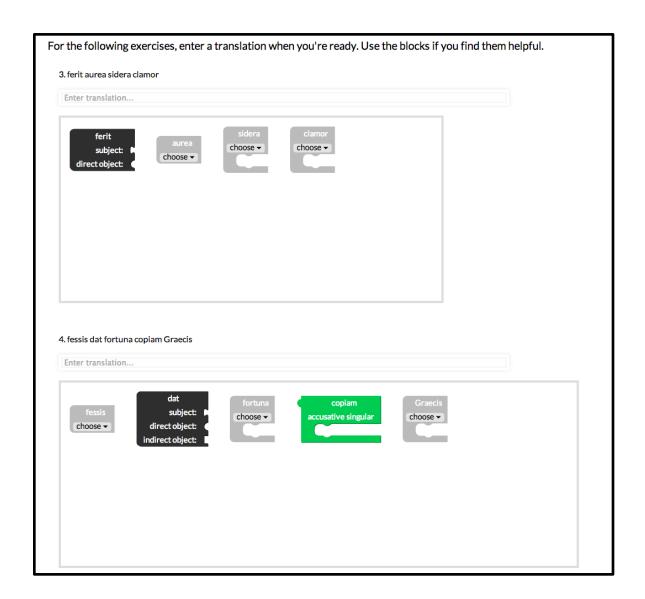


Figure 16: The top image is the first set of directions on the experimental condition with the external interface. The middle image is the first set of directions on the experimental condition with the inline interface. The bottom image is the first set of directions on the control condition. Please refer to the *External Interface* and *Inline Interface* subsections of the *Experimental Condition* section and the *Control Condition* section in the *Appendix: Flows of the Experiment* for the full collection of slides in all three sets.

Participants alternated between translation and fill-in-the-blank exercises for a total of three rounds of each.²⁹ Each set of translation exercises included four sentences or phrases, only the first two of which required the participant to use the blocks in the experimental condition; the use of the blocks was not mandated in the latter two. The fill-in-the-blank exercises all indicated that the blocks were completely optional (see Figure 17).

²⁹ For complete flows of the experiment, please refer to the images in the *Appendix: Flows of the Experiment*.



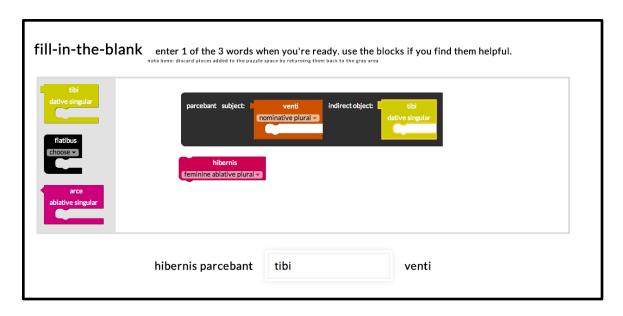
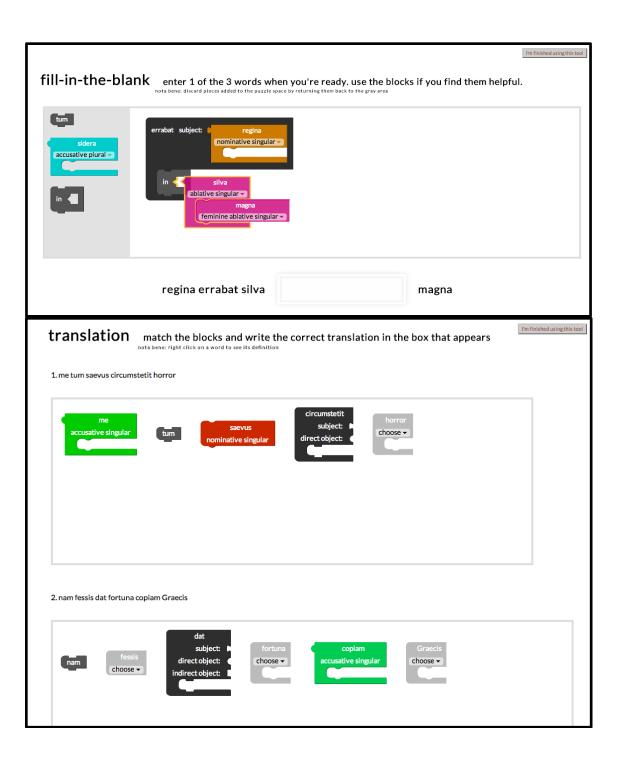


Figure 17: The top image shows the directions for the last two of the four translation exercises in each set, allowing participants to opt out of using the blocks and entering a translation directly: "For the following exercises, enter a translation when you're ready. Use the blocks if you find them helpful." The bottom image shows instructions to opt out of using the blocks in all the fill-in-the-blank exercises: "enter 1 of the 3 words when you're ready. use the blocks if you find them helpful." Please refer to the *External Interface* and *Inline Interface* subsections of the *Experimental Condition* section in the *Appendix: Flows of the Experiment* for the full collection of slides in the experimental condition.

At the 3rd set of translation exercises and on the 3rd fill-in-the-blank exercise, a skip button stating "I'm finished using this tool" appeared and stuck to the top-right corner of the participant's screen as they moved about the page so that it was omnipresent (see Figure 18). This skip button allowed the participant to skip to the end of the condition.



[adverb] [adjective] [adjective] [verb] [noun] [noun] [noun]	1
for the teacher gives a reward to no lazy boy	
nec tristitia rursus reginae persuadebit perpetua [adverb] [noun] [adverb] [noun] [verb] [adjective]	
nor shall eternal sorrow persuade the queen	saved
3. dea mollibus nautis proprium dedit domicilium prope [noun] [adjective] [noun] [adjective] [verb] [noun] [preposi domicilium, domicilii, n.: dwelling, home	
[noun] [adjective] [noun] [adjective] [verb] [noun] [preposition of the content o	
[noun] [adjective] [noun] [adjective] [verb] [noun] [preposidomicilium, domicilii, n.: dwelling, home	
[noun] [adjective] [noun] [adjective] [verb] [noun] [preposidomicilium, domicilii, n.: dwelling, home Enter translation	

Figure 18: All three images show the third round of exercises, in which participants have the option to exit their exercise at any point and click the skip button to the corresponding condition's post-assessment in the top right hand corner. The top image is of the third fill-in-the-blank exercise using the inline interface of the experimental condition. The middle image is of the third set of translation exercises using the external interface of the experimental condition. The bottom image is of the third set of translation exercises in the control condition. Please refer to the *External Interface* and *Inline Interface* subsections of the *Experimental Condition* section and the *Control Condition* section in the *Appendix: Flows of the Experiment* for the full collection of views in all three sets.

After completing or skipping to the end of their 1st condition, participants were presented with a post-assessment form asking them about their experience with the tool on the same 1–7 Likert scale as in the pre-assessment, with adjusted 1 and 7 values for certain questions where "strongly disagree" and "strongly agree"

were not suitable answers.³⁰ Questions pertained to subjective cognitive and emotional engagement, subjective learning, cognitive load, and subjective self-efficacy, and were identical in both conditions. The questions on self-efficacy were identical to those given in the pre-assessment. All questions were required to proceed. Once completed, participants then continued to the 2nd condition they were assigned and performed the tasks in this condition in the same order and manner. After both conditions and post-assessments were completed, participants were asked which tool they preferred: the 1st or the 2nd.³¹ Names were not given to the multiple choice to avoid biasing the responses and jeopardizing the validity of the question.

Lastly, participants were thanked for their time and participation. In-person participants received \$10 in cash at this time. Remote participants were asked for their email for their Amazon account, and were informed by email that their payment of \$10 in Amazon gift credit would be processed as soon as possible.³²

Design and Data Analysis

The design of my primary analysis was both within-subjects³³ and betweensubjects³⁴ on the following **factors**:

³⁰ See the *Post-Assessment* form in the *Appendix: Form Questions*.

³¹ See the *Final Preference* form in the *Appendix: Form Questions*.

³² See the *Compensation* form in the *Appendix: Form Questions*.

³³ "Within-subjects" means that it measures variations within the data set per participant, as opposed to variations between data sets of different participants.

³⁴ "Between-subjects" means that it measures variations between data sets of different participants, as opposed to variations within the data set of a single participant at a time.

- Experimental condition using Ingenium
 - Inline interface
 - External interface
- Control condition using a traditional system

My **measures** were:

Overall Preference

I examined participants' final tool preference, either the 1st or the 2nd tool, alongside which condition they were assigned first. I used the exact binomial test.³⁵ This test was suitable here, because I was observing whether the difference between preferring Ingenium and preferring the traditional system was statistically significant.

• Objective engagement

I examined the number of optional exercises participants performed before choosing to click the skip button, "I'm finished using this tool" to evaluate objective engagement between the control condition and the experimental condition. I used the Wilcoxon Signed Rank test.³⁶ This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each

³⁵ The exact binomial test measures whether the probability of an event occurrence is statistically significant from its nonoccurrence.

³⁶ This test is a within-subjects nonparametric statistical test. "Nonparametric" means that I do not assume a specific data distribution, e.g. normal distribution.

individual's results. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also observed the number of participants who chose to use the blocks in the experimental condition, when using the blocks was completely optional. Of those who opted to use it, I also observed the number who chose to complete the blocks into full sentences. I compared, and analyzed the differences between, the inline interface and the external interface. I used the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both and I was observing differences between different participants.

Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

• Subjective cognitive engagement

I examined the averages of two cognitive engagement questions in the post-assessment results of the control condition and the experimental condition.³⁹ I analyzed their differences using the Wilcoxon Signed Rank test. This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each individual's results. Observing the distribution, I

³⁷ As a reminder, this was the case for the last two exercises in every translation set and for every fill-in-the-blank exercise.

³⁸ This test is a between-subjects nonparametric statistical test. "Nonparametric" means that I do not assume a specific data distribution, e.g. normal distribution.

³⁹ See the *Post-Assessment* form in the *Appendix: Form Questions*.

cannot conclude that the data was normal, so I kept the test nonparametric.

Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed variances in cognitive engagement between the inline interface and external interface of the experimental condition with the same questions, using the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

• Subjective emotional engagement

I examined the averages of two emotional engagement questions in the post-assessment results of the control condition and the experimental condition. ⁴⁰ I analyzed their differences using the Wilcoxon Signed Rank test. This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each individual's results. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed variances in emotional engagement between the inline interface and external interface of the experimental condition with the same

⁴⁰ See the *Post-Assessment* form in the *Appendix: Form Questions*.

questions, using the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

Subjective learning

I examined the average of five learning questions in the post-assessment results in the control condition and the experimental condition.⁴¹
I analyzed their differences using the Wilcoxon Signed Rank test. This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each individual's results. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed variances in learning between the inline interface and external interface of the experimental condition with the same questions, using the Wilcoxon Rank Sums test. This test is appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants.

Observing the distribution, I could conclude that the data was normal, so I kept the test nonparametric.

Cognitive load

⁴¹ See the *Post-Assessment* form in the *Appendix: Form Questions*.

I examined the average of four cognitive load questions in the post-assessment results in the control condition and the experimental condition.⁴² I analyzed their differences using the Wilcoxon Signed Rank test. This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each individual's results. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed variances in cognitive load between the inline interface and external interface of the experimental condition with the same questions, using the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

Subjective change in self-efficacy

I determined the mean value of five self-efficacy questions in each condition for each participant.⁴³ I looked at the difference in average self-efficacy between the measure of self-efficacy in each condition and the preceding measure, which was either in the pre-assessment if the condition came first or the post-assessment of the other condition if the condition came

⁴³ See the *Post-Assessment* form in the *Appendix: Form Questions*.

⁴² See the *Post-Assessment* form in the *Appendix: Form Questions*.

second.⁴⁴ I used the Wilcoxon Signed Rank test for the delta of self-efficacy in both the control condition and the experimental condition. This test was appropriate here, because participants completed both the control condition and the experimental condition and I was observing differences within each individual's results. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed variances in the delta of self-efficacy between the inline interface and external interface of the experimental condition with the same questions, using the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants. Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

Accuracy

I examined the accuracy of all translation exercises with and without using Ingenium. I then specifically examined the optional translations and the variation in accuracy when using and not using Ingenium. I used the Wilcoxon Signed Rank test. This test was appropriate here, because participants completed both the control condition and the experimental

⁴⁴ For example, if the control condition were to come first, I would calculate the delta in self-efficacy for the control by subtracting the value in the pre-assessment from the value in the control condition, because the control condition was preceded by the pre-assessment. I would calculate the delta in self-efficacy for the experimental condition by subtracting the value in the post-assessment of the control from the value in the experimental condition, because the experimental condition was preceded by the control condition.

condition and I was observing differences within each individual's results.

Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

I also analyzed differences in translation accuracy between the external interface and the inline interface of the experimental condition. I used the Wilcoxon Rank Sums test. This test was appropriate here, because participants were assigned either the inline or the external interface, not both, and I was observing the differences between different participants.

Observing the distribution, I could not conclude that the data was normal, so I kept the test nonparametric.

For all translation exercises for which the use of blocks was optional, I observed the accuracy between those who used and did not use the blocks, as well as between those who completed and did not complete the blocks into full sentences. I used logistic regression to analyze this.⁴⁵ This analysis was suitable here because I had as a binomial variable whether the user used the blocks, in the one case, and whether the user completed the blocks, in the other case. With logistic regression analysis, I was able to examine how each variable—to use or not to use the blocks for one, to complete or not to complete the blocks for the other—impacted the user's accuracy in translation.

⁴⁵ Logistic regression is a type of statistical analysis that measures how the occurrence of an event, as opposed to its nonoccurrence, impacts something else, or another variable(s).

Participants

Of the 171 participants who began the study, 77 participants reached the end of the study. Ages of all participants ranged from 18–22 to 51+ years old. Most were college students, who hailed from over 42 different colleges and universities across the United States. Participants included current students, self-learners, and former students of Latin. Their formal study of Latin included representation from over 14 different Latin textbooks.

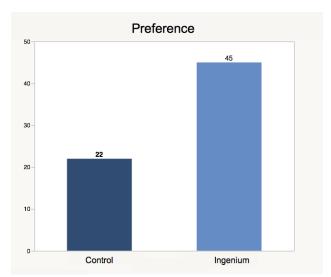
Adjustments of Data

I filtered out 10 participants who reached the end of the study, but did not complete the required sets of exercises. Additionally, there was an error in some studies that allowed them to skip the demographics and pre-assessment forms. This impacted 17 participants who reached the end of the study. Because these 17 participants still completed the required sets of exercises, they were excluded in subjective self-efficacy analysis, which required pre-assessment data, but still included in the analysis of other measures. Thus, for subjective self-efficacy, I only analyzed the 50 participants who were able to complete the pre-assessment. For all other measures, which did not require pre-assessment data, I analyzed the 67 participants who had completed the required activities.

Results

The means and p-values of my results comparing Ingenium against the control, and the inline interface against the external interface, are provided in Table 1 and Table 2 at the end of this section.

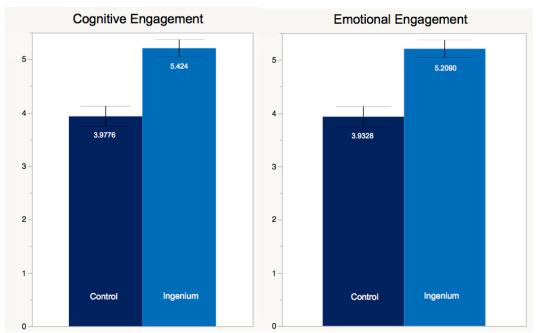
Hypothesis 0 (see Graph 1): 45 of the 67 participants preferred using Ingenium over the control, and this preference was statistically significant (p = 0.0067).



Graph 1: Shows student preference of Ingenium over control with count values.

Hypothesis 1 (see Graph 2): Participants reported higher levels of cognitive engagement in the experimental condition than in the control condition, and this difference was statistically significant (S = 629.000, p < 0.0001). For this measure, I also observed higher levels of cognitive engagement in the external interface than in the inline interface, and this difference was statistically significant (Z = 2.8975, p = 0.0038).

Hypothesis 2 (see Graph 2): Participants reported higher levels of emotional engagement in the experimental condition than in the control condition, and this difference was statistically significant (S = 576.000, p < 0.0001). For this measure, I also observed higher levels of emotional engagement in the external interface than in the inline interface, and this difference was statistically significant (Z = 2.2663, p = 0.0234).

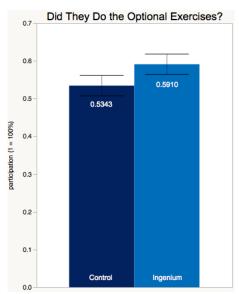


Graph 2: On the left, cognitive engagement and on the right, emotional engagement of Ingenium and control are shown with mean values and standard error bars. Ingenium had substantially greater subjective engagement overall than the control condition.

Hypothesis 3 (see Graph 3): Participants chose to complete more optional translation and fill-in-the-blank exercises when using Ingenium than when using the baseline control condition, and this difference was statistically significant (S = 389.500, p = 0.0339). For the fill-in-the-blank exercises alone, where using the

blocks in Ingenium was always optional, 79.60% of participants chose to use them and 44.28% chose to complete the sentence with them before finishing the exercise.

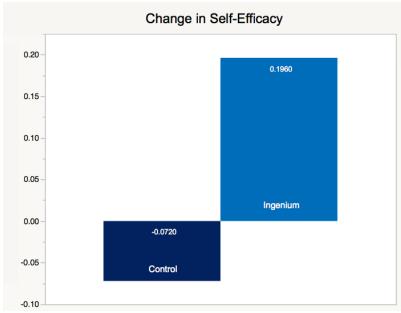
I observed slightly less participation in the optional exercises for the external interface than for the inline interface, though this difference was not statistically significant (Z = -1.2616, p = 0.2071). I observed decreased use of the blocks in the external interface than in the inline interface, yet this difference was not statistically significant (Z = -0.2346, p = 0.8145). I also observed a slightly higher instance of participants completing the blocks into full sentences in the external interface than in the inline interface, yet this difference was not statistically significant (Z = 1.2243, p = 0.2208).



Graph 3: Shows greater optional participation in Ingenium than in control with mean values and standard error bars. A statistically significant increase in objective engagement was observed for Ingenium over the control.

Hypothesis 4 (see Graph 4): Participants' self-efficacy increased in the experimental condition, while it decreased in the baseline control condition, yet the difference was not statistically significant (S = 139.000, p = 0.0602). For this

measure, I observed a minor effect from the external interface over the inline interface, but the difference was not statistically significant (Z = 0.9625, p = 0.3358).

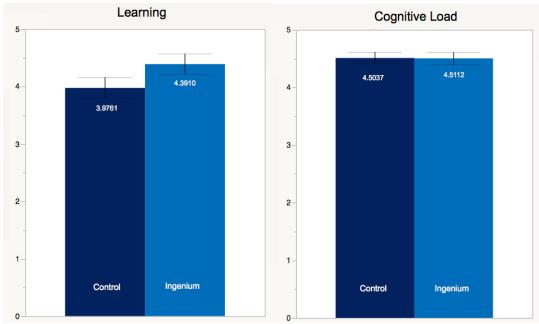


Graph 4: Shows a positive change for Ingenium and a negative change for the control with mean values displayed.

Hypothesis 5 (see Graph 5): Participants reported higher levels of learning in the experimental condition than in the control condition, and this difference was statistically significant (S = 242.000, p = 0.0471). For this measure, I also observed a higher levels of learning in the external interface than in the inline interface (Z = 2.3857, p = 0.0170).

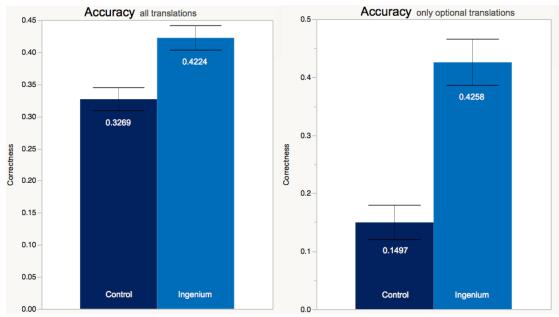
Hypothesis 6 (see Graph 5): Participants reported approximately equal levels of cognitive load in the experimental condition as in the control condition, though there was a slight increase for the experimental condition, though the difference was not statistically significant (S = -37.000, p = 0.7646). For this measure, I

observed a minor effect from the external interface over the inline interface, but this difference was not statistically significant (Z = -1.0215, p = 0.3070).



Graph 5: On the left is a graph showing participants' reports on learning and on the right is a graph showing their reports on cognitive load of Ingenium and control, both with mean values and standard error bars. A statistically significant increase in learning was observed in Ingenium, while a nearly equal amount of cognitive load was reported in both conditions.

Hypothesis 7 (see Graphs 6 and 7): Participants were substantially more accurate in completing translation exercises when using Ingenium in the experimental condition than in the baseline control condition, and this difference was statistically significant (S = 4095.00, p < 0.0001). Examining only optional translations, I observed that, of those that were completed, participants were also substantially more accurate when using Ingenium than in the control condition, and this difference was statistically significant (S = 379.500, p < 0.0001).

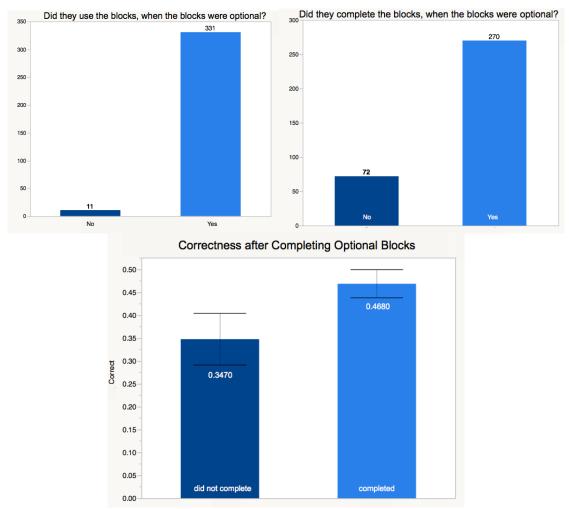


Graph 6: On the left is a graph showing the statistically significant increase in accuracy using Ingenium over the control among all translation exercises. On the right is a graph showing the greater accuracy using Ingenium for the optional translation exercises alone. Both graphs display the mean values and standard error bars for each condition.

I observed a minor effect from the external interface over the inline interface on the accuracy of all translation exercises, but this difference was not statistically significant (Z = 1.7965, p = 0.0724). I also observed a minor effect from the external interface over the inline interface on the accuracy of only optional translation exercises, but this difference was not statistically significant (Z = 1.0510, D = 0.2933).

In the 342 completed translation exercises where the use of blocks was completely optional, participants still used the blocks 96.78% (331 out of 342) of the time and completed the blocks 78.95% (270 out of 342) of the time. Those who completed the blocks (M = 0.4680) had greater accuracy than those who did not (M = 0.3470), though this difference was not statistically significant, X^2 (1, N = 339) =

3.4174, p = 0.0645. I did not perform logistic regression analysis for the effect of using the blocks on translation accuracy due to the insufficient representation of exercises in which participants chose not to use the optional blocks (only 11 out of 320).



Graph 7: On the top left is a graph showing the counts of whether participants used the optional blocks, which was nearly all the time. On the top right is a graph showing the count of whether participants opted to complete the optional blocks into full sentences, which was most of the time. The bottom graph shows the increase in accuracy for those who completed the optional blocks. It also displays the mean values and standard error bars for both conditions.

Table 1: A summary of means and p-values of measures pertaining to differences between using Ingenium and using the control. Statistically significant p-values are bolded and highlighted in green. The relevant hypothesis is shown parenthetically after the measure.

	Control	Ingenium	p-value
Cognitive engagement (1)	3.9776	5.4254	< 0.0001
Emotional engagement (2)	3.9328	5.2090	< 0.0001
Objective engagement (3) did optional exercises	0.5343	0.5910	0.0339
Change in self-efficacy (4)	-0.0720	0.1960	0.0602
Learning (5)	3.9761	4.3910	0.0471
Cognitive load (6)	4.5112	4.5037	0.7646
Accuracy (7) all translations	0.3269	0.4224	< 0.0001
Accuracy (7) optional translations	0.1497	0.4258	< 0.0001

Table 2: A summary of means and p-values of measures pertaining to differences between the external interface and the inline interface. Statistically significant p-values are bolded and highlighted in green. The relevant hypothesis is shown parenthetically after the measure.

	External	Inline	p-value
Cognitive engagement (1)	5.8333	5.0946	0.0038
Emotional engagement (2)	5.5500	4.9324	0.0234
Objective engagement (3) did optional exercises	0.5533	0.6216	0.2071
Objective engagement (3) used blocks	0.4844	0.4920	0.8145
Objective engagement (3) completed blocks	0.3778	0.3405	0.2208
Change in self-efficacy (4)	0.3455	0.0786	0.3358
Learning (5)	4.7467	4.1027	0.0170

Cognitive load (6)	4.2083	4.1581	0.3070
Accuracy (7) all translations	0.4603	0.3913	0.0724
Accuracy (7) optional translations	0.5000	0.3708	0.2933

5 Discussion

Ingenium vs. Control

My study demonstrates that Ingenium can increase student engagement, both objectively, in the larger number of optional exercises they were willing to complete using Ingenium, and subjectively, in their self-reported assessments of emotional engagement and cognitive engagement. Students also indicated a strong preference for Ingenium over the traditional system. When the use of blocks in the translation exercises was optional, nearly all students still used them (about 97%) and most completed them into full sentences (about 79%). Ingenium thus demonstrates that it can effectively capture student interest in Latin grammar and exercises. These findings support hypotheses 0, 1, 2, and 3.

Students using Ingenium translated their sentences more accurately, suggesting that Ingenium could have helped students reflect on the grammar of a sentence before they drew conclusions on the meaning. In fact, this finding strongly supports the Michigan Latin approach of focusing students' attention on grammatical concepts, rather than word order and vocabulary. Even among the

students who chose to continue through with the optional translation exercises instead of skipping them, those who opted to use Ingenium were much more accurate in their translations. Hypothesis 7 is supported by these findings.

Students reported having learned more through the exercises when using Ingenium than without, which supports hypothesis 5. Measuring subjective, or self-reported, learning is particularly interesting because while it does not prove the actual learning and mastery of concepts, it does measure how much students believe they have learned. Along with increased accuracy and engagement, student confidence in the effectiveness of instructional material also suggests that Ingenium is particularly well-suited for novice students whose initial enthusiasm and devotion to the subject build from their confidence in both their own ability and the tools of instruction.

While the change in self-efficacy in hypothesis 4 was not statistically significant between Ingenium and the control, the actual values of this measure suggest a notable difference. With Ingenium, the average change in self-efficacy was a positive number, while in the control, the average change in self-efficacy was negative. This signifies that students felt increased self-efficacy after using Ingenium, but decreased self-efficacy after using the control.

Students did not report lower measures of cognitive load with Ingenium than without. In fact, students felt equal amounts of cognitive load using both systems—there was even a negligible increase for Ingenium. This finding does not support hypothesis 6, but does point to an interesting conclusion when analyzed with the

other results. While students did not find Ingenium to relieve the mental burdens of the exercise any more so than the baseline control, they still found Ingenium more engaging and more effective on their learning of grammatical concepts. While students find Ingenium to be just as mentally demanding, they still found it more enjoyable and were willing to spend more time with it. It is also possible that, since this was their first exposure to Ingenium, the mental energy needed to learn the new system contributed to the perceived increase in cognitive load. The questions on cognitive load did not capture the mental demand of the translation or fill-in-the-blank exercises specifically, but rather asked about the overall effort and mental demand that students put into using each tool.

External vs. Inline

In my hypotheses, I did not predict there to be a significant difference in the two designs of Ingenium for any measure. My findings support hypotheses 3, 4, 6, and 7, as I found no statistically significant difference between the two interfaces in objective engagement, change in self-efficacy, cognitive load, and accuracy, respectively. Students, however, found the external interface substantially more engaging both cognitively and emotionally, as well as more effective in their learning. These do not support hypotheses 1, 2, and 5, suggesting that there was preference for one interface over another.

While objective engagement, measured by whether students opted to do the optional translation and fill-in-the-blank exercises, was nearly equal for both interfaces, students' perceptions and feelings of engagement skewed strongly for

the external interface. One possible reason is that the external interface is more stimulating and thought-provoking because it removes the sentence from its typical horizontal linearity. This action forces students to realign the sentence in a way that deconstructs the original word order and prevents them from relying on word order to make meaning out of the sentence. This finding is interesting because it suggests that students found it both more helpful and more exciting to depart from a sentence's natural word order, and hence to reflect and engage more critically on the sentence's grammar.

6 Future Work

Though not included in the study, I also created a lookup functionality and a sandbox interface. The lookup functionality allows the user to input a Latin word, phrase, or sentence. Using the *Lewis & Short Latin Dictionary* on Tufts University's Perseus Digital Library, ⁴⁶ the tool generates blocks from the entered Latin text. Because words in Latin often have multiple lemmas, all possible lemmas are offered as separate blocks. I envision creating an instructor interface, in which Latin instructors can make their own sentence blocks by entering sentences, choosing the correct lemma of a word if necessary and desired, so that they can upload these blocks to a separate student interface for their students to play and engage with.

⁴⁶ See the "Latin Word Study Tool" on the Perseus Digital Library's website.

Nevertheless, the lookup functionality could be used by all: teachers, students, and, in particular, self-learners.

The sandbox interface is a free-play zone, in which students can manipulate a wide array of blocks, deleting and adding new ones from a toolbox of various words. Students have full creative license to generate grammatically logical and sound, yet semantically unrestricted, sentences and phrases. The sandbox allows students to see changes at their own pace and without the constraints of a fixed set of sentences. This interface was inspired by Scratch's open playground where students can also freely manipulate blocks. I envision allowing students to save blocks that they have put together. These saved blocks could used to supplement or be incorporated into the Latin classroom, e.g. submitted for homework. Moreover, should students either want to return to see the logical connections or wish to ask their instructors about a confusing grammatical concept that can be communicated through the visualization that they are able to do so easily by saving a set of blocks. I also foresee creating this interface into a collaborative space in which students, in the classroom or virtually with others, can work together on sentence construction and decomposition.

The most immediate next steps in the system are to develop ways of visualizing interactions and grammatical concepts of greater complexity in order to include even more varied and intricate sentences in Latin. While particularly fitting for Latin, Ingenium can also be easily extendable to other inflected languages,⁴⁷ as well as scalable to a wide variety of languages with appropriate adaptations to the

⁴⁷ Greek is particularly fitting, because there is readily available parsing information available on the Perseus Digital Library in the same format as the information on Latin.

visualization of blocks. In fact, with visualizations of different languages, instructors and learners could potentially juxtapose and compare the logical visualizations of the target language and the learner's native tongue. In this way, variations in two languages' grammars could be easily observed and grasped without technical and textual explanation that often confounds the novice learner.

Conclusion

Initial testing has shown that Ingenium is effective in helping beginning Latin students overcome common grammatical problems. Most apparently, the system enables students to make literal connections between associated words, regardless of their position in an idiomatic Latin sentence. As a result, Ingenium may counteract preconceptions based on word order, while also reducing cognitive load on the learner, who no longer has to hold every possible combination of form and function until reaching the end of a sentence. Additionally, the printing of the Latin words on each block accustoms the learner to seeing that words whose endings are not spelled the same way may nevertheless agree and that apparent "matching" of endings does not necessarily indicate agreement; the interlocking of the blocks and the matching colors give immediate feedback as to possibly correct combinations. By default, definitions of words are hidden on the blocks, but available with a simple right-click. This decision guides the student to examine the structure of the sentence first, only to consider "looking up" words and to rely on vocabulary as a secondary thought. Ingenium not only improves students' understanding and translating of

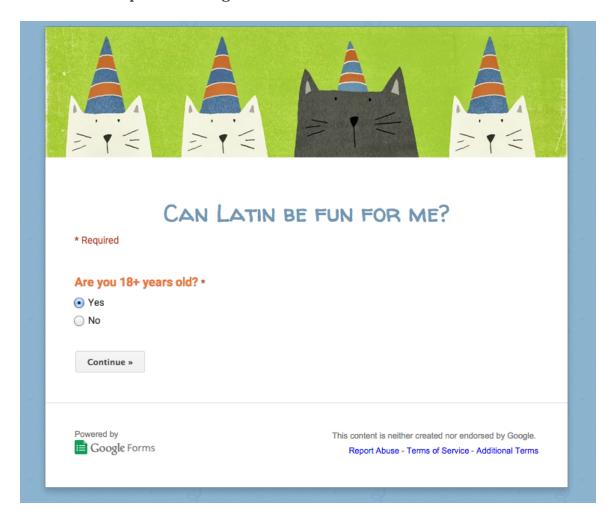
Latin sentences through its visual emphasis on grammatical concepts and structures, but also increases student engagement by providing a dynamic environment for experimentation. The further increase in engagement found in the system's external interface, which departs more from natural word order, reveals that students enjoy restructuring the sentence in a more schematic and less linear way, finding the activity both more helpful and more exciting.

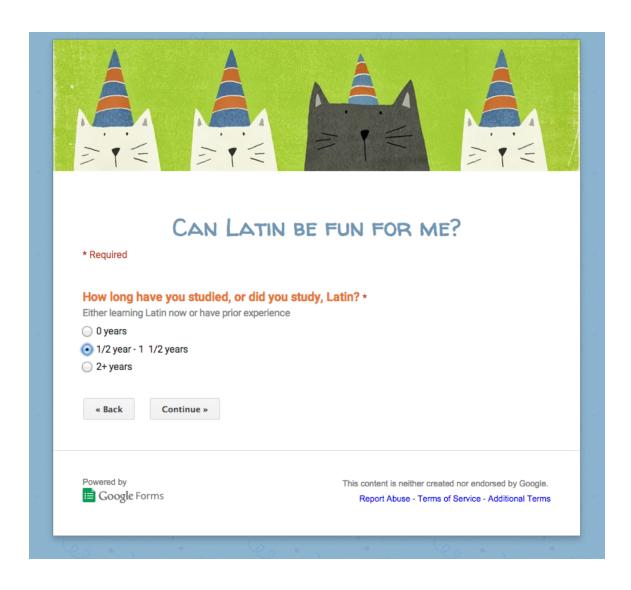
Ingenium has demonstrated its potential to be a powerful tool in improving students' engagement, learning, and accuracy in reading Latin. Motivated by existing methods to address the problems in learning Latin and supported by theory in computer science, psychology, and education, Ingenium has the ability to fundamentally change the way students approach and interact with Latin grammar. Since it does not presuppose or establish a new set of terminology of its own and can be accessed online, Ingenium serves as an easily adoptable piece of supplementary instructional material for Latin classrooms worldwide.

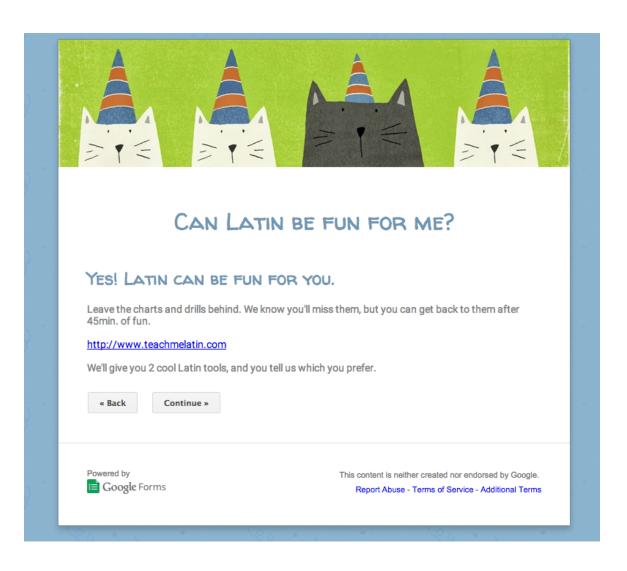
Appendix

Form Questions

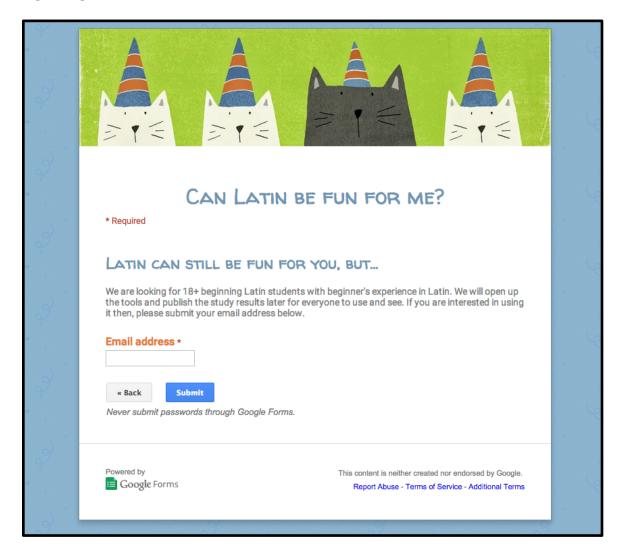
Remote Participant Filtering







If participant was not 18+ or a novice:



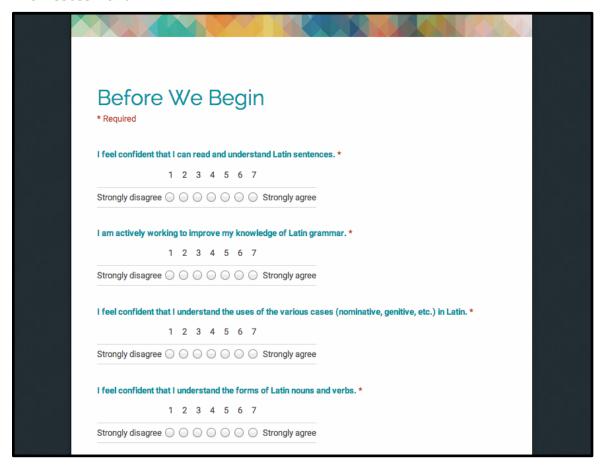
Basic Demographics

Basic Demographics & Latin Background *Required	d
Age *	100000
○ 18-22	
○ 23-35	1000
○ 36-50	200000
○ 51+	1607 (0) (0)
Gender *	
○ M	576-777-783
○ F	63873734
Other	
In School: how long have you studied (or did you study) Latin?*	
○ 0 years	200000
○ 1/2 year	1507 207 207
○ 1 year	67676 657
○ 11/2 years	1000000000
Other:	1444
At what school level(s)?* Check all that apply.	
☐ Middle school	
☐ High school	100300
☐ College / University	ALVA VE
Other:	

Out of school: how long have you studied (or did you study) Latin? * Any self-learning or informal instruction	
○ 0 years	
○ 1/2 year	
1 year	
11/2 years	
Other:	
When did you start learning Latin? * MM/YYYY	
O Aug/Sept 2014 (08/2014 or 09/2014)	
○ Jan 2014 (01/2014)	
Other:	
What textbook(s) did you use?* Check all that apply.	
☐ Introduction to Latin by Shelmerdine	
☐ Ecce Romani	
☐ Learn to Read Latin by Keller and Russell	
☐ Traditio by Patricia Johnston	
☐ Latin: An Intensive Course by Moreland and Fleischer	
☐ Wheelock's Latin	
Oxford Latin Course	
☐ Cambridge Latin Course	
☐ Jenney's	
Other:	
Current chapter in your textbook, if you're enrolled in Latin now. eg. Chapter 8: 3rd Declension i-stems	

University *	Marie Care Marie
Agnes Scott College	
Brandeis University	
Brown University	
○ College of the Holy Cross	
○ College of William & Mary	
○ Columbia University	
Orew University	
George Washington University	
○ Gordon College	
○ Harvard College	
Harvard Extension School	
O John Hopkins University	
Northwestern University	
○ Skidmore College	
Stanford University	
○ Trinity College	
○ Tufts University	
University of Massachusetts Boston	
○ University of Michigan	
 University of North Carolina at Greensboro 	
○ Yale University	Wash data Black
Other:	
Year in University *	
○ Freshman	
○ Sophomore	
○ Junior	
○ Senior	
Graduate student	
○ Independent learner	
Other:	
Please do not change the text below * Verifies your completion of this survey after submission without storing a personally identifiable name (so that we can compensate you later).	
Submit	
Never submit passwords through Google Forms. 100%: You made it.	

Pre-Assessment



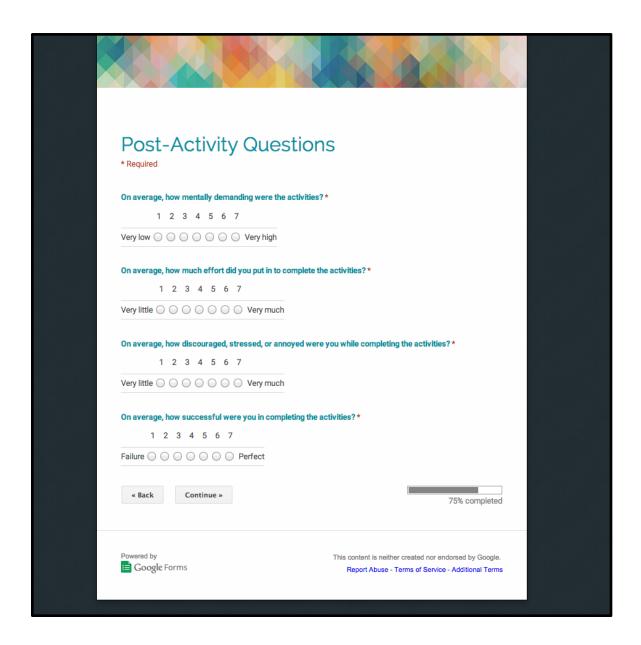
	1	2	3	4	5	6	7			
Strongly disagre	ee 🔾	0	0 (\bigcirc	0	0	0	Strongly agree		
I feel confident	that I	und	ersta	nd	l the	COI	nce	pt of noun-adjec	ctive agreement. *	
	1	2	3	4	5	6	7			
Strongly disagre	ee 🔾	0	0)	0	0	0	Strongly agree		
Please do not co							fter	r submission.		
	mplet	ion (of this	SS	urve	еу а				

Post-Assessment (identical for each condition)

Post-Activity Questions *Required	
How much have you learned about the uses of the various cases (nominative, genitive, etc.) in Latin? *	
1 2 3 4 5 6 7	
Not at all OOOO Substantially	
How much have you learned about the forms of Latin nouns and verbs?*	10/25
1 2 3 4 5 6 7	
Not at all O O O Substantially	
How much have you learned about the concepts of person and number of verbs?*	
1 2 3 4 5 6 7	Maria L
Not at all O O O Substantially	
How much have you learned about the concept of noun-adjective agreement? *	
1 2 3 4 5 6 7	KAMES
Not at all OOOO Substantially	
How well do you think you would do if you were given an exam now to measure your retention of the	
concepts from these activities?*	
1 2 3 4 5 6 7	
Not at all OOOO Substantially	
These activities are generally helpful to me in my current Latin studies. *	
1 2 3 4 5 6 7	
Not at all O O O Substantially	Marine 1

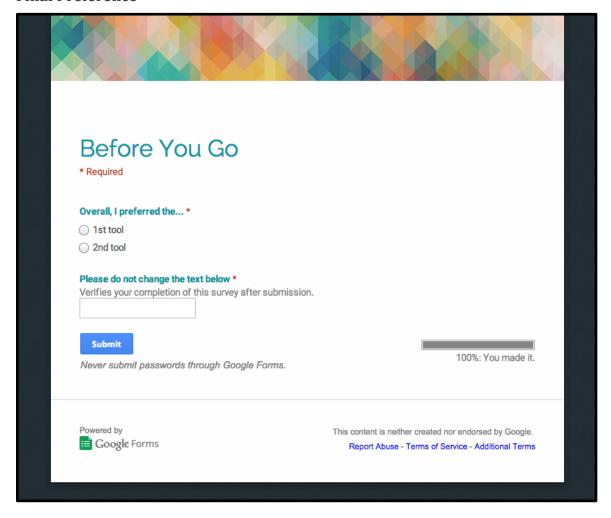
1	2	2 3	4	5	6	7					
Not at all) () ()	0	0	0	0	Substantially				
These activ	rities	s wo	uld h	nave	bee	n m	ore helpful to	me earlier in my Latin st	tudies than nov	w. *	
1	2	2 3	4	5	6	7					
Not at all) () ()	0	0	0	0	Substantially				
Continue	è »									25%	completed
owered by								This content is	neither created r	nor endorsed	by Google.

Post-Ac * Required	tivity	Qu	estion	S	
I feel confident that I	I can read and	d underst	and Latin senter	ices.*	
1	2 3 4	5 6 7			
Strongly disagree	0000	000	Strongly agree		
I am actively working				ammar. *	
Strongly disagree	2 3 4				
on ongry disagree			outorigiy agree		
	l understand t			ses (nominative, genit	ive, etc.) in Latin. *
Strongly disagree					
I feel confident that I	2 3 4			and verbs. *	
Strongly disagree	0000	000	Strongly agree		
I feel confident that	Lunderstand	the conce	ents of person a	nd number of verbs. *	
	2 3 4				
Strongly disagree	0000	000	Strongly agree		
I feel confident that I	I understand f	the conce	ept of noun-adject	ctive agreement. *	
1	2 3 4	5 6 7			
Strongly disagree	0000	000	Strongly agree		
« Back Co	ontinue »				
- Back Co	manue »				50% completed

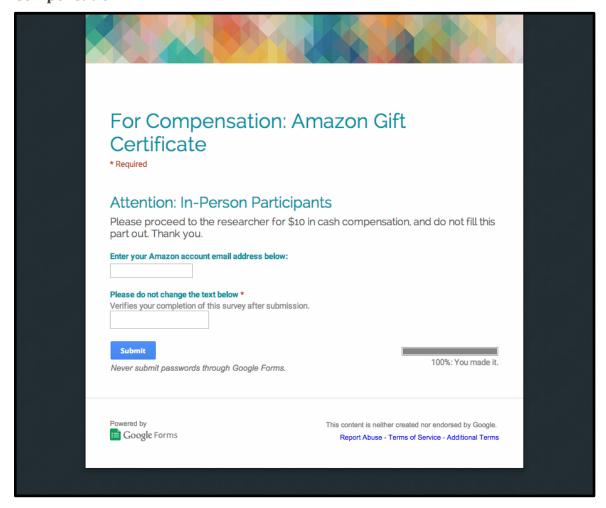


Post-Activity C	uestions		
I enjoyed doing the translation exercis	es. *		
1 2 3 4 5	5 7		
Strongly disagree O O O O) C Strongly agree		
l enjoyed doing the fill-in-the-blank ex			
Strongly disagree O O O O	Strongly agree		
When I translate a Latin sentence, I re 1 2 3 4 5 6 Strongly disagree	5 7		
I found it helpful for my learning to do			
Strongly disagree O O O O	Strongly agree		
I found it helpful for my learning to do 1 2 3 4 5 Strongly disagree	5 7	s.*	
Please do not change the text below * Verifies your completion of this survey			
« Back Submit Never submit passwords through God		100%: Y	ou made it.

Final Preference

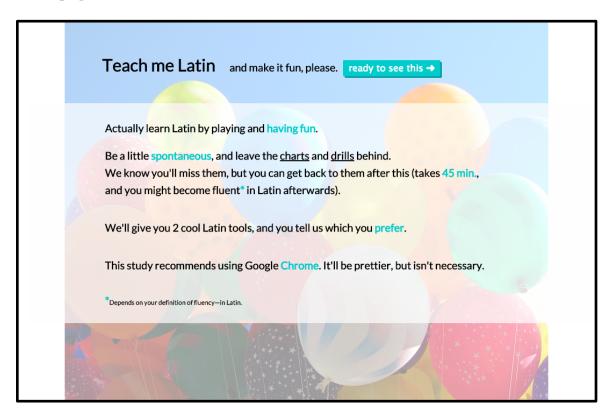


Compensation



Flows of the Experiment

Homepage



Consent Form

Ready! check 'I agree' and...

✓ I agree let's go →

Form of Consent: please read and check 'I agree' above to begin!

Study Title: Efficacy of Latin Grammar Tools for Language Instruction: An Undergraduate Computer Science and Classics Thesis

Researcher: Sharon Zhou

Version Date: February 12, 2015

Participation is voluntary

It is your choice whether or not to participate in this research. If you choose to participate, you may change your mind and leave the study at any time. Refusal to participate or stopping your participation will involve no penalty or loss of benefits to which you are otherwise entitled

What is the purpose of this research?

The purpose of this research is to study the efficacy of different Latin grammar tools on student learning, engagement, and confidence in learning Latin grammar.

How long will I take part in this research?

What can I expect if I take part in this research?

As a participant, you will be introduced to two Latin grammar tools and be asked to perform translation and fill-in-the-blank exercises with each. Before beginning, you will be asked to provide basic demographic information and given a pre-assessment questionnaire. After each tool, you will be asked a post-assessment questionnaire. After both tools, you will be asked which tool you preferred.

What are the risks and possible discomforts?

If you choose to participate, there are no or minimal foreseeable risks or discomforts.

Are there any benefits from being in this research study?

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits include introducing new Latin grammar study tools to the Latin classroom.

Will I be compensated for participating in this research?

download form of consent

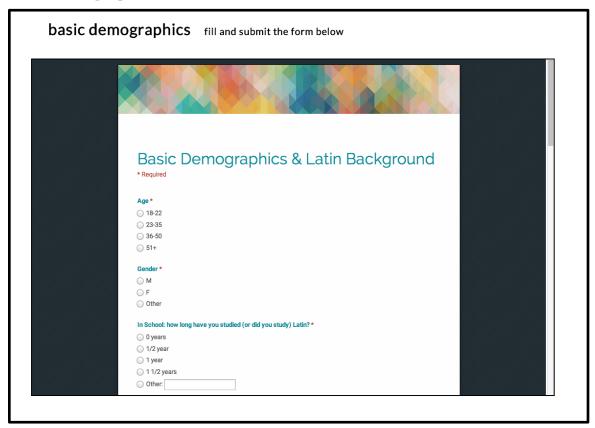
General Directions

Coming up:

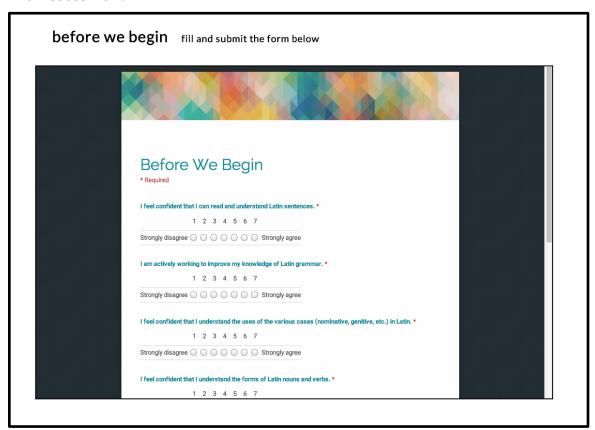
- 1. You'll be asked basic demographics and about your general Latin experience.
- 2. You'll be shown two different designs of Latin grammar tools.
- 3. After each, you'll be asked questions about using it.
- 4. At the end, you'll be asked which you preferred.
- 5. If you are online, you'll be asked for your Amazon account email address, for \$10 on Amazon.
- 6. It will all take about 45 minutes. You will have fun.

let's go →

Basic Demographics



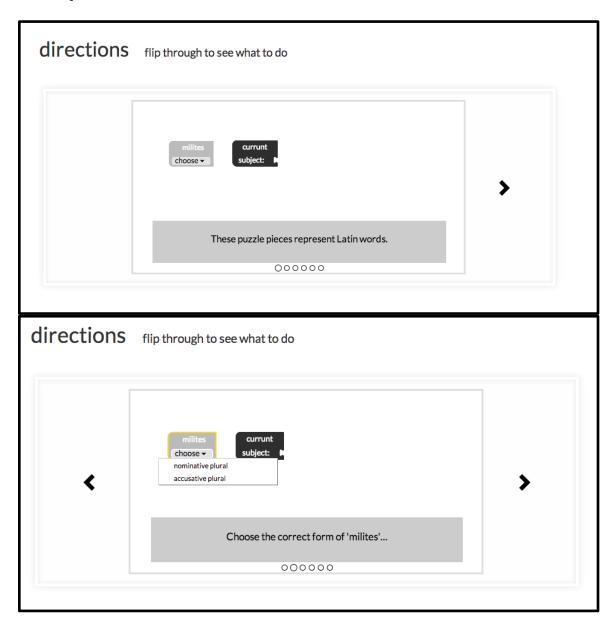
Pre-Assessment

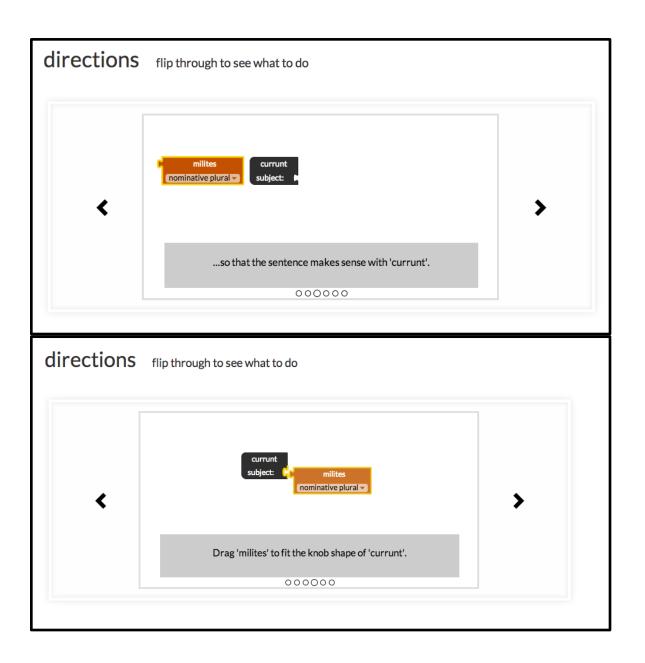


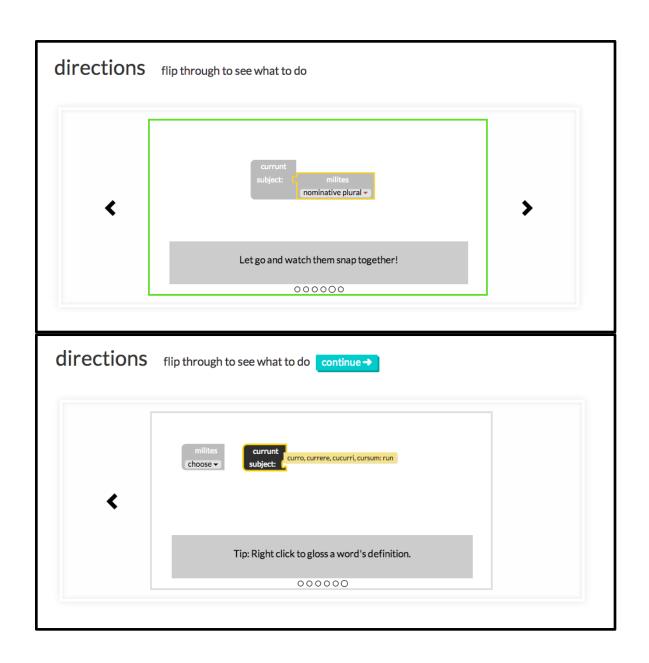
Experimental Condition

External Interface

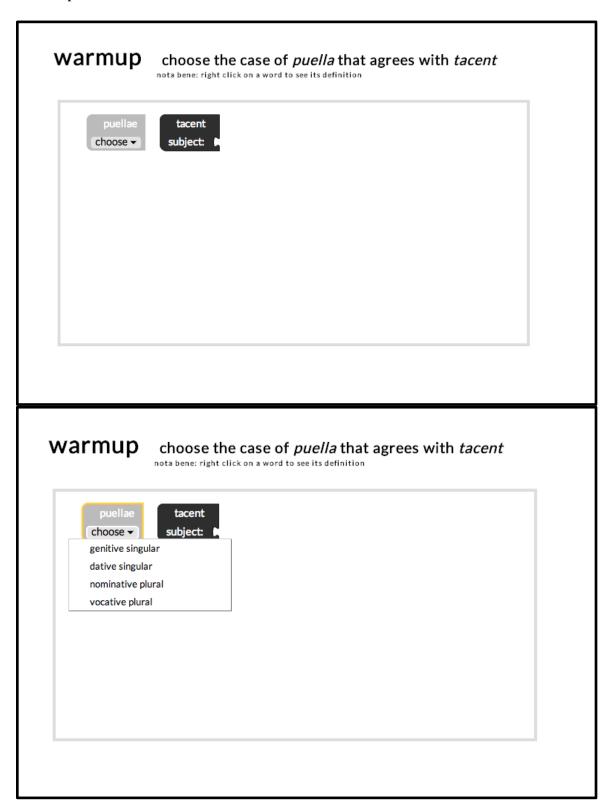
Warmup - Directions



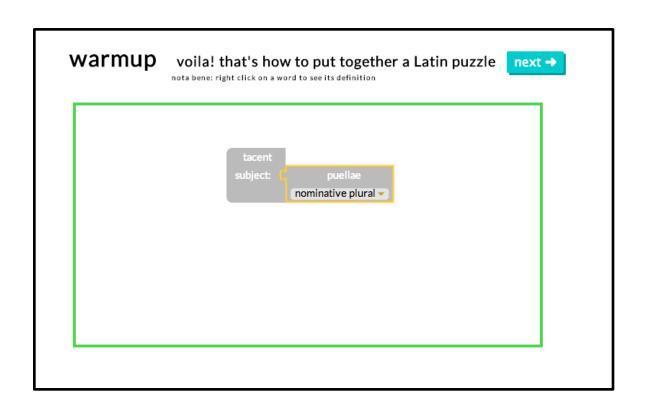




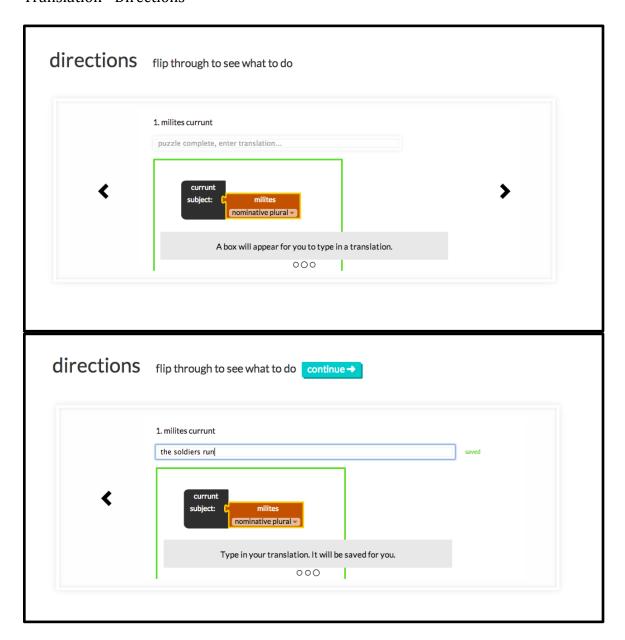
Warmup - 1



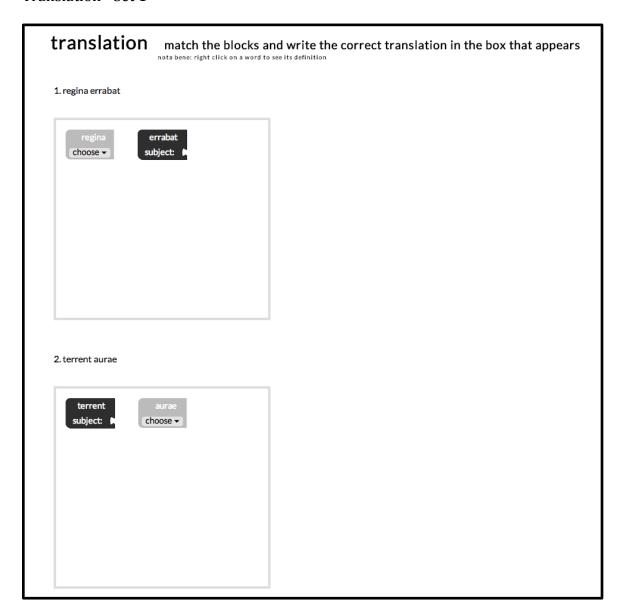
warmup drag and drop *puellae* to match the knob shape of *tacent* nota bene: right click on a word to see its definition puellae nominative plural 🔻 tacent subject: warmup drag and drop puellae to match the knob shape of tacentnota bene: right click on a word to see its definition tacent subject: nominative plural -

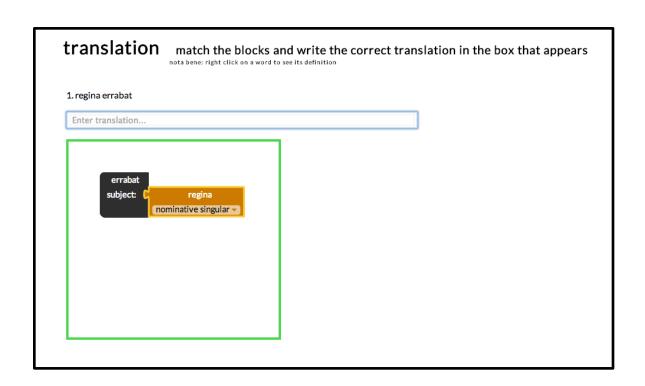


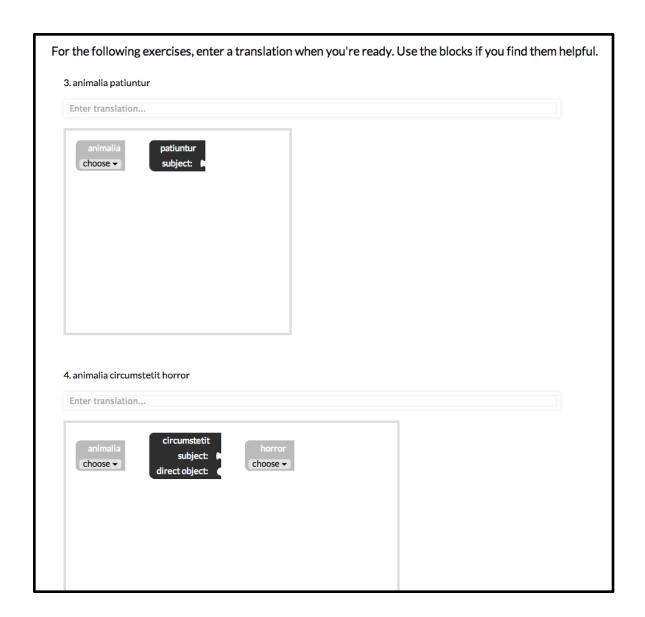
Translation - Directions



Translation - Set 1

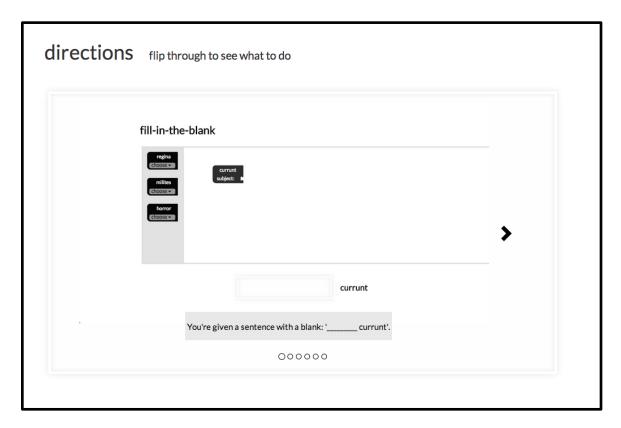


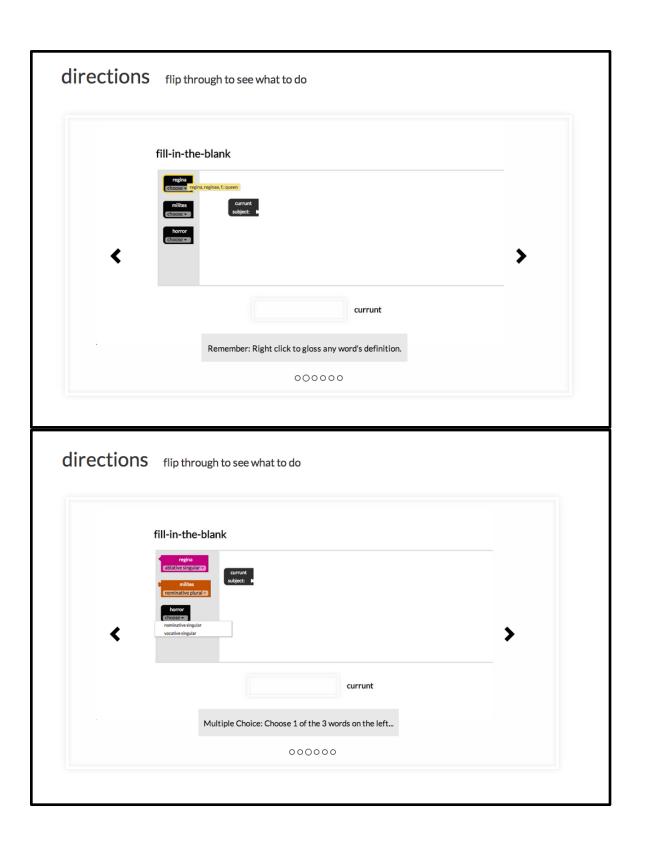


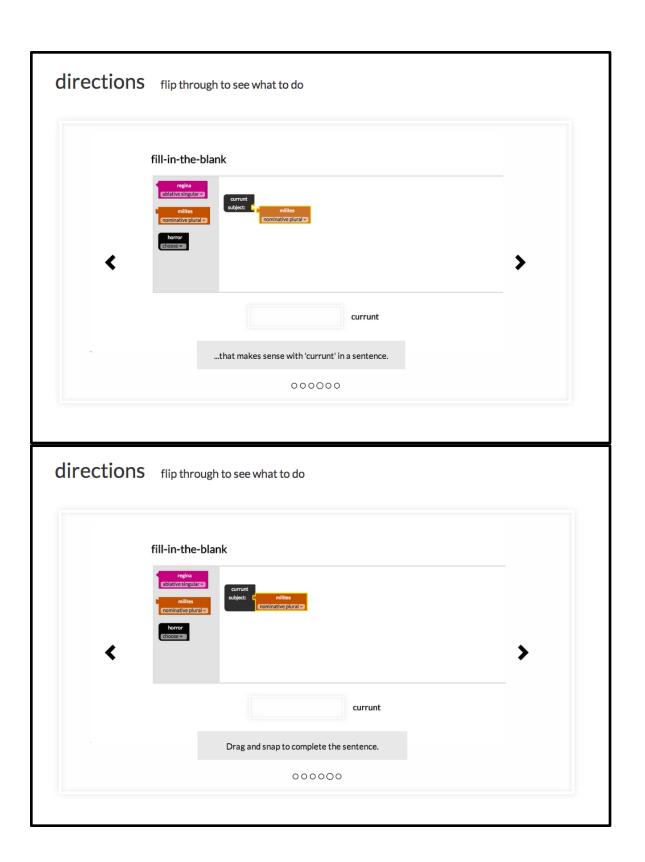


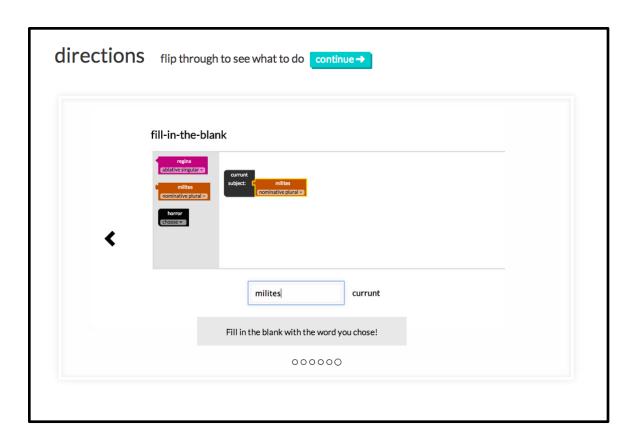


Fill-in-the-Blank - Directions



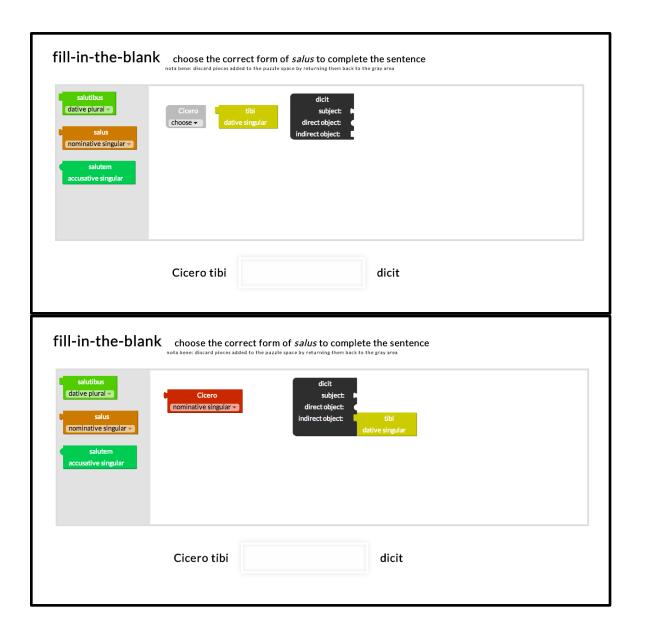


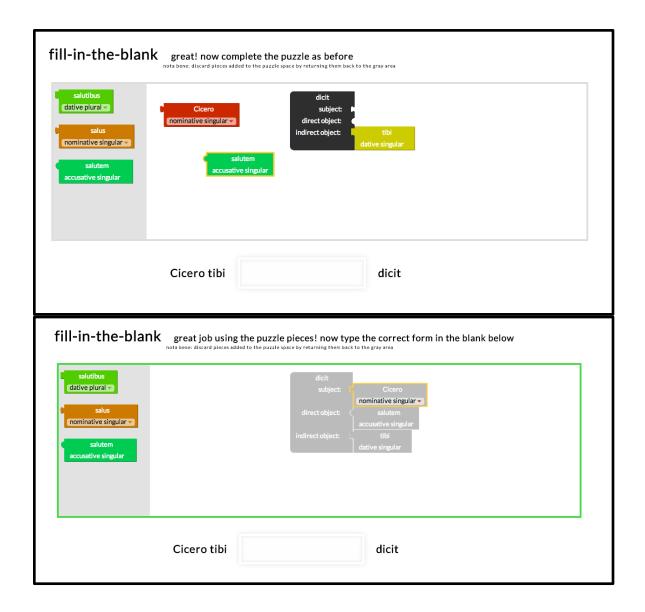


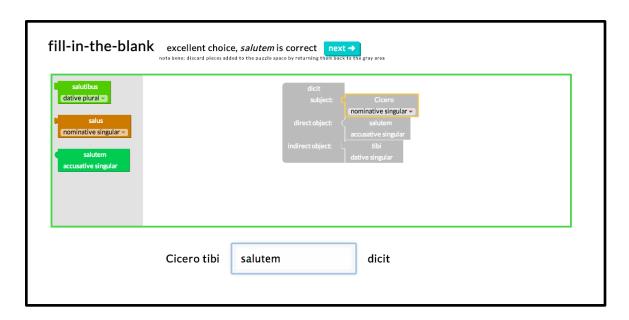


Fill-in-the-Blank - 1



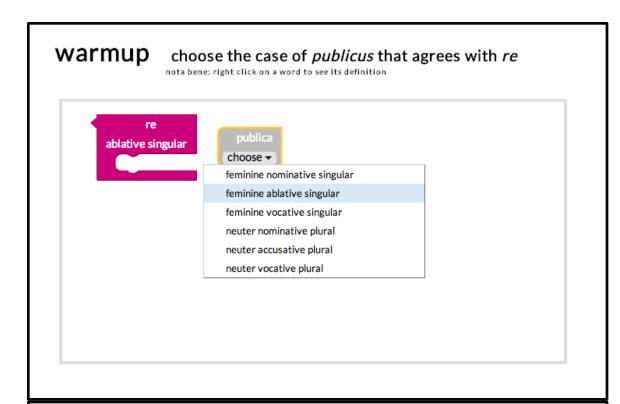


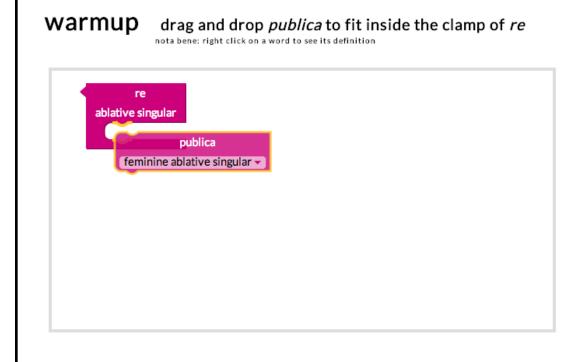


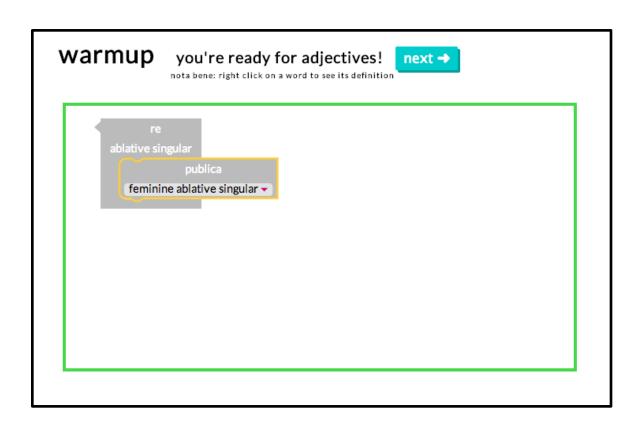


Warmup - 2

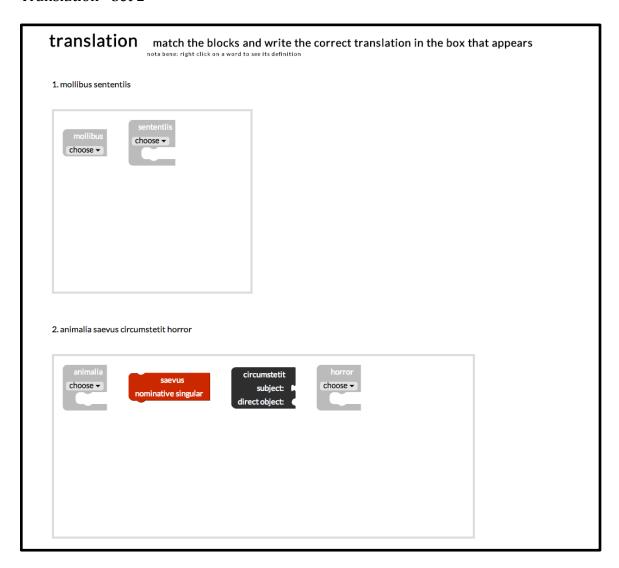


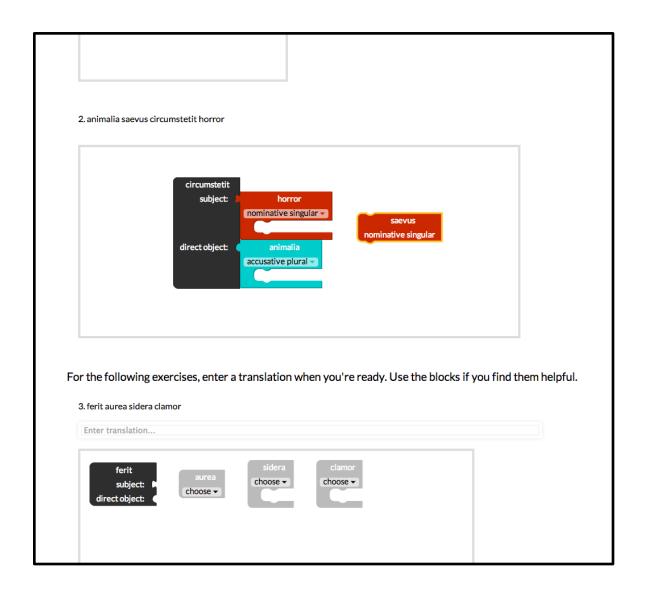


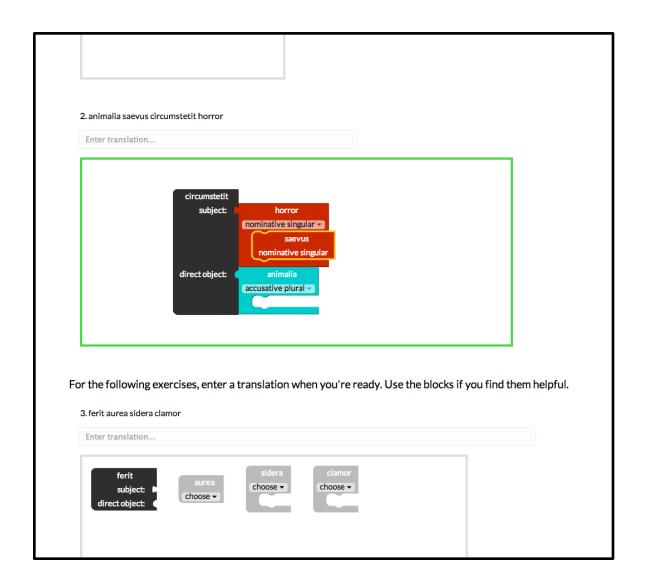


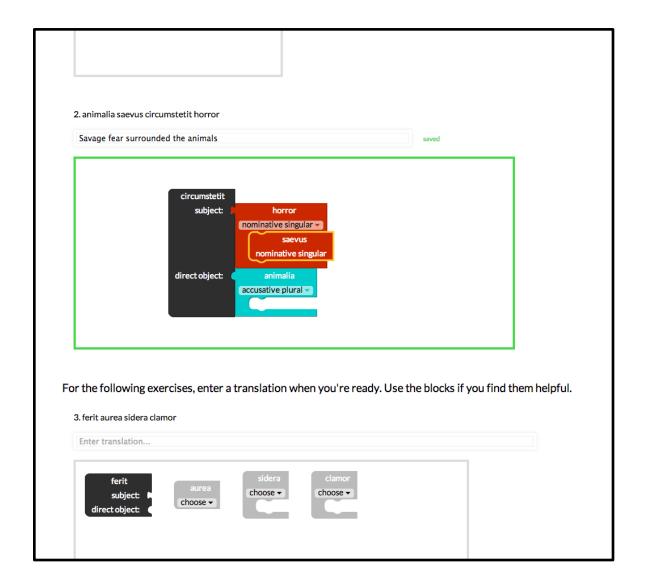


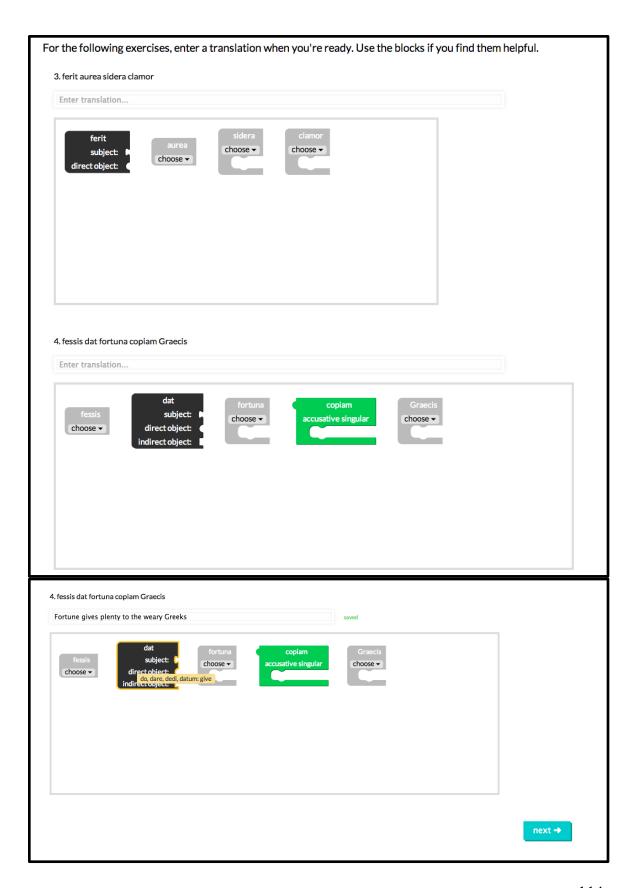
Translation - Set 2



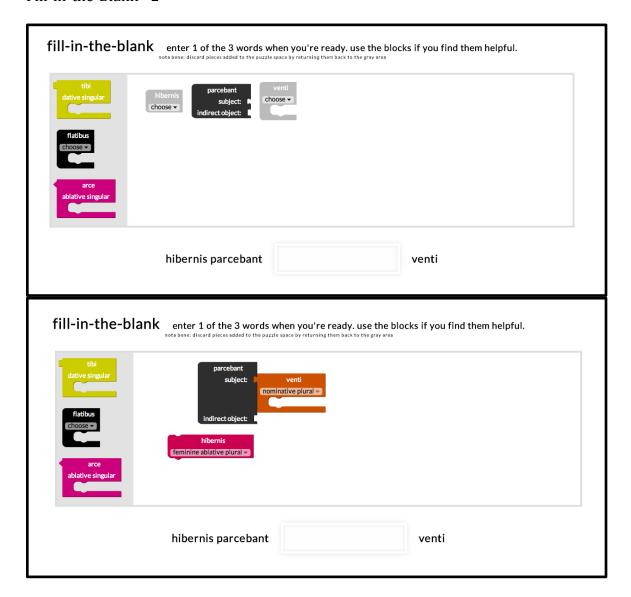


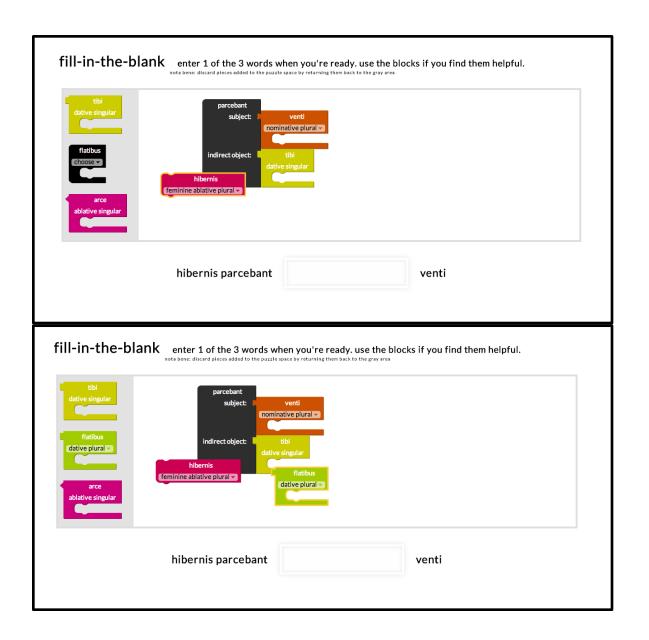


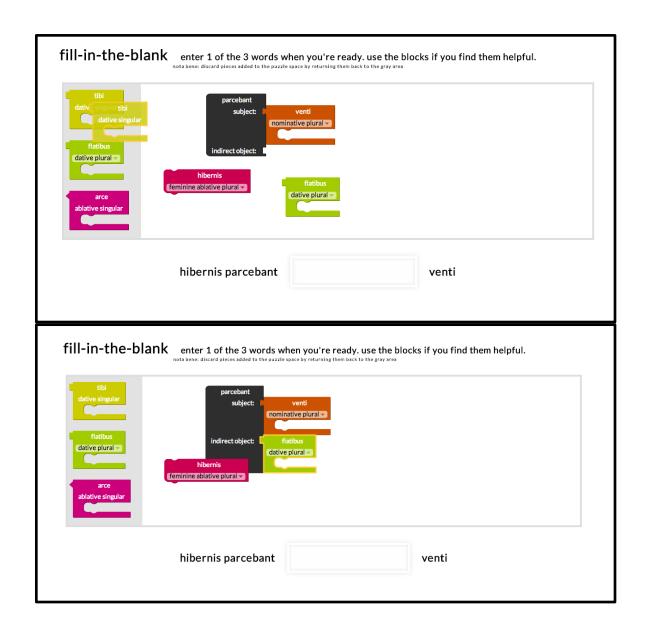


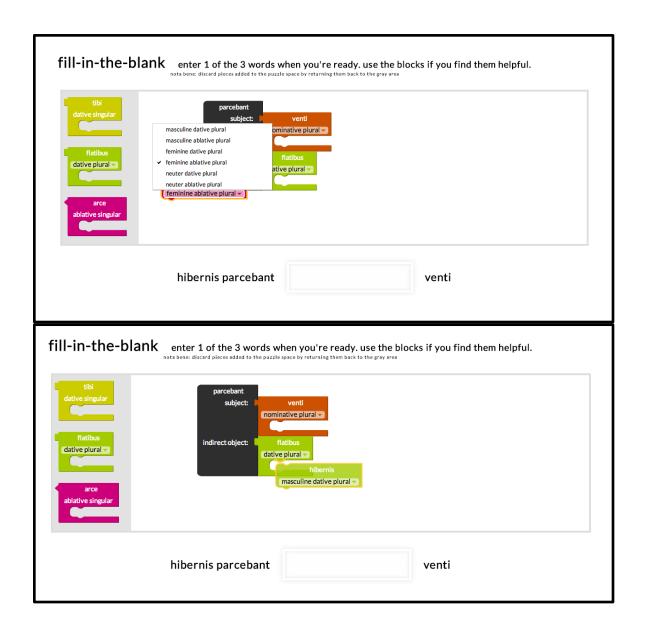


Fill-in-the-Blank - 2



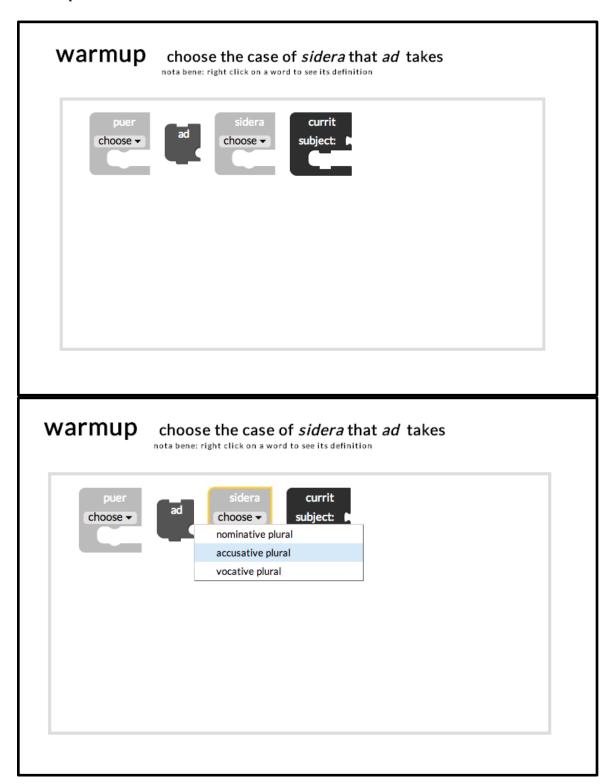


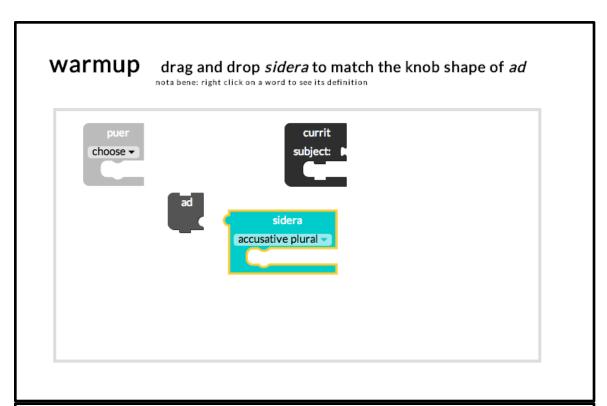


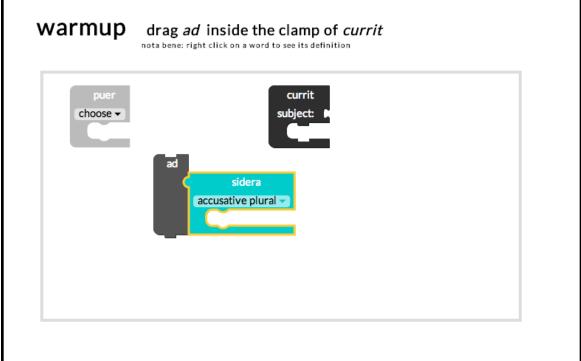


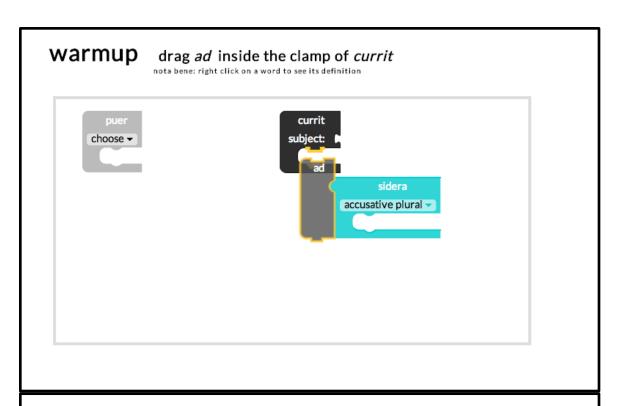


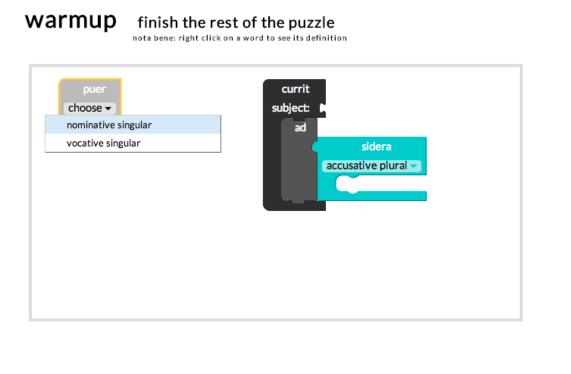
Warmup - 3

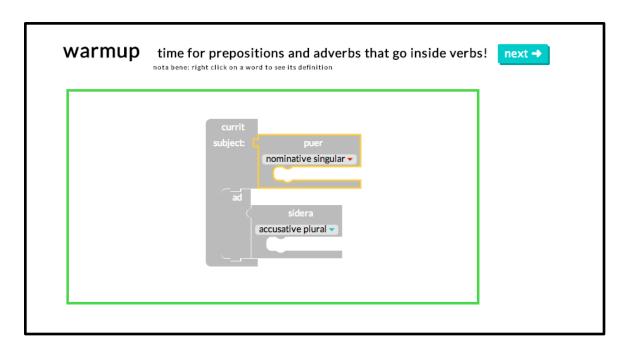




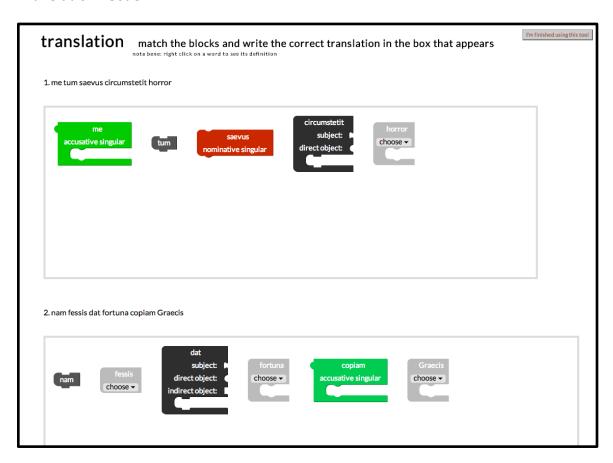


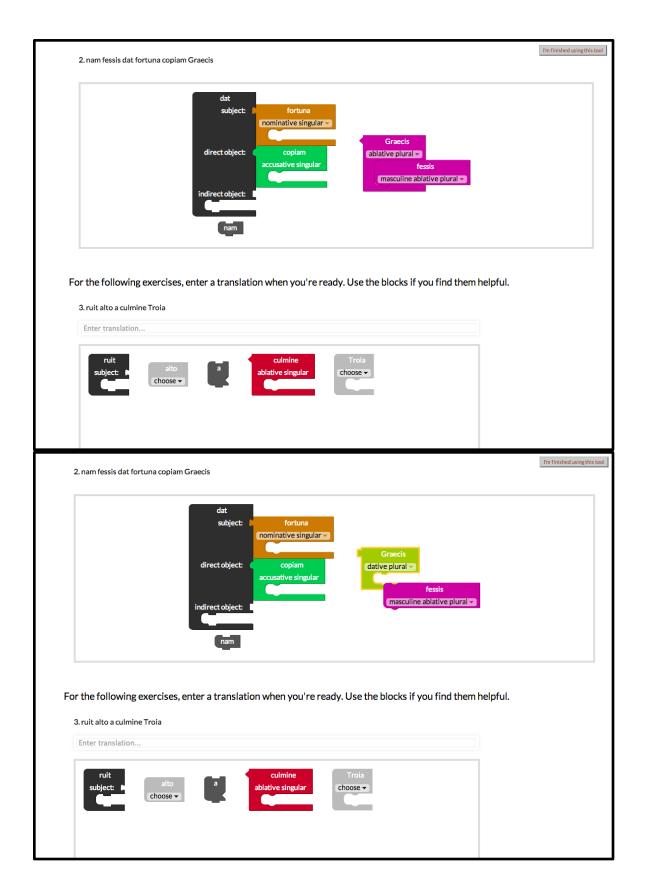


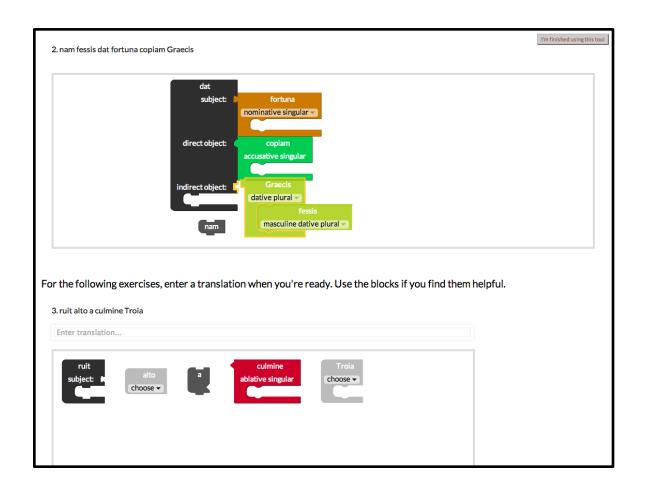


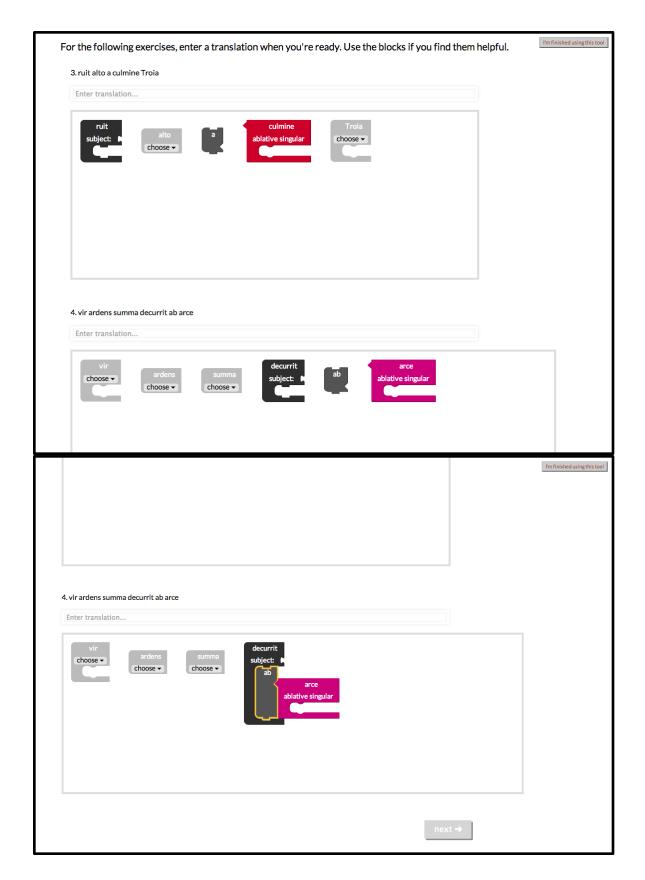


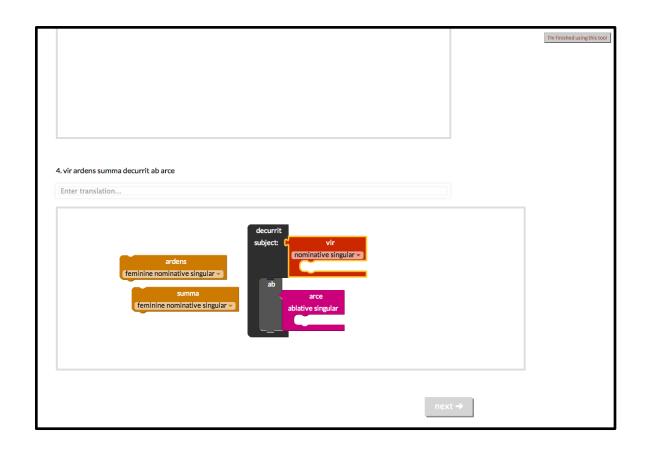
Translation - Set 3

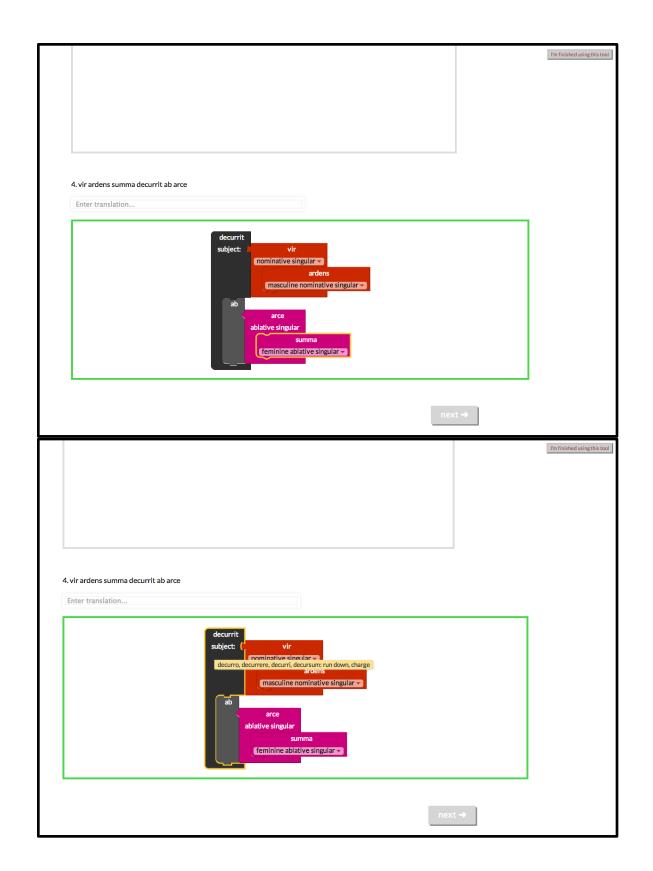






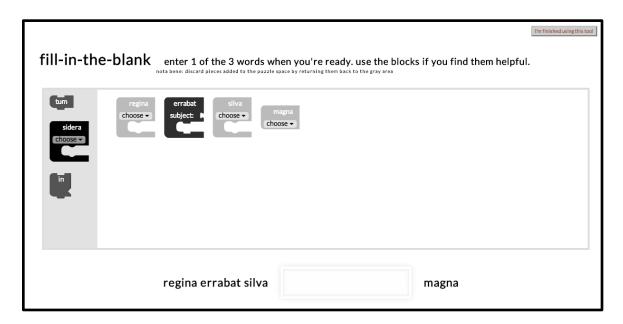




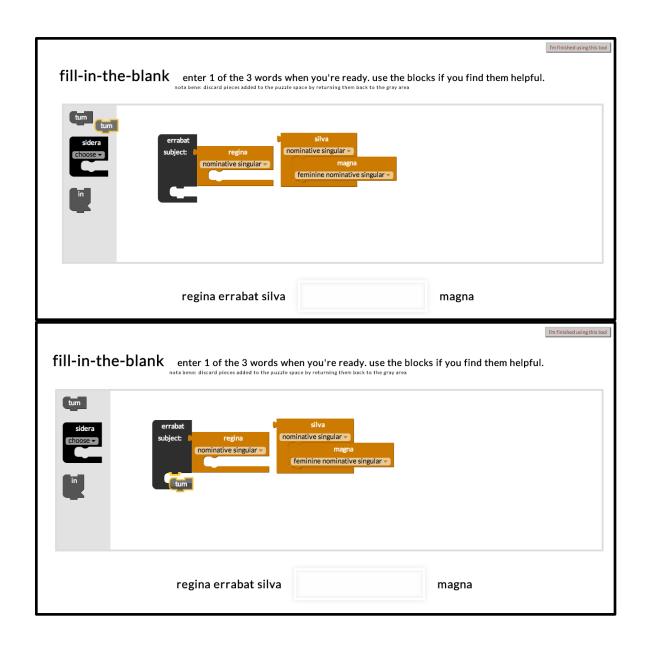


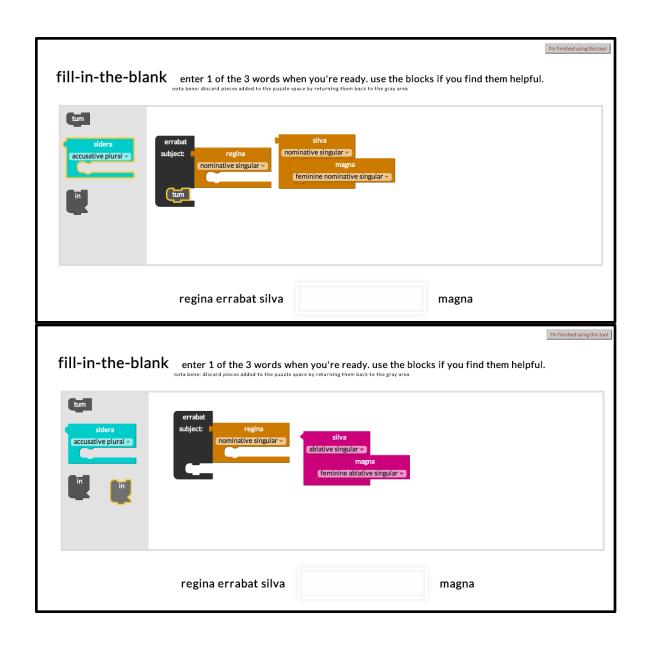
	I'm finished using this	tool
4. vir ardens summa decurrit ab arce The eager man runs down the highest fortress decurrit subject vir nominative singular arce ablative singular summa feminine ablative singular summa feminine ablative singular		
next	t →	

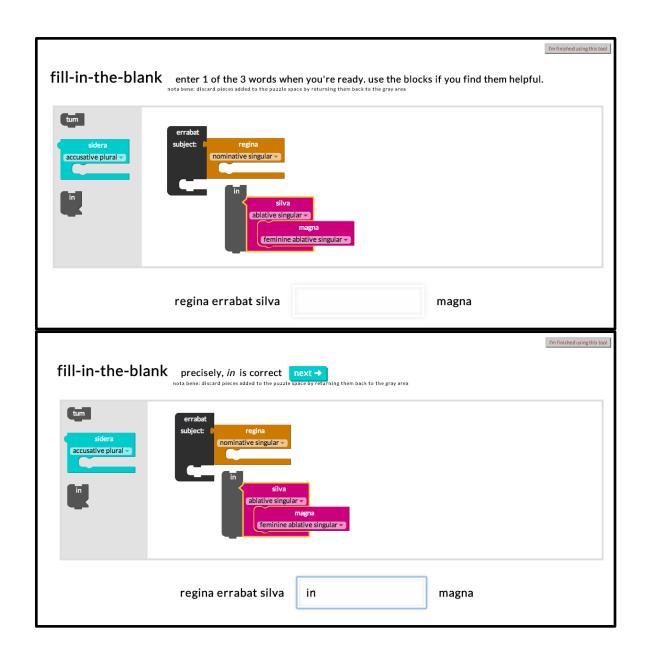
Fill-in-the-Blank - 3







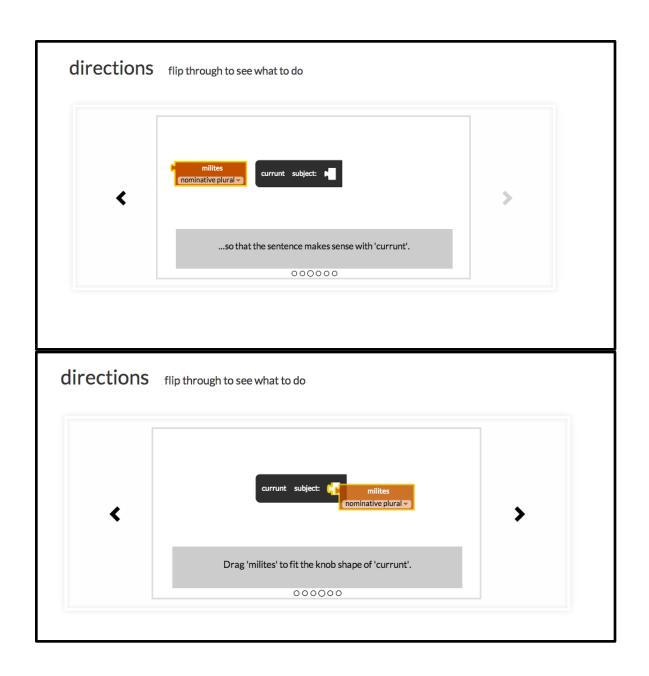


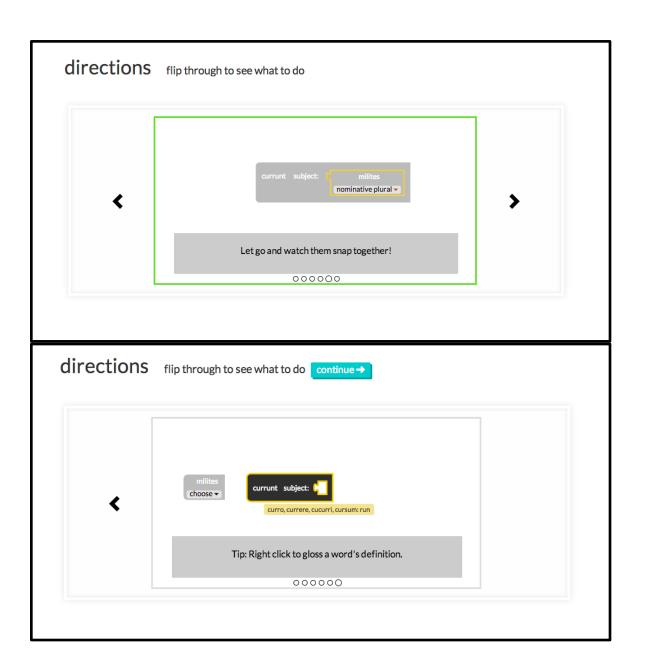


Inline Interface

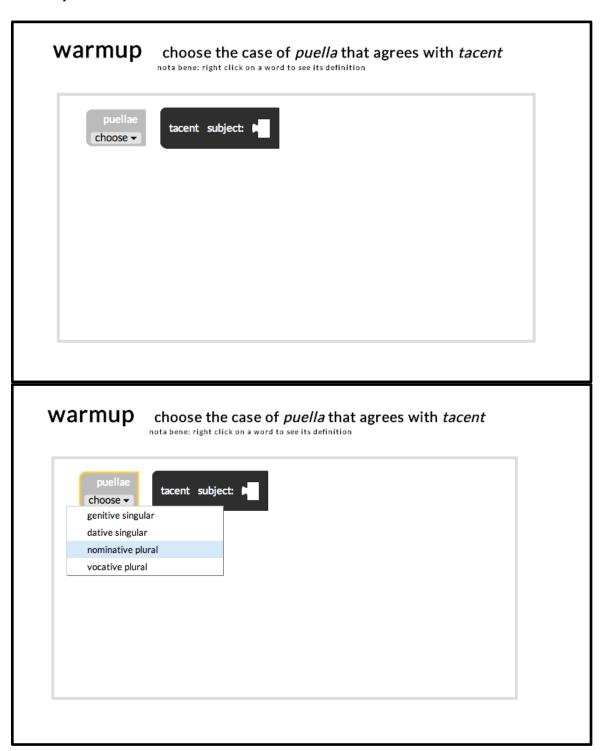
Warmup - Directions

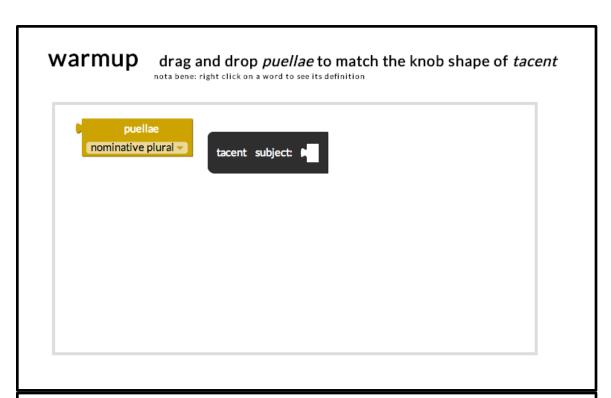


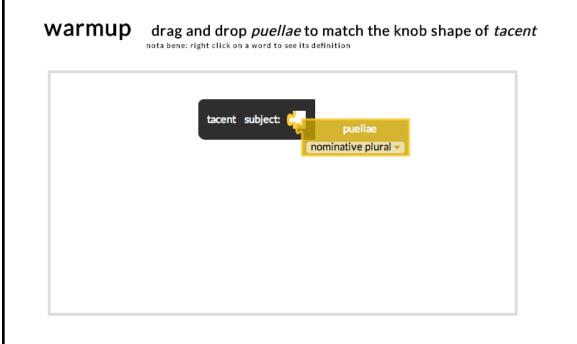




Warmup - 1

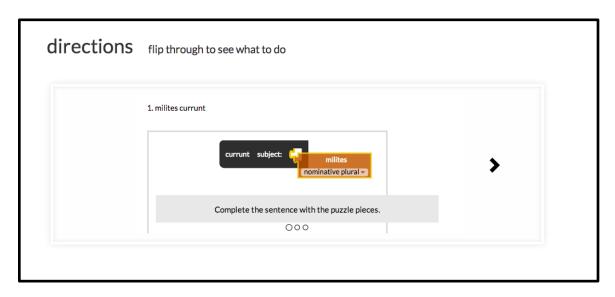


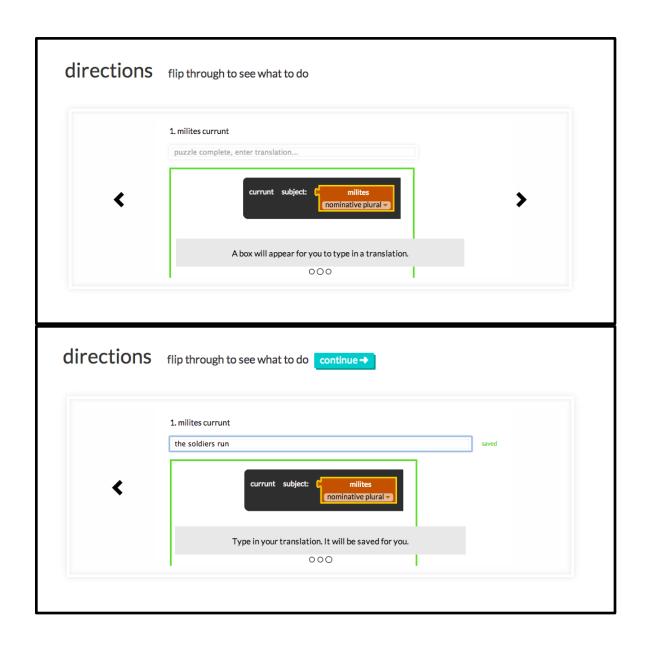


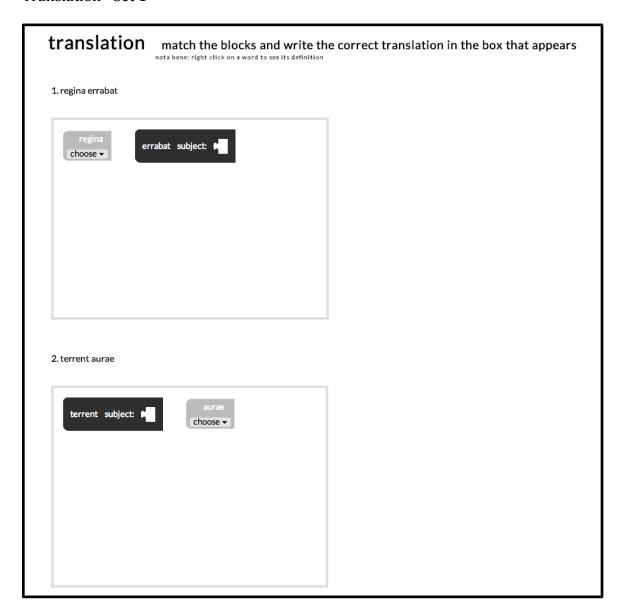


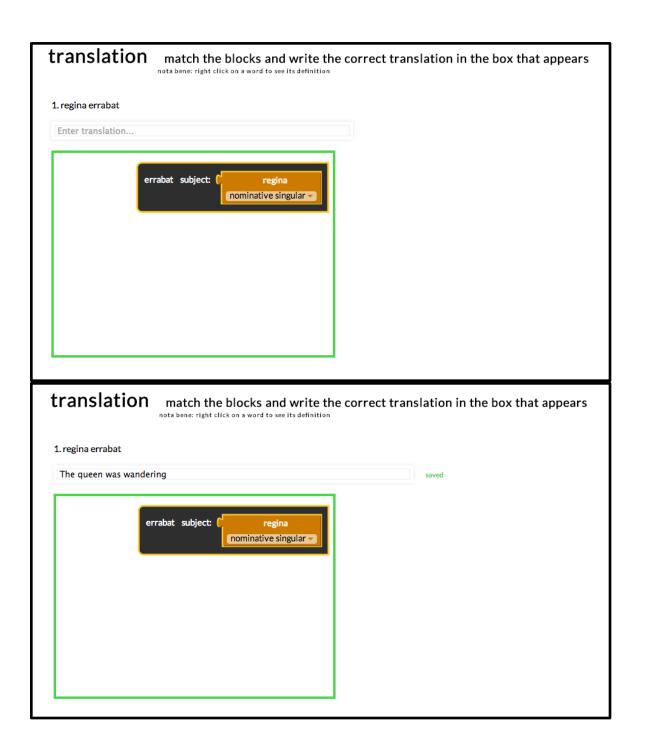
warmup	voila! that's how to put together a Latin puzzle next →
	tacent subject: puellae nominative plural

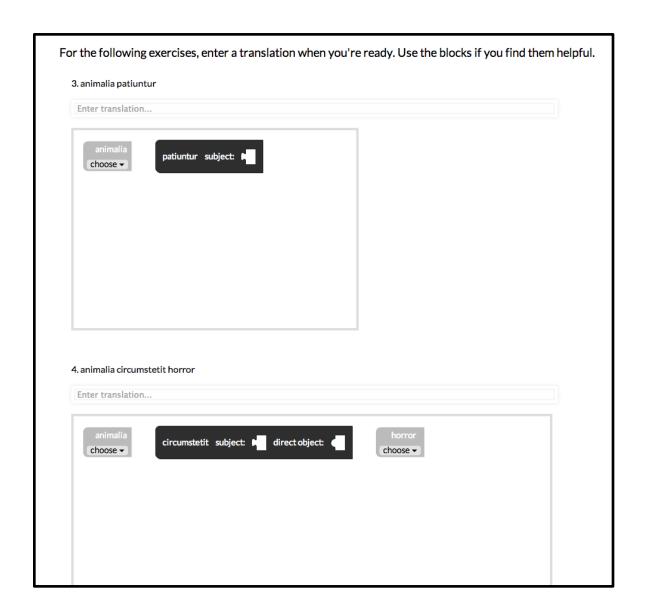
Translation - Directions

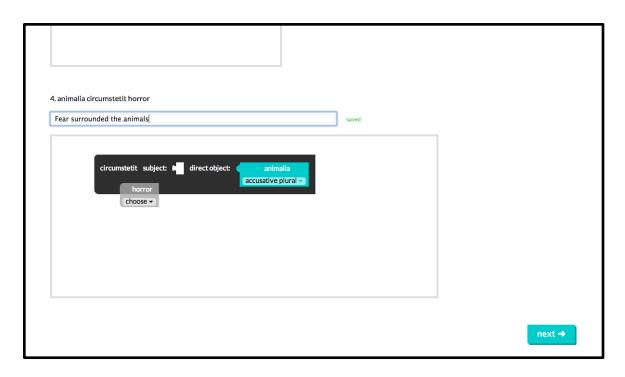




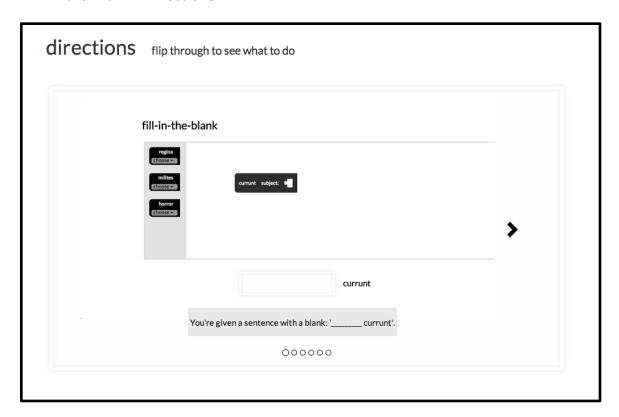


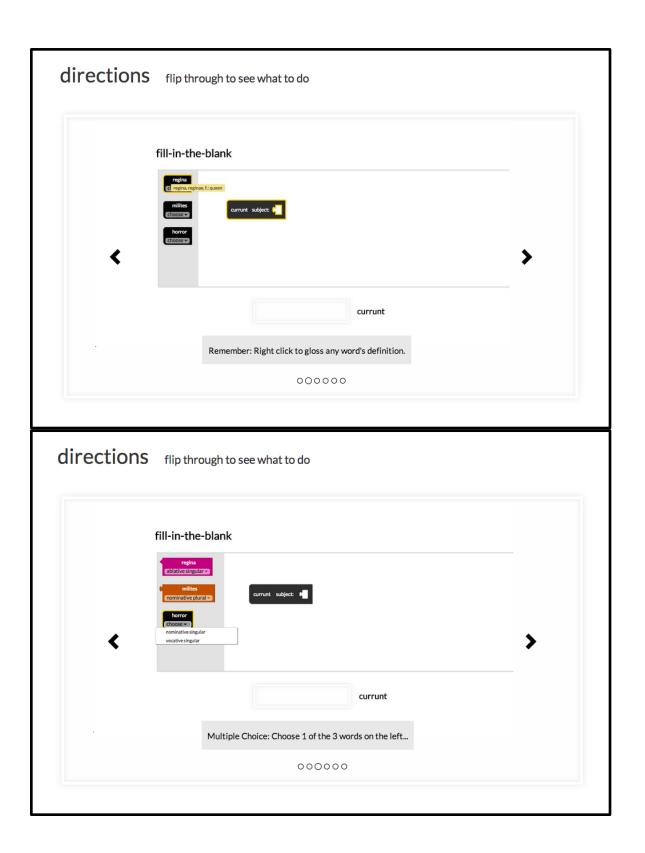


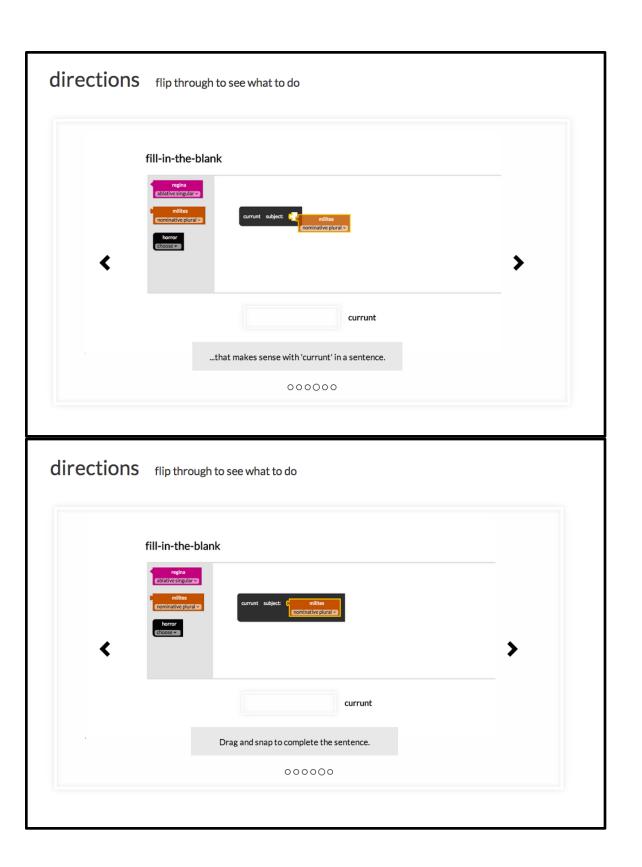


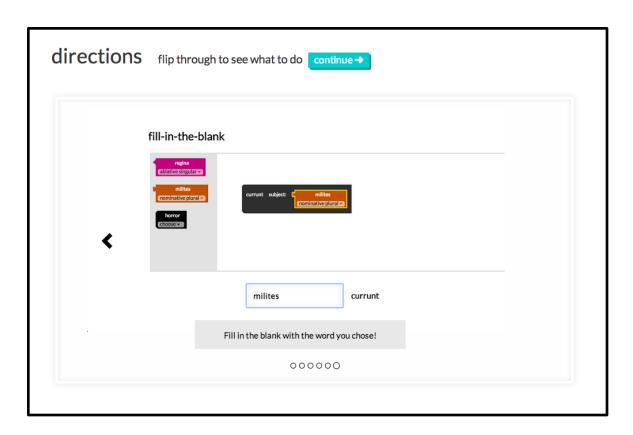


Fill-in-the-Blank - Directions

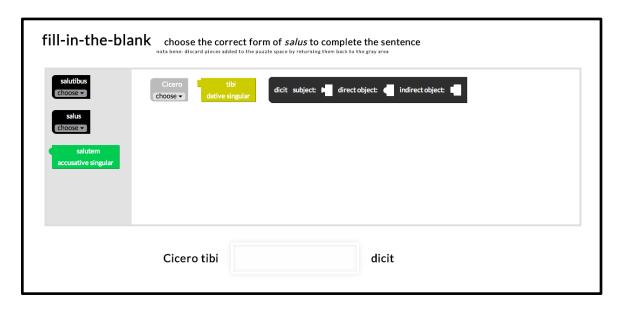


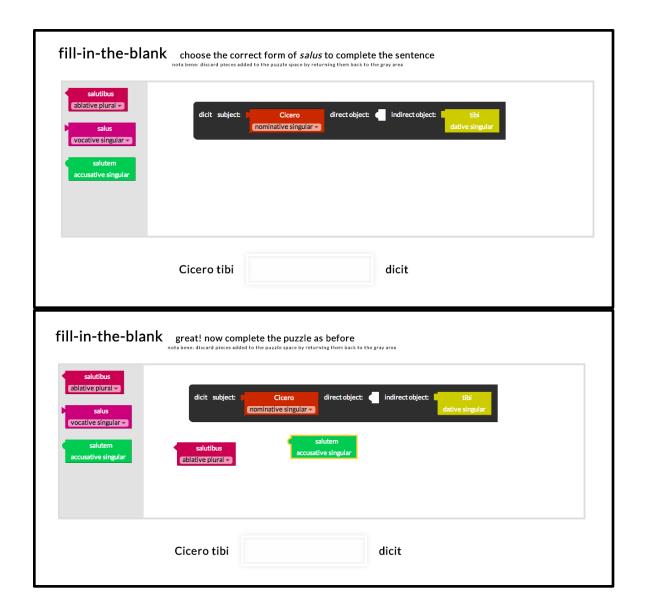


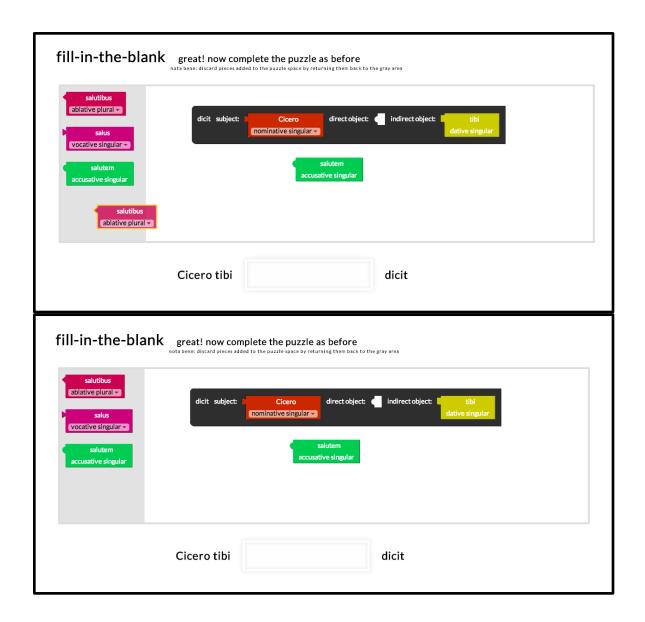


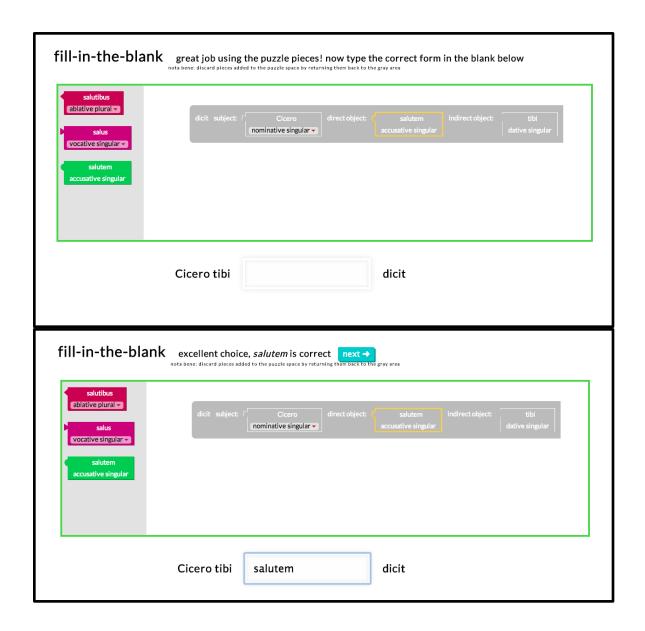


Fill-in-the-Blank - 1

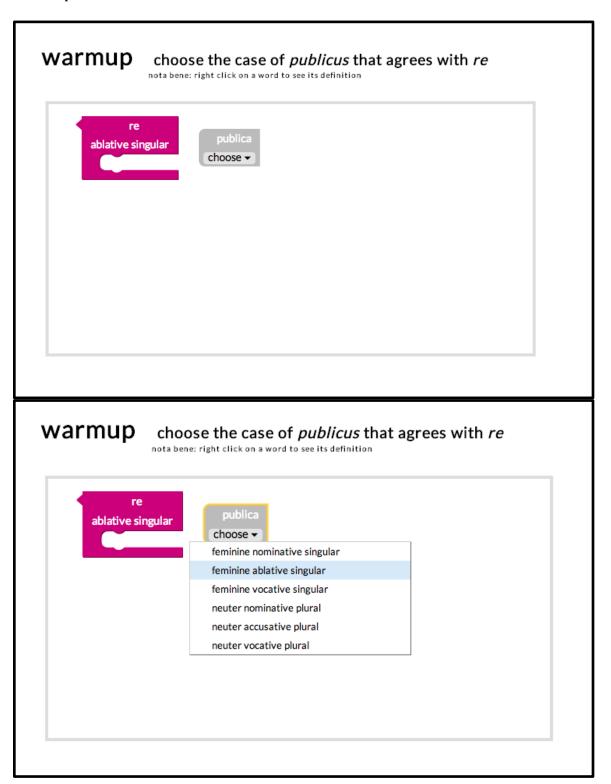


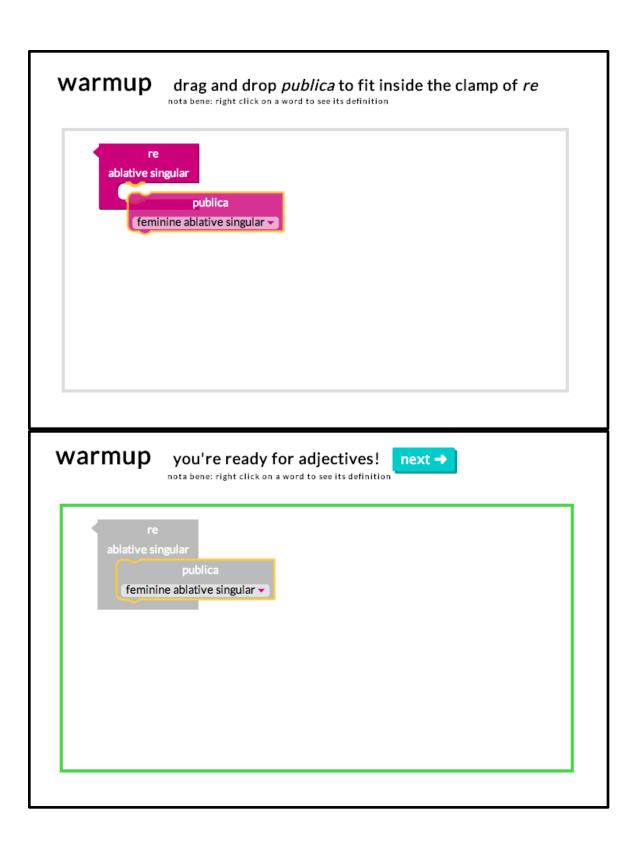




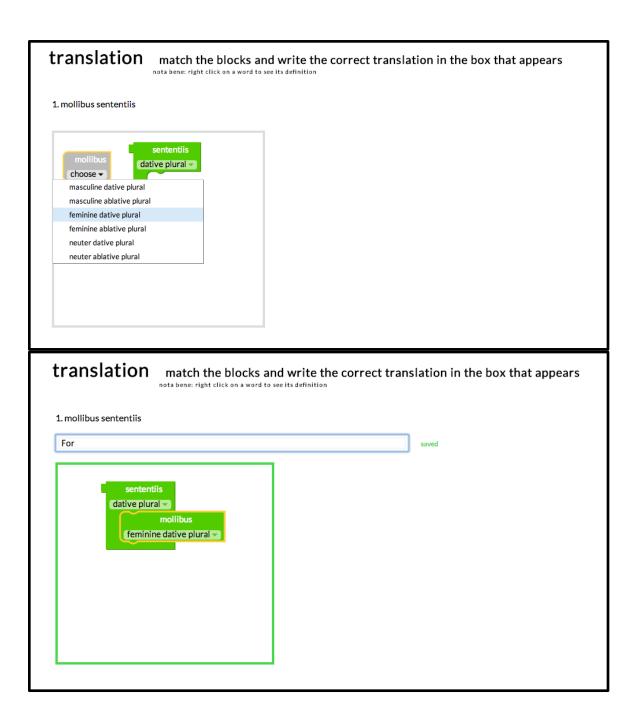


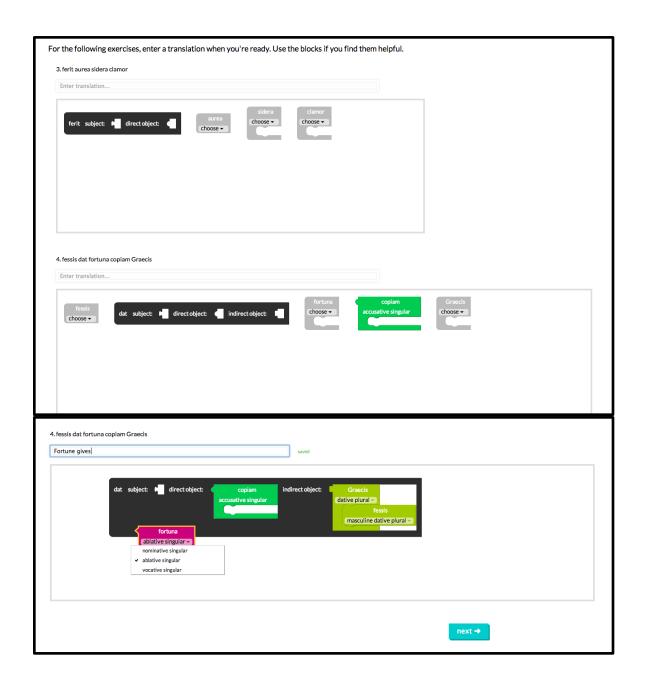
Warmup - 2



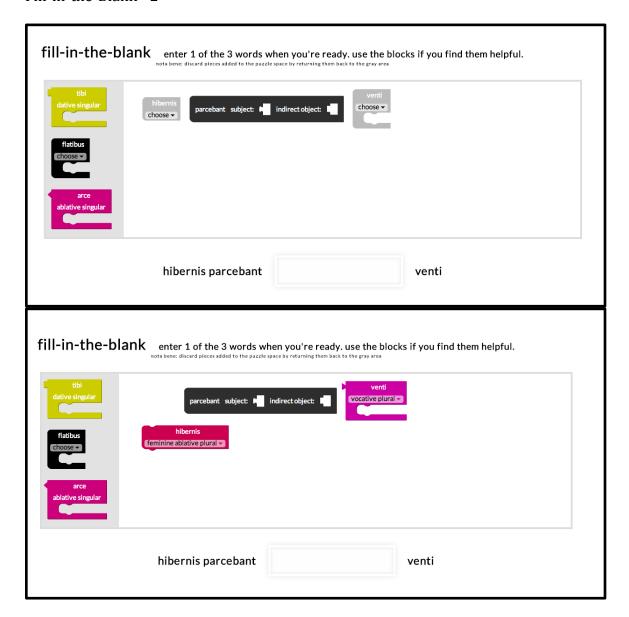


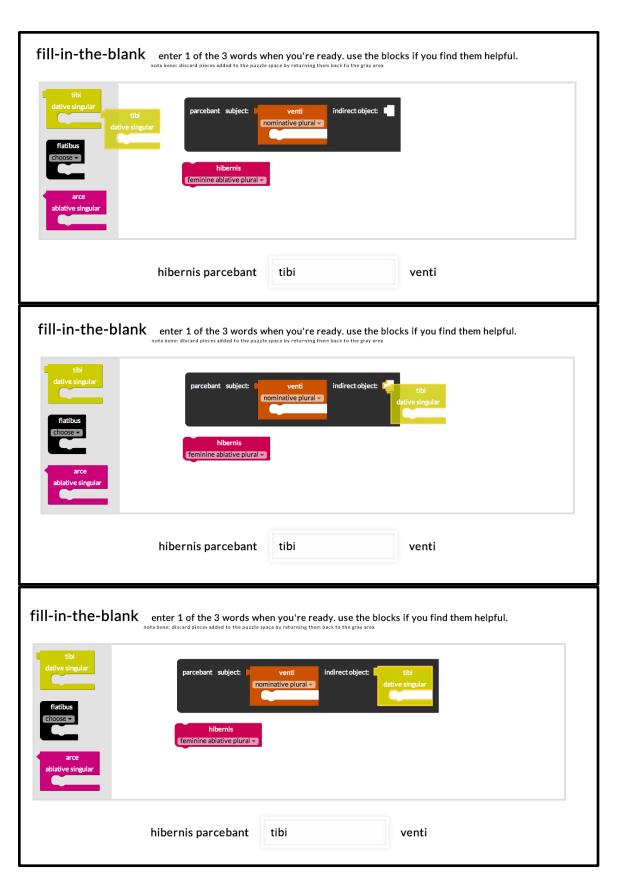


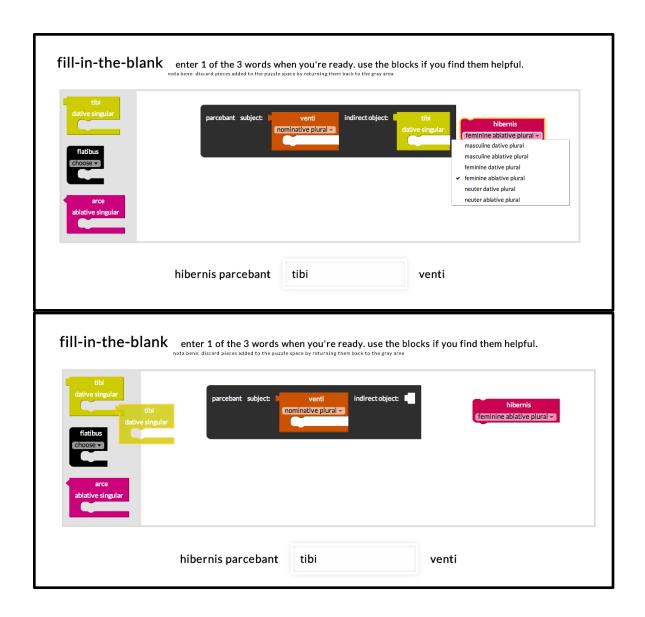


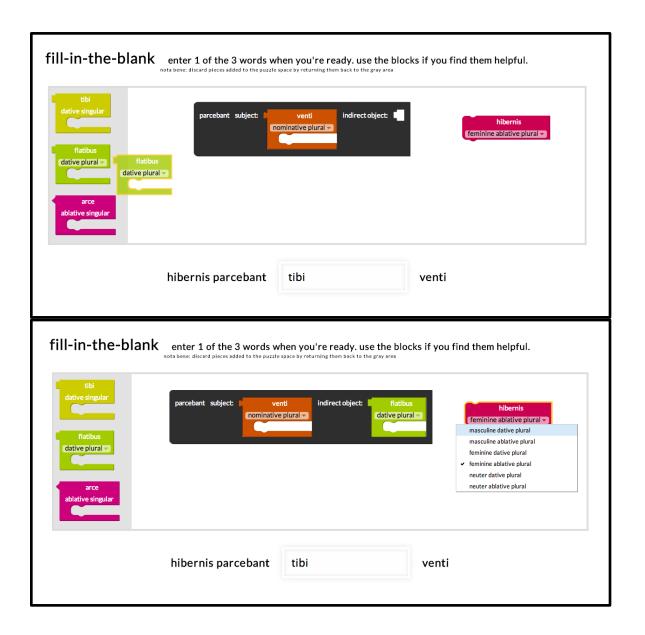


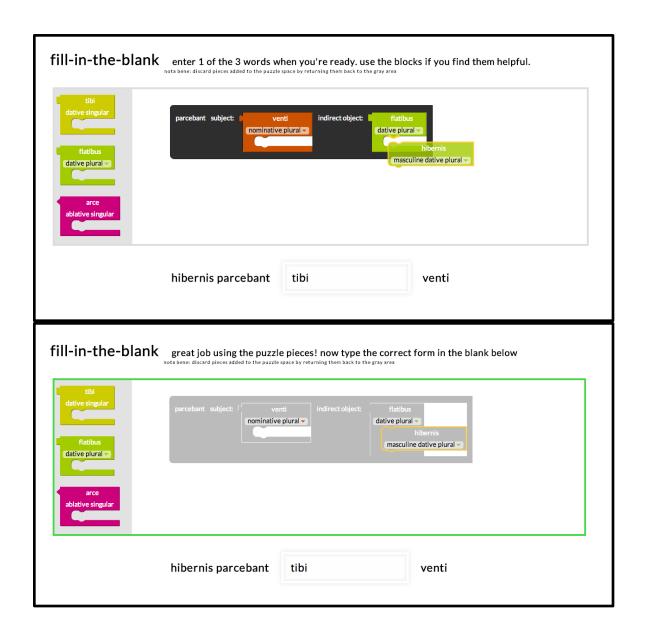
Fill-in-the-Blank - 2

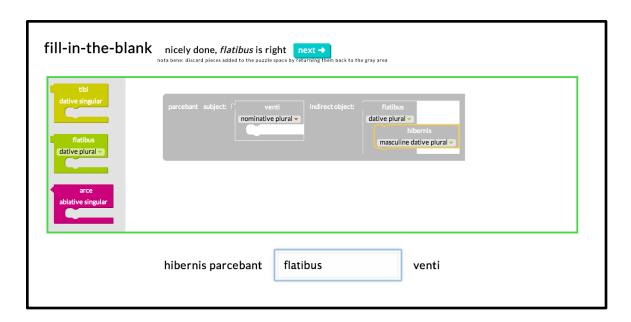




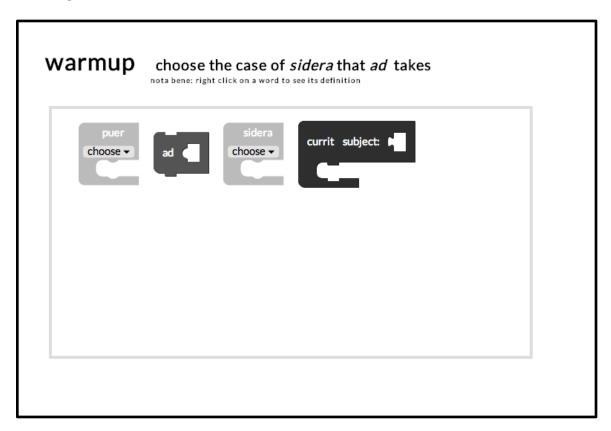


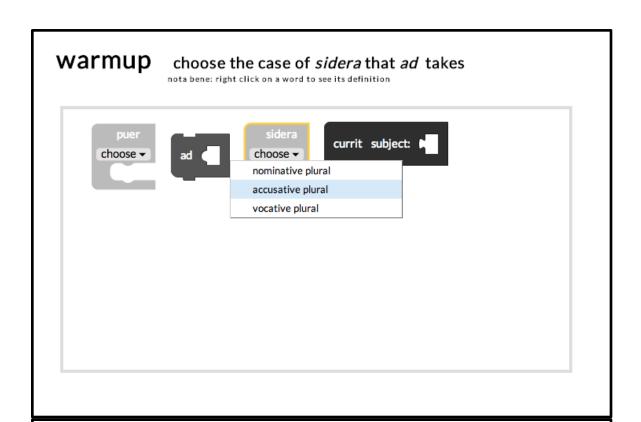


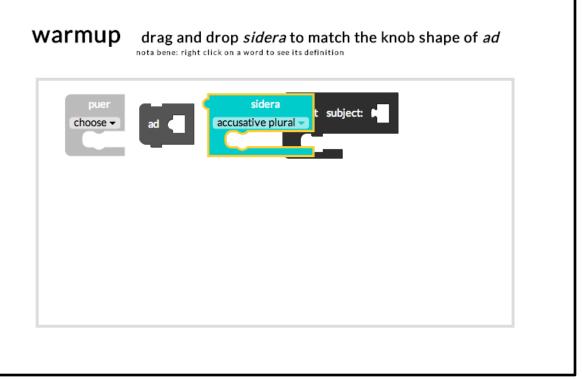




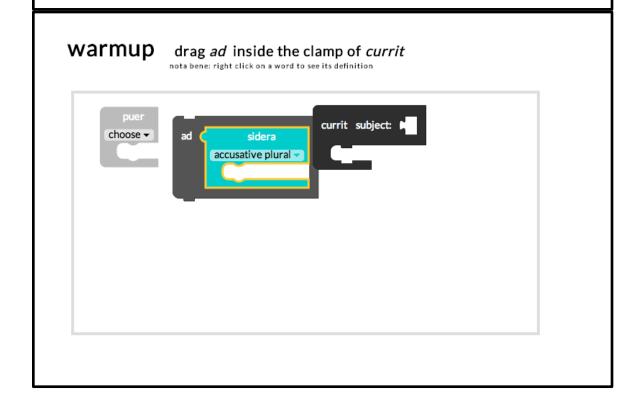
Warmup - 3

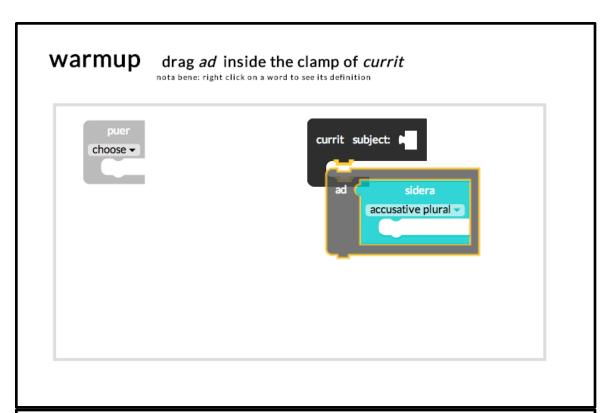


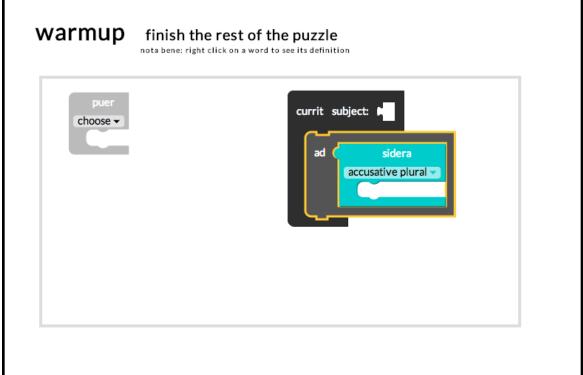


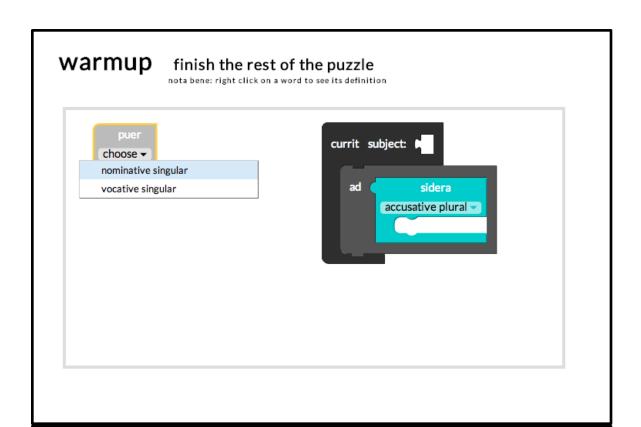


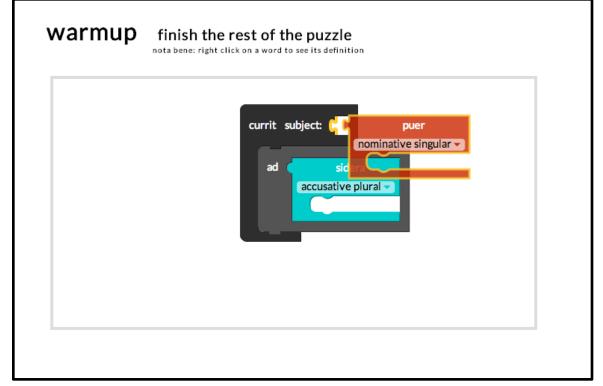
warmup drag and drop sidera to match the knob shape of ad nota bene: right click on a word to see its definition puer choose v sidera accusative plural v currit subject:

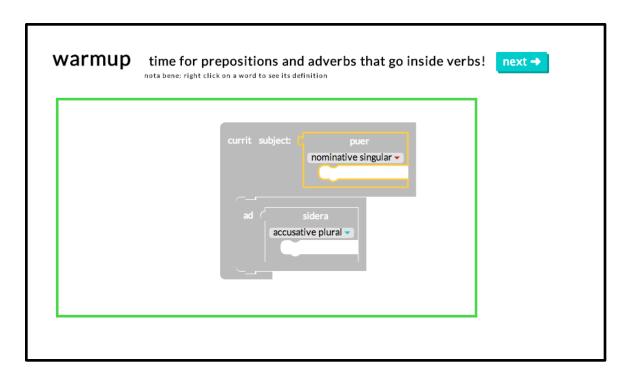


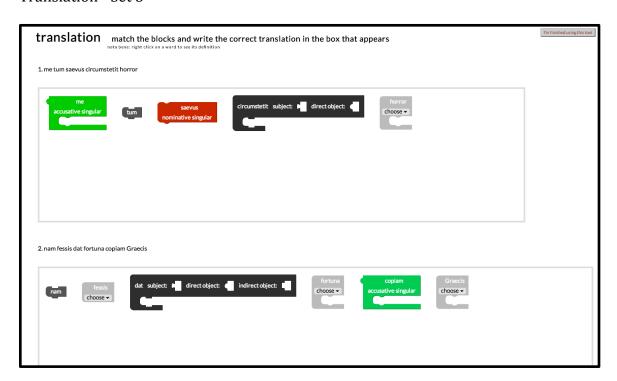


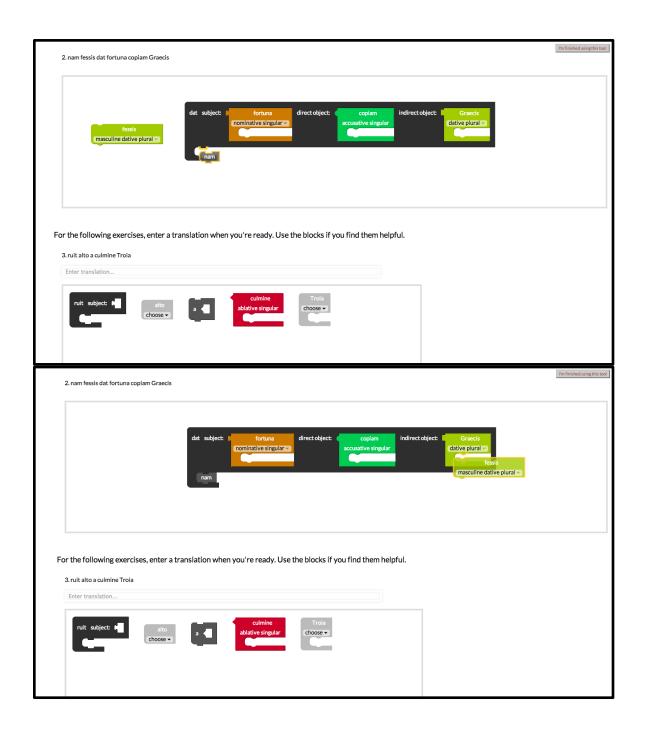


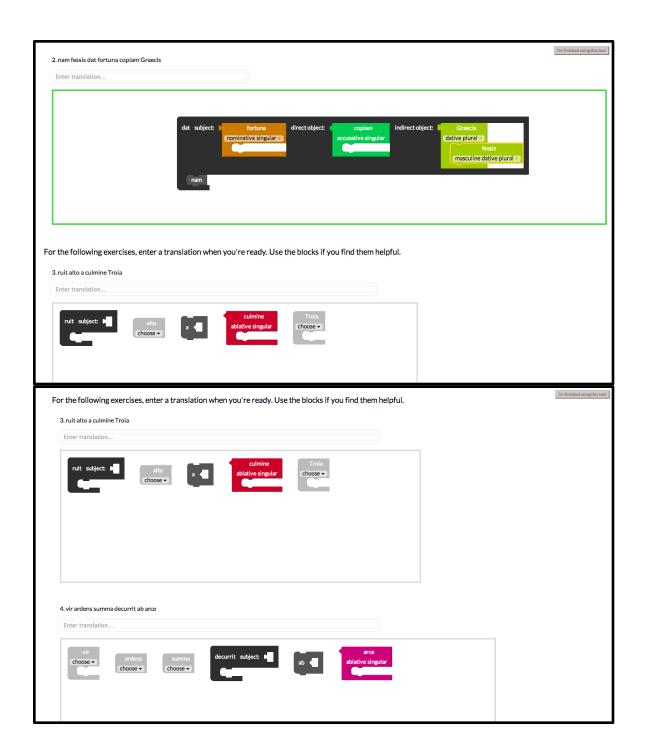


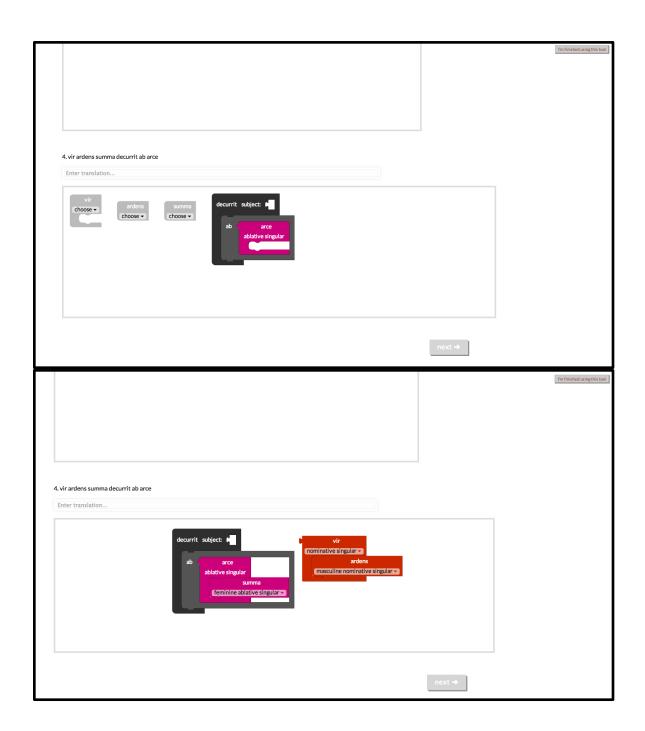






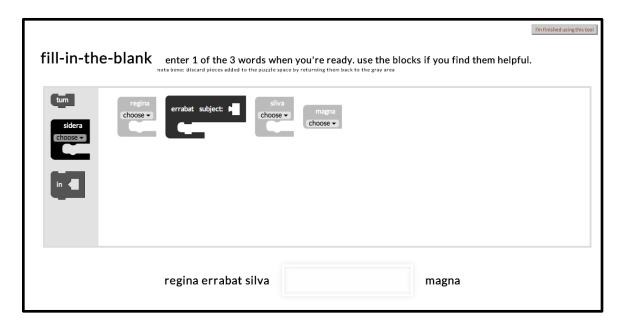


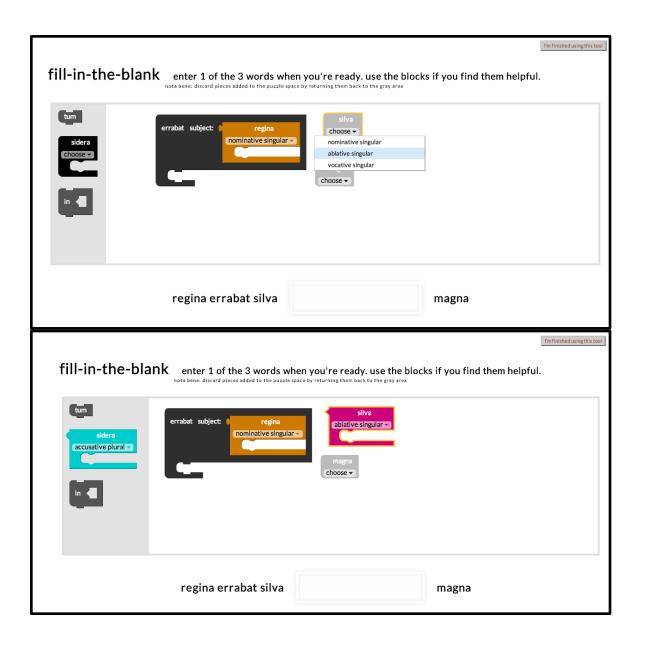


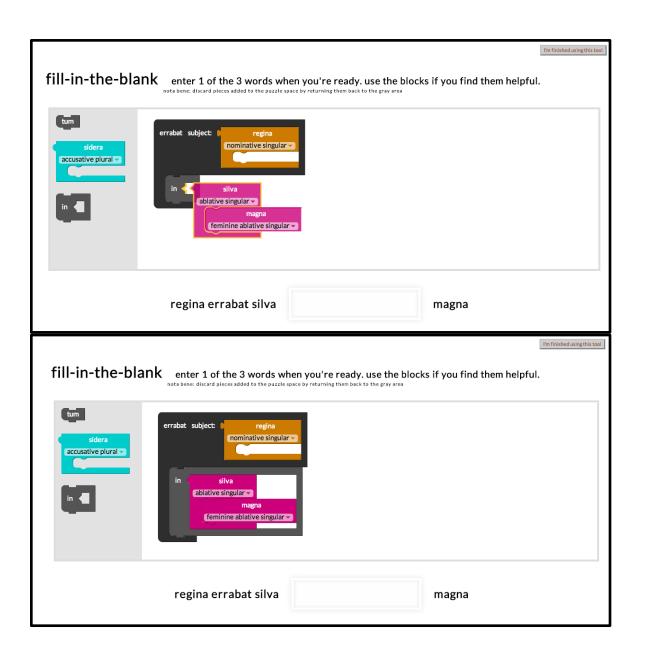




Fill-in-the-Blank - 3



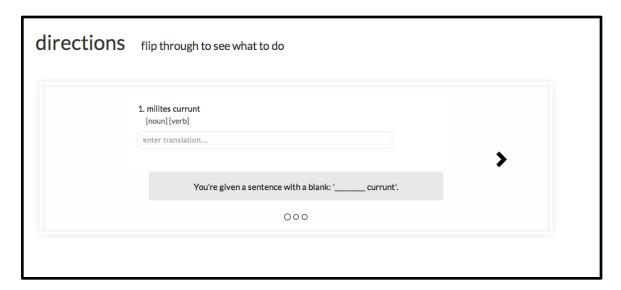


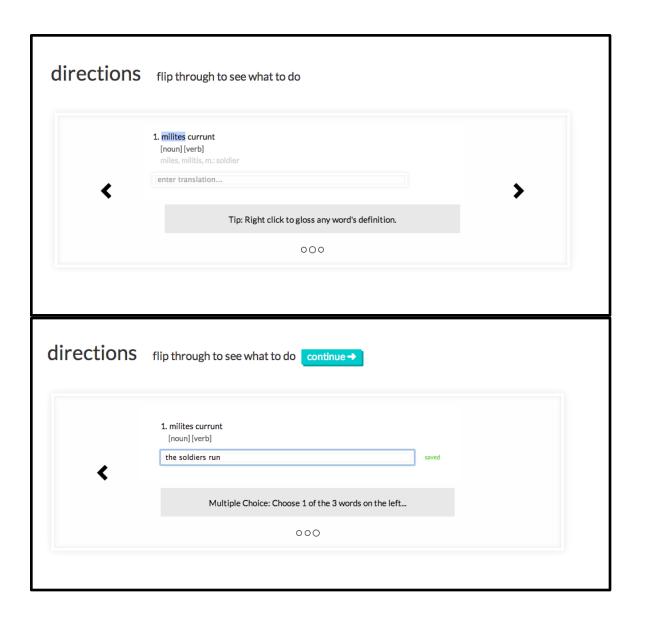


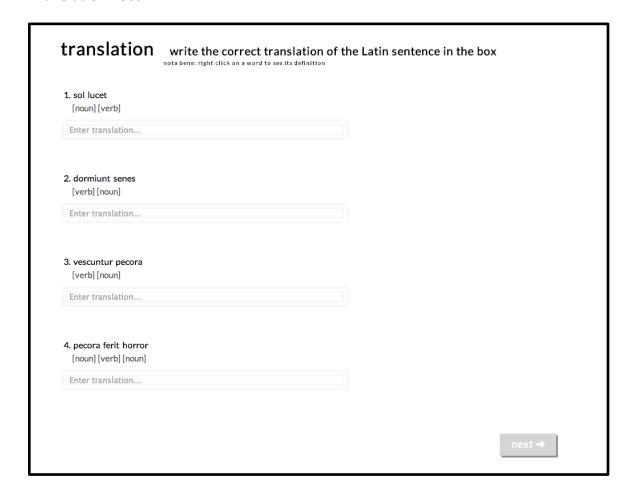


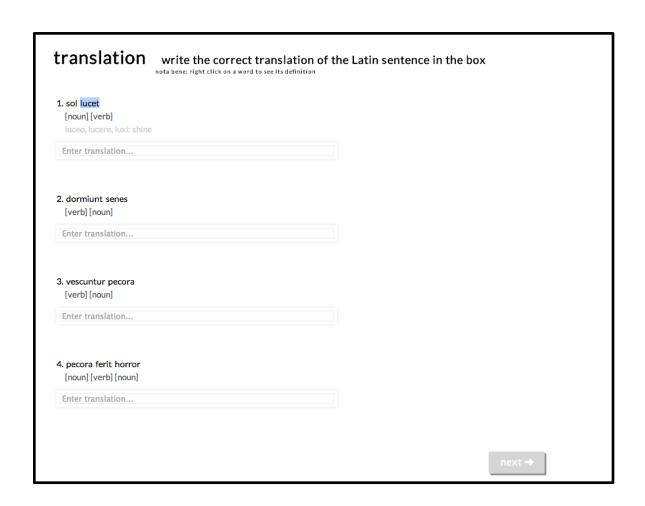
Control Condition

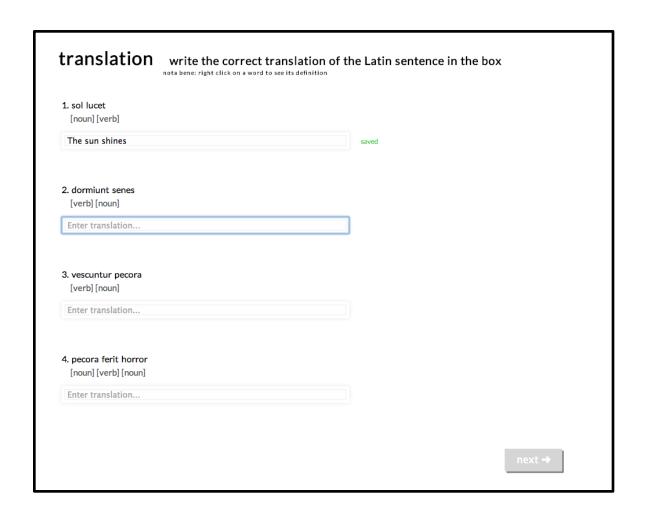
Translation - Directions

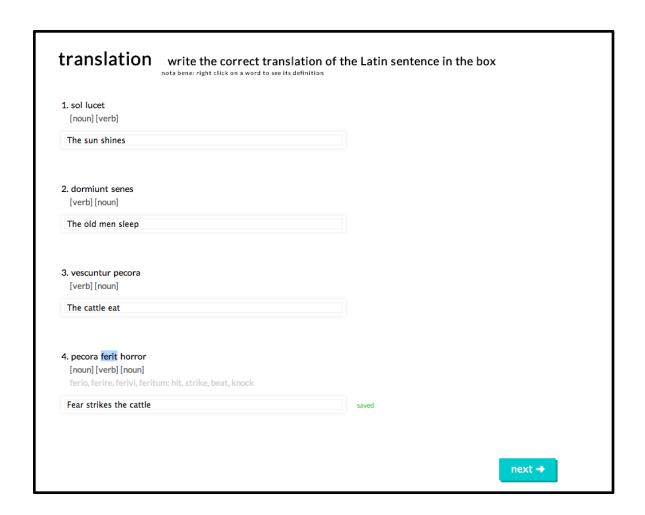






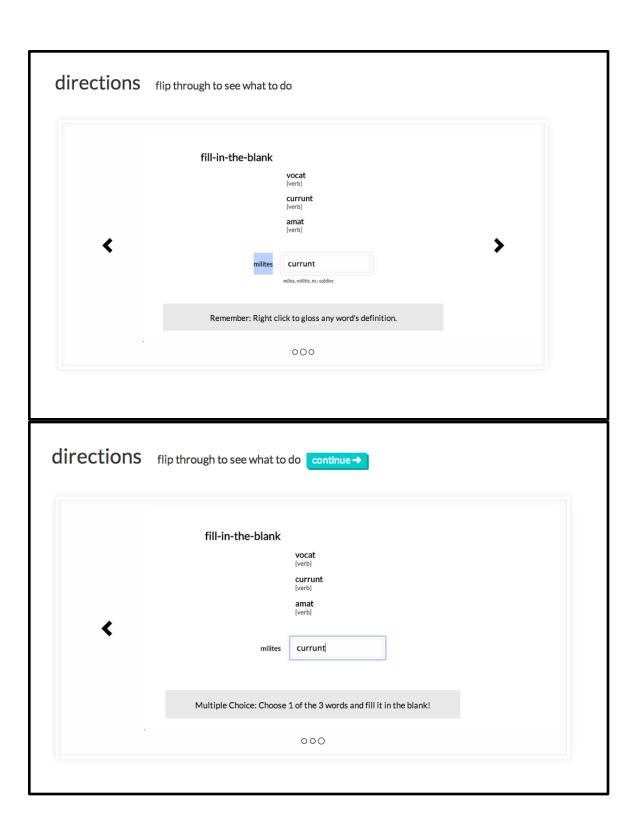






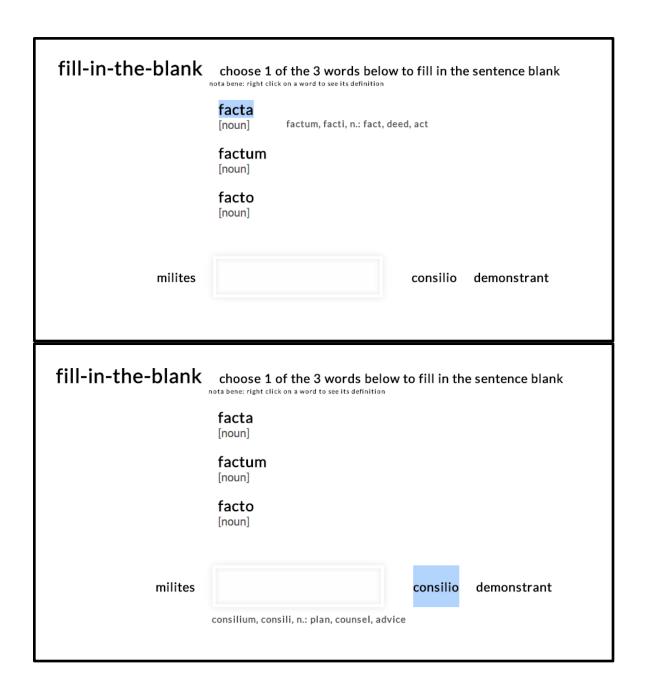
Fill-in-the-Blank - Directions





Fill-in-the-Blank - 1

fill-in-the-blank	choose 1 of the 3 words belo	ow to fill in the	e sentence blank
	facta [noun]		
	factum [noun]		
	facto [noun]		
milites		consilio	demonstrant
fill-in-the-blank	choose 1 of the 3 words belo	ow to fill in the	e sentence blank
fill-in-the-blank		ow to fill in the	e sentence blank
fill-in-the-blank	nota bene: right click on a word to see its definition facta	ow to fill in the	e sentence blank
fill-in-the-blank	nota bene: right click on a word to see its definition facta [noun] factum	ow to fill in the	e sentence blank
	facta [noun] factum [noun] facto		
fill-in-the-blank	facta [noun] factum [noun] facto	ow to fill in the	e sentence blank demonstrant

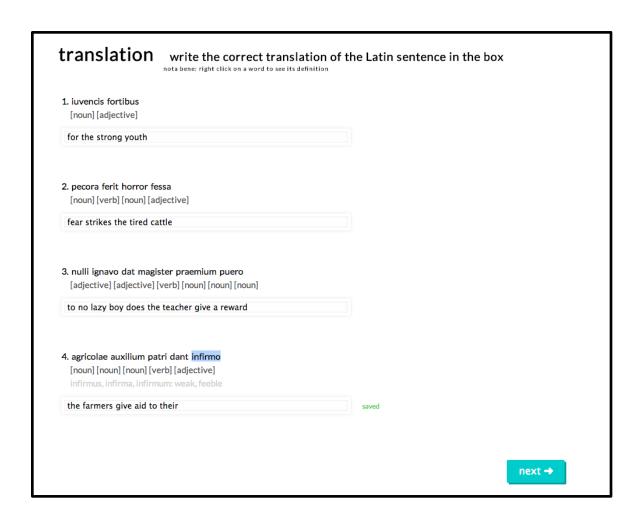


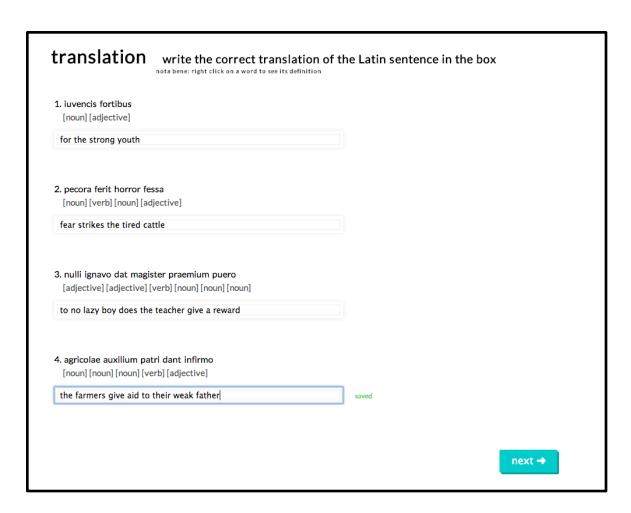
fill-in-the-blank	excellent choice, factum is co	errect next →
	facta [noun]	
	factum [noun]	
	facto [noun]	
milites	factum	consilio demonstrant

Translation - Set 2

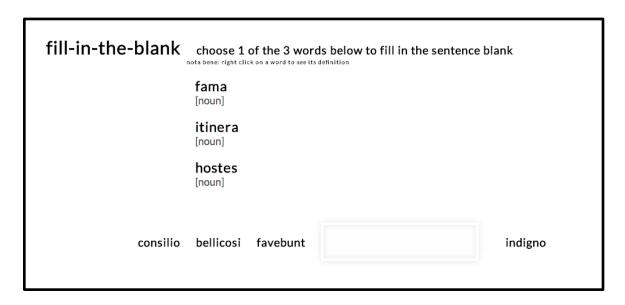


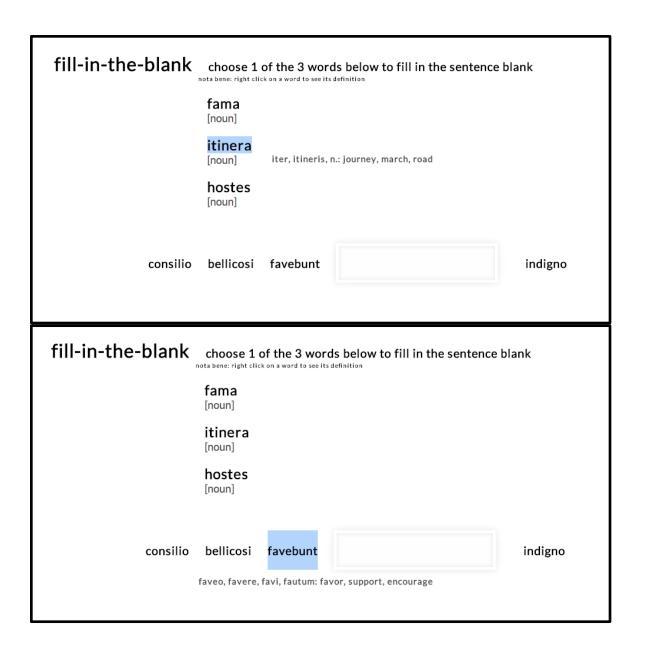
[noun] [adjective]			
iuvencus, iuvenci, m.: you	uth		
Enter translation			
2. pecora ferit horror fes	ssa		
[noun] [verb] [noun] [adj			
Enter translation			
nulli ignavo dat magist [adjective] [adjective] [v Enter translation	ter praemium puero verb] [noun] [noun]		
[adjective] [adjective] [v	-		
[adjective] [adjective] [v	verb] [noun] [noun] [noun]		
[adjective] [adjective] [v	verb] [noun] [noun] [noun]		
[adjective] [adjective] [v	verb] [noun] [noun] [noun]		

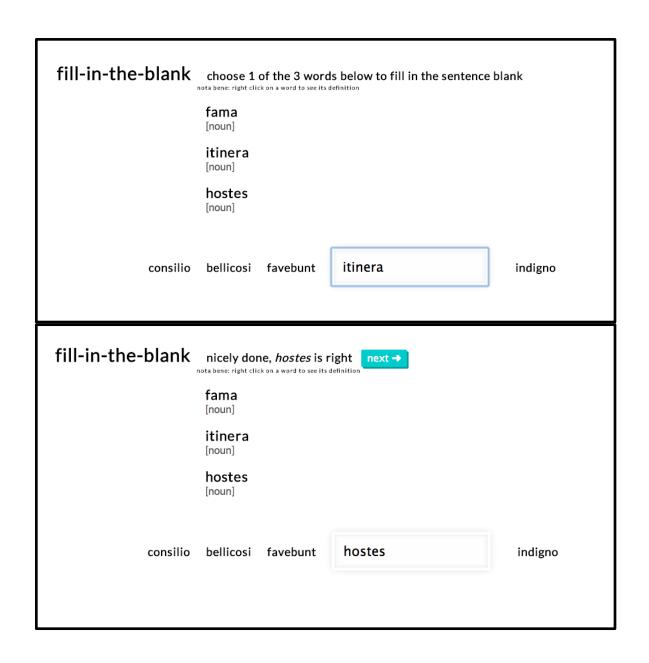




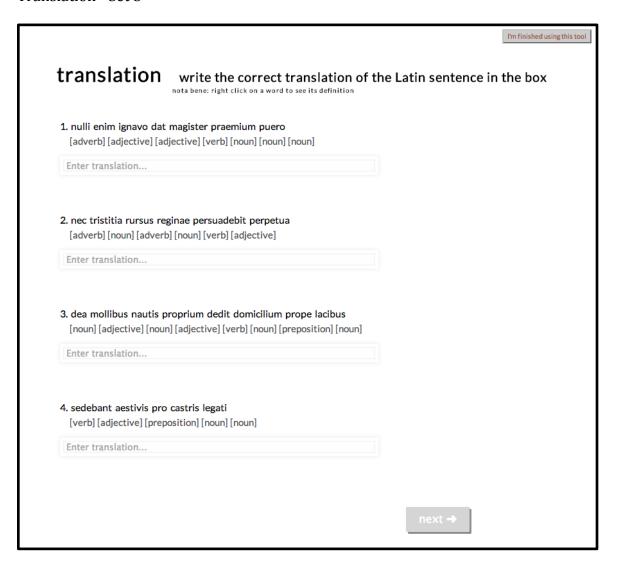
Fill-in-the-Blank - 2







Translation - Set 3



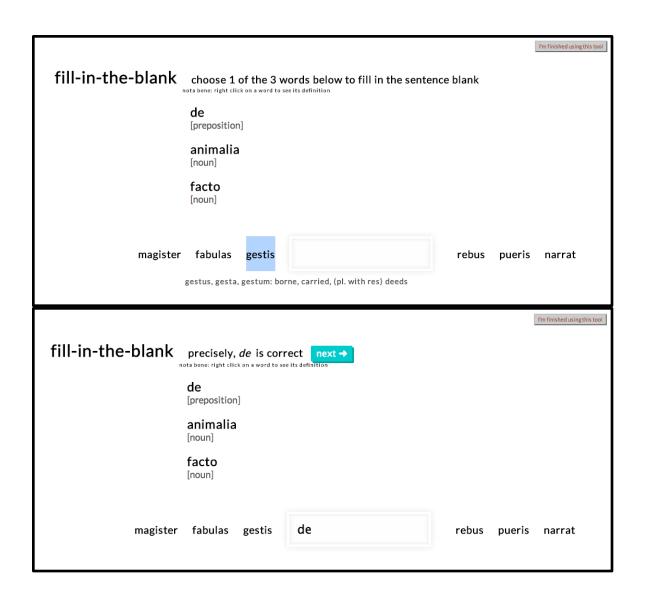
	I'm finished using
ranslation write the correct translation of nota bene: right click on a word to see its definition	the Latin sentence in the box
1. nulli enim ignavo dat magister praemium puero [adverb] [adjective] [adjective] [verb] [noun] [noun]	
for the teacher gives a reward to no lazy boy	
2. nec tristitia rursus reginae persuadebit perpetua	
[advorb] [noun] [advorb] [noun] [vorb] [adjoctive]	
[adverb] [noun] [adverb] [noun] [verb] [adjective] nor shall eternal sorrow persuade the queen	saved
	saved
nor shall eternal sorrow persuade the queen 3. dea mollibus nautis proprium dedit domicilium prope lacibus	saved
nor shall eternal sorrow persuade the queen	saved
nor shall eternal sorrow persuade the queen 3. dea mollibus nautis proprium dedit domicilium prope lacibus [noun] [adjective] [noun] [quipersition] [noun]	saved
nor shall eternal sorrow persuade the queen 3. dea mollibus nautis proprium dedit domicilium prope lacibus [noun] [adjective] [noun] [adjective] [verb] [noun] [preposition] [noun] domicilium, domicilii, n.: dwelling, home	saved
nor shall eternal sorrow persuade the queen 3. dea mollibus nautis proprium dedit domicilium prope lacibus [noun] [adjective] [noun] [adjective] [verb] [noun] [preposition] [noun] domicilium, domicilii, n.: dwelling, home	saved
nor shall eternal sorrow persuade the queen 3. dea mollibus nautis proprium dedit domicilium prope lacibus [noun] [adjective] [noun] [adjective] [noun] [preposition] [noun] domicilium, domicilii, n.: dwelling, home Enter translation	saved

	ion
nulli enim ignavo dat magister praemium puero [adverb] [adjective] [adjective] [verb] [noun] [noun]	
for the teacher gives a reward to no lazy boy	
nec tristitia rursus reginae persuadebit perpetua [adverb] [noun] [adverb] [noun] [verb] [adjective]	
nor shall eternal sorrow persuade the queen	
3. dea mollibus nautis proprium dedit domicilium prope lacibus [noun] [adjective] [noun] [preposition] [n	
the goddess gave her own home near the lakes to the mild sails	ors
4. sedebant aestivis pro castris legati	
[verb] [adjective] [preposition] [noun] [noun]	
	saved

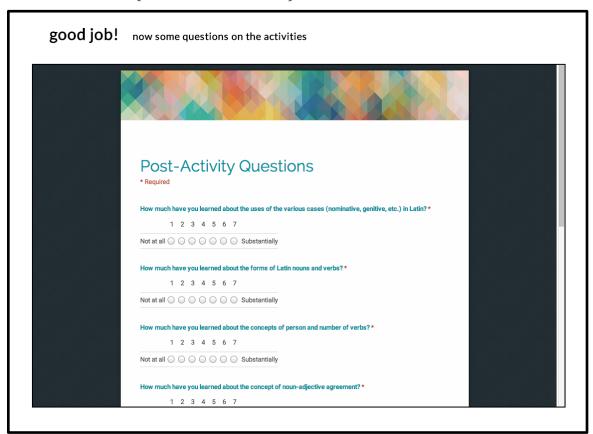
1. nulli enim ignavo dat n [adverb] [adjective] [adje	nagister praemium puero ective] [verb] [noun] [noun] [noun]	
for the teacher gives a re	ward to no lazy boy	
	nae persuadebit perpetua	
[adverb] [noun] [adverb]	[noun] [verb] [adjective]	
nor shall eternal sorrow	persuade the queen	
•	oprium dedit domicilium prope lacibus [adjective] [verb] [noun] [preposition] [noun]	
	the second of the second of the second	
the goddess gave her ow	n home near the lakes to the mild sailors	
the goddess gave her ow 4. sedebant aestivis pro o		
	astris legati	
4. sedebant aestivis pro c [verb] [adjective] [prepos	astris legati	saved

Fill-in-the-Blank - 3

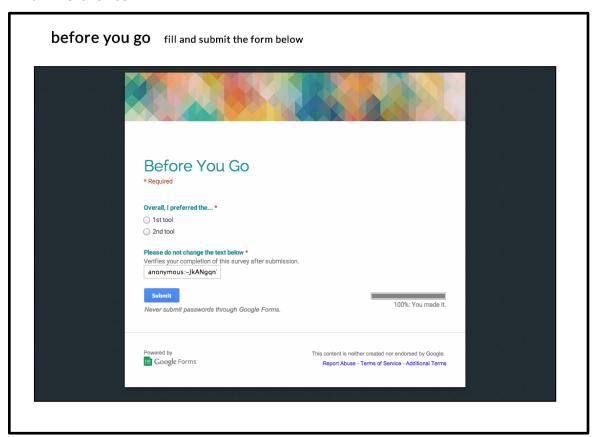




Post-Assessment (after each condition)



Final Preference



Thank you

thank you for participating!

We really appreciate it, and are excited to introduce new tools into Latin grammar study.

Bibliography

- Harrell, Robert. TCI FAQ Sheet. Creative Commons Attribution, 2013. PDF.
- Johnson, James F. "Alternative Approaches for the College Elementary Latin Sequence." *The Classical Journal* 82.3 (1987): 246–55. *JSTOR*. The Classical Association of the Middle West and South. Web. 24 Jan. 2015.
- Keller, Andrew, and Stephanie Russell. *Learn to Read Latin*. New Haven: Yale UP, 2006. Print.
- Kirschner, Paul A., John Sweller, and Richard E. Clark. "Why Minimal Guidance

 During Instruction Does Not Work: An Analysis of the Failure of Constructivist,

 Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching."

 Educational Psychologist 41.2 (2006): 75–86. Web. 22 Dec. 2014.
- Knudsvig, Glenn M., Gerda M. Seligson, and Ruth S. Craig. *Latin for Reading: A Beginner's Textbook with Exercises*. Ann Arbor: U of Michigan, 1986. Print.
- Krashen, Stephen. "The Case for Non-Targeted, Comprehensible Input." *Journal of Bilingual Education Research & Instruction* 5.1 (2013): 102–10. Web. 28 Feb. 2015.
- "Latin Word Study Tool." *Perseus Digital Library*. Ed. Gregory R. Crane. Tufts University, 2015. Web. 15 Jan. 2015.
- Lifelong Kindergarten Group. *Scratch*. Computer software. *Scratch Imagine, Program, Share*. Vers. 2.0. MIT Media Lab, n.d. Web. 2 Mar. 2015.
- Lister, Bob. "Latin in Transition." *Arts & Humanities in Higher Education* 8.2 (2009): 191–200. Sage Publications. Web. 24 Jan. 2015.

- Malan, David J., and Henry H. Leitner. "Scratch for Budding Computer Scientists." Proc. of SIGCSE'07, Covington, Kentucky. ACM, 2007. Web. 15 Dec. 2014.
- Markus, Donka D., and Deborah Pennell Ross. "Reading Proficiency In Latin Through Expectations And Visualization Free Ebook Download." *The Classical World* 98.1 (2004): 79–93. *JSTOR*. The Johns Hopkins University Press. Web. 20 Feb. 2015.
- Olson, Judith S., and Wendy Kellogg, eds. *Ways of Knowing in HCI*. New York: Springer Science+Business Media, 2014. Print.
- Phc Documentation. Program documentation. The Phc Developer's Manual. Tree
 Traversal API Tutorials, 20 Sept. 2011. Web. 1 Mar. 2015.
- Place, Perley Oakland, Curtis C. Bushnell, Harold L. Cleasby, and Alvah T. Otis.

 *Beginning Latin; an Introduction, by Way of English, to the Latin Language. New York: American Book, 1919. Print.
- Robert, Patrick. "Latin Is Not Different." *Septentrionale Americanum Latinitatis Vivae Institutum*. Proc. of SALVI Board Summit, Claymont Estate. SALVI, 21 July 2011. Web. 4 Jan. 2015.
- Rodrigo, Victoria, Stephen Krashen, and Barry Gribbons. "The Effectiveness of Two Comprehensible-Input Approaches to Foreign Language Instruction at the Intermediate Level." *System* 32.1 (2004): 53–60. Web.
- Ross, Deborah Pennell. "History of Michigan Latin." *Michigan Latin*. University of Michigan, n.d. Web. 1 Jan. 2015.

- Ross, Deborah Pennell. "Teaching Materials." *Michigan Latin*. University of Michigan, n.d. Web. 11 Feb. 2015.
- Sag, Ivan A., Ronald Kaplan, Lauri Karttunen, Martin Kay, Carl Pollard, Stuart

 Shieber, and Annie Zaenen. "Unification and Grammatical Theory."

 Proceedings of West Coast Conference on Formal Linguistics. 1986. 5:238–54.

 Web.
- Shelmerdine, Susan C.. Introduction to Latin. Indianapolis: Focus, 2013. Print.
- Sweller, John, Jeroen J. G. Van Merrienboer, and Fred G. W. C. Paas. "Cognitive Architecture and Instructional Design." *Educational Psychology Review* 10.3 (1998): 251–96. Plenum Publishing Corporation. Web.
- TJ, Howell. "Delivering Understandable Messages in Latin–A Report from the Field."

 Web log post. *Classical Association of New England*. Classical Association of

 New England, 11 Feb. 2014. Web. 08 Jan. 2014.
- University of Michigan. Reading with Expectations. N.p.: n.p., Feb. 2014. PDF.
- Wheelock, Frederic M., and Richard A. LaFleur. *Wheelock's Latin 7th Edition*. New York, NY: HarperCollins, 2011. Print.
- Wilkes, Angela, and John Shackell. *Latin for Beginners*. London: Usborne, 2001. Print.