Project Title:
INFRASTRUCTURE AND INTEGRATED TOOLS FOR PERSONALIZED LEARNING OF READING SKILL

Project Acronym:

iREAD

Grant Agreement number:
731724 — iRead H2020-ICT-2016-2017/H2020-ICT-2016-1

Subject:
D4.1 Dyslexia Domain Models (English and Greek)

Dissemination Level:
PUBLIC

Lead Beneficiary:
UOI

Project Coordinator:
UCL

Contributors:
UOI, UCL, NTUA

Revision	Preparation date	Period covered	Project start date	Project duration
V1	December 2017	Month 1-12	01/01/2017	48 Months

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 731724

Table of contents

List of figures

List of Tables

1 Executive Summary

This deliverable describes the rationale for, and the outline of the developed English and Greek domain models of reading skills for children with dyslexia. Domain models are is essential for the development of the literacy games, the iRead Reader, and content classification at later stages of the iRead project. A wide range of linguistic skills are assessed within domain models across languages, going beyond the skills of decoding of sound-letter correspondence and word recognition to more complex reading skills, including morphological, syntactic and discourse processing of text. Language category selection within domain models was based on the outcomes of extensive literature reviews, frequency counts derived from children's books and corpora available in the languages under investigation, as well as empirical data.

The overarching goal of the domain models' development is not only to incorporate different type of information (lexical, morphosyntactic etc.), but to also provide (a) ratings of difficulty, indicating each feature's difficulty or complexity relative to the rest of the features within the same category, as well as (b) progression schemes, in the form of prerequisites, indicating the order of teaching of those features. Specifically, rich texts in words/linguistic structures are likely to cause a barrier in reading. Therefore, all language features included in the models were classified as relatively easier or harder based on an initial taxonomy of features per category formulated by empirical data. This will be used for game and e-reader development to sequence the features.

The present report is structured as follows: Chapter 2 provides an introduction to domain models, including the following sections: Section 2.1 analyses the rationale and function of domain models in iRead; section 2.2 describes the role of domain models in iRead and the way they interact with other sub-systems; section 2.3 lists the types of information that domain models store; section 2.4 explains the terminology used in domain models, providing helpful definitions; and section 2.5 lists the various domain models included in the iRead system. Chapter 3 analyses the general methodological steps followed in building the domain models, while Chapters 4 and 5 describe the specific methodologies used in building the English and Greek domain models respectively. Finally, Chapter 6 provides an account of the supportive actions taken for the development and implementation of domain models.

2 INTRODUCTION: Domain Models

This work package focuses on the development of a comprehensive domain model of reading skills that will guide the development of domain models for the other languages involved in the iRead project from the vantage point of the students' needs. Although separate domain-models will be developed to tap on reading processes across the different child learner groups involved in the project (children with dyslexia, typically developing readers (novice readers), readers with English as a Foreign Language (EFL)) as well as across different languages considered (English, Greek, German and Spanish), this deliverable focuses on the development of the Greek and English Domain Models for children with dyslexia.

The discussed models were developed based on the revision and the significant extension of the previous Greek and English domain models developed for a-typical reading development (attributed to dyslexia) within iLearnRW FP7 EU project[1]. Note that the iLearnRW dyslexia domain models include only word level difficulties related to phonological and decoding skills in the two languages, whereas the new domain models are enhanced to incorporate more advanced reading skills that draw on a wider range of linguistic knowledge bases, going beyond the skills of decoding of sound-letter correspondences and word recognition to address more complex reading skills such as morphological, syntactic, and discourse processing of texts.

2.1 Rationale and function of Domain Models

Domain models constitute a central component in iRead, which will inform all sub-systems (Games, Reader, Text Classification) and determine the individualized pedagogical design used for each child. The role of each domain model is to (a) specify the linguistic structures that will constitute teaching material, (b) provide the basis for the formulation of individual user models, which will enable the iRead system to record each user's individual strengths and weaknesses, e.g. specific linguistic structures that he/she finds more challenging and needs to work on, and mark the user's progress as he/she uses the iRead system. In order to achieve these tasks, domain models contain all the necessary information that will render them able to provide specific information about each user's linguistic and learning particularities and individual needs. The linguistic information included in domain models was

[1] http://www.ilearnrw.eu/

especially selected in order to address all the linguistic sub-skills that are relevant or necessary for reading development, which are usually affected in dyslexia.

Based on the fact that we can read words with regular and irregular spelling as well as novel words and pseudo words with great ease, a dual-route model of reading was put forward (Coltheart et al. 2001), according to which reading can follow two routes for turning printed words into sound: a *direct* route (the *lexical orthographic* route), which is used when real words are read and involves accessing the lexical entry in the speaker's lexicon and activating semantic information during reading. However, reading new words or pseudo words makes use of an *indirect* (*sub-lexical phonological* route), which involves applying grapheme-phoneme conversion (GPC) rules (phonological recoding) without accessing the lexicon. According to this model, accessing the phonological information of a word during reading is not obligatory, while the lexical route is also divided into one part that accesses semantic information and one that does not. The Dual Route model can explain cases of readers who can comprehend printed words without being able to pronounce them, but also readers who achieve target-like loud reading without being able to retrieve any meaning from what they read.

Learning to read involves moving from a stage of alphabetic reading (applying grapheme-phoneme conversion (GPC) rules) to direct access reading. Various models have attempted to describe the process whereby a child shifts from alphabetic to direct access reading. Although there has been considerable debate on whether children start with large units (i.e. rimes, word beginnings or endings) and then move to small (i.e. individual phonemes) or the reverse, it is now generally agreed that they develop their reading skills by learning grapheme-phoneme correspondences first and moving to reading by analogy at a later stage (e.g. Brown & Deavers 1999; Ehri 1992).

Disruption in one of the two routes of the dual-route model can lead to different types of dyslexia. If the lexical route is impaired, the person has difficulty reading irregularly spelled words or exceptions, thus producing over-regularization errors such as reading 'island' as /ˈaɪslənd/ and 'steak' as /stik/ (Harley 2001). Conversely, if the sub-lexical route is impaired, the reader has difficulty applying GPC rules, thus faces great difficulties reading novel words or nonwords. However, in many cases the person can pronounce some nonwords, while in others there are also difficulties with certain types of existing words. In these cases, the individual has difficulty reading function words (e.g. articles), low-frequency words and morphologically complex words, while derivational errors (e.g. reading 'performing' as 'performance') and visual errors (e.g. reading 'perform' as 'perfume') are also very common (Harley 2001). Dyslexia has often been claimed to arise from problems with phonological processing, since people with this type of reading impairment often exhibit more general phonological problems (Harm & Seidenberg 2001).

Based on the above, domain models partly focus on those skills that constitute *phonemic awareness*, which includes knowledge of the phonological patterns and irregularities of one's language, as well as the skills to handle phonological units in producing and perceiving language. Phonemic awareness, which is usually enhanced in education by a set of activities that address *phonics* (e.g. phoneme discrimination, phonemic analysis and synthesis etc.), also includes the ability to relate phonological units to visual symbols of written language, i.e. *graphemes*. This specific ability, which is called *grapho-phonemic awareness*, is crucial for the development of reading and is very frequently disrupted in dyslexia. Therefore, domain models address all the sub-skills that are included in phonemic and grapho-phonemic awareness, skills that are actively used in the process of decoding. However, since decoding only constitutes part of the reading process, domain models also attempt to address higher-level language skills that map onto reading comprehension, that is, morphological awareness and syntactic processing.

All the linguistic information stored in domain models is used as the basis for the creation of individual user models, which will in away assess each child's abilities and place him/her at a specific point in his/her language's repertoire, aiming to provide him/her with linguistic material that is most relevant to his/her own needs and mark his/her progress. Specifically, the domain models will inform the focus of the adaptivity for which a combination of methodologies will be used (e.g. knowledge elicitation with teachers, data from literacy games usage and e-reader) to develop a machine learning component that automatically selects content for activities within the literacy games. This information will not only be used by the games and activities to allow users to interact with the words at the level of phonology, orthography, and syntax, but will also also provide educators with toolkits in order to enhance teachers' decision-making and lesson planning.

2.2 Role of Domain Models in iRead

Domain models play a central role in iRead, as they interact almost with every component of the iRead system. Domain models are stored within the iRead system as directed acyclic graphs (refer to Section 6.1 for more details). The vertices of the graph correspond to the language features of the domain model, while edges indicate prerequisites. In each vertex of the graph all attributes defined by the linguist experts are stored. Domain models are used for creating ad updating user models, for practising language features using the literacy games or the iRead reader application, and for content classification. In the following we give a short description of the aforementioned usages.

User Models

User models are treated by the iRead system as "instantiations" of the domain model. Each user model is a clone-graph of the corresponding domain model where all vertices and edges are present. In order to support personalized learning in reading, every vertex is equipped with an additional attribute, called "competence", which is used for storing the learner's current mastering of a language feature. This attribute not only allows for monitoring the user's progress, but also plays an important role in the status of the user's model: as the user practises various language features, their competence is updated, and more language features become available to the user (see Section 6.2 for details on this mechanism).

Literacy Games / iRead Reader

Literacy games and the iRead Reader are the main tools of iRead for practising language features. When a user interacts with these tools, her success or failure on the proposed tasks is recorded and a decision is made on whether the competence of the language features related to the tasks should be updated. The execution flow for playing a literacy game is as follows: before the game starts, the application asks for the user's available language features. Based on the user's usage data, the application choses the next language feature (or a group of them) for the user to practice. Then, appropriate and relevant content is provided to the game. While the user plays her actions are recorded and stored as logs. Afterwards, the application decides whether the competence of the language feature should be updated. If this is the case, then the user's model is updated and, depending on the new model, more language features might become available.

Content Classification

Content classification refers to the task of classifying texts based on their "difficulty". Here difficulty is related to the language features that appear within the texts. The same text though might be easier for a reader who has mastered most of the language features appearing in the text. Personalization of the content classification is achieved by taking into consideration the user's model; in particular her competence on the language features.

For supporting content classification, literacy games and the iRead Reader, it is essential to develop resources such as child appropriate dictionaries; these resources are bound to the domain model, meaning that for a dictionary-entry or a sentence of a text, one should be able to track all language features that are relevant.

2.3 Information stored in Domain Models

In order to address complex reading skills, the described domain models incorporate different *linguistic levels* including phonology, morphology and syntax. Each linguistic level is represented by a number of language phenomena or structures, called *language categories*, each of which includes a set of specific instances, the *features*. For instance, at the phonological level domain models include *categories* such as syllabification, grapheme-phoneme correspondence (GPC) and phonemes, and a number of *features* for each category such as open, closed syllables, consonant clusters and diphthongs for syllable patterns, or consonants and vowels in the initial and internal position for the category of phonemes. At the morphological level domain models include the categories of derivational and inflectional suffixes, as well as prefixes used for derivational processes. Specific suffixes (e.g. the English past tense –*ed*) constitute the linguistic features of these categories. Within the syntactic level we included categories like function words (some features of this category would be clitics and articles), prepositions, negative particles, embedded constructions, passives, and discourse anaphors. Orthography is also part of the domain models with an emphasis on letter similarity.

For each of the selected linguistic features included in domain models, the following types of information are stored: (a) characteristic examples, which include typical instances of each feature, (b) a rating of difficulty, indicating each feature's difficulty or complexity relative to the rest of the features within the same category, and (c) prerequisites, which are those features that must be acquired *before* a feature can be addressed in the iRead application. Difficulty is instantiated by number values that denote a scaling of the features that belong to a single category with respect to their relative difficulty or linguistic complexity. The exact specification of the relative difficulty of features differs slightly between the English and the Greek domain model: an interval 4-level scale was used to rate features of the English domain model, specifying difficulty levels between 1 (easiest) and 4 (hardest). Using a slightly different method, the features of the Greek domain model were ranked with respect to difficulty and assigned a number value specifying their position in the within-category difficulty scale. For instance, in a category with 10 features, these features would be ranked (ordered) from the easiest to the hardest and each given a number value that denotes this order. In this category, difficulty levels would range from 1 to 10 or lower, as each level of the scale may correspond to more than one feature (i.e. several features may be placed on the same level of difficulty). Difficulty levels are then used to produce the learning *prerequisites* for each feature, which define the linguistic material that needs to have been mastered before addressing a particular feature. Specifically, the prerequisites of a feature are selected out of the features within the same category that are of lower difficulty level and are most closely related to the target feature. Crucially, prerequisites cannot be of a higher difficulty level than the feature itself.

Additionally, frequency counts for each feature were also used in building the domain models. Frequency counts were drawn from (a) children's texts and (b) general corpora within each language. Frequency is used in order to ensure that domain models include features that are frequently encountered by the target population or constitute part of their learning curriculum, rather than features that children are unlikely to deal with. To achieve this aim, frequency counts are used to remove features that are very infrequent or absent from children's text.

Language categories were selected based on three criteria: (a) those that are most relevant to reading in general as a language skill and linguistic process, (b) those that are most frequent in language use and often appear in children's' books, and (c) those that are found to be more difficult to develop and are, therefore, often emphasized in teaching. Each of these is described in detail for English and Greek domain models in Chapters 4 and 5 respectively.

In some models[2], features within each category that are closely related are clustered into groups to denote which features can be addressed or taught in parallel in both the literacy games and the reader. When grouping features we took into account their linguistic characteristics (e.g. phonemes /p/ and /t/ would be grouped together as they share phonetic characteristics and articulation properties), as well as the level of difficulty or complexity of each feature, so that no features with substantial differences with respect to complexity would be grouped together. In other words, the members of a group can be features of equal or similar frequency and difficulty, which involve the same or similar linguistic process and require the same specific skill to master. Additionally, for the purposes of this project, prerequisites were included to indicate progression within language categories. Progression here relates to the order of teaching, the order of acquisition and the prerequisites for the establishment of new knowledge.

2.4 Terminology used across Domain Models

The linguistic information stored in Domain Models is organised in terms of *Linguistic Levels*, *Language Categories* and *Language Features*. These are defined as follows:

- **Linguistic level** refers to the level of analysis each language structure belongs to. For example, phonemes and syllables refer to the *phonological* level of analysis, morphemes and word constituents like prefixes and suffixes refer to the *morphological* level, while phrase and sentence structure refers to the level of *syntax*. Additionally, word suffixes that are determined by syntactic relations, such as suffixes that reflect subject-verb agreement, refer to the level of *morphosyntax,*

[2] Grouping of features was used in the Greek domain model, as well as in the models that are being built for German and Spanish, but not in the English domain model. This difference is due to the highly inflectional nature of the three languages, which render grouping similar items together necessary.

the interface between morphology and syntax. Finally, elements that reflect the visual processing of graphemes or words belong to the level of *orthography*.

- **Language category** refers to categories of linguistic structures, like *phonemes* and *syllabification*, *prefixes* or *suffixes*, *embedding* or *passive voice*. Each language category belongs to one of the linguistic levels mentioned in the previous paragraph.
- **Language feature** refers to the specific instances of each category included in domain models. For instance, the category *phonemes* of the Greek domain model includes the Phonemes /p/, /b/, /d/, /g/, /s/ and so on in word-initial and word-internal position, while the category of Syllabification of the Greek domain model includes different syllable structure combinations like CV-CV, CVC-CV, VC-CV, etc. (see Appendix II). As another example, the category of Grapheme-Phone correspondence (GPC) of the English domain model includes features like "c pronounced as /s/" and "s pronounced as /s/", while the Greek domain model includes consonant clusters such as "μπρ pronounced as /br/" and "γκρ pronounced as /gr/" in initial/internal position for the GPC category. Finally, the category of Embedding of the English domain model includes specific types of embedded clauses like relative clauses, *that*-clauses etc. (see Appendix I).

The following table provides an illustration of how the information stored in the domain models is structured (examples are given in English).

Table 1 Structure of information in domain models

Linguistic Level	Language Category	Example features
Phonology	Phonemes[3]	/p/, /t/, /e/ etc.
	Syllabification	CV, CVC, VC etc.
	Grapheme-Phoneme correspondence	English: p-/p/, pp-/p/, s-/z/-, s-/s/ etc. Greek: γκρ, μπρ, στρ.
Word recognition		As, science, etc.
Morphology	Suffixes	-ed, -ing etc.
	Prefixes	pre-, un-, dis- etc.
Morphosyntax	Agreement	3rd person –s, etc.
	Adjectives: comparative	-er than, -ier than, etc.
Syntax	Embedding	*That* clauses, relative clauses, etc.

[3] Included in the Greek domain model only.

	Passive	Be + PP + (by phrase)

2.5 Domain Models included in iRead

Within the iRead project, several domain models have been developed for reading development, while there are still two that will be delivered in the following months. The ones that are discussed in this deliverable include models of (a) English and Greek children with dyslexia, while domain models for (b) English and Greek novice readers and (c) readers learning English as a foreign language (EFL) are described in Deliverable 4.2. Two more domain models will be delivered in Month 16 (Deliverable 4.3, "Novice Reader Domain Models: Spanish/German"), including the domain models for German and Spanish typically developing novice readers. As already mentioned, a number of sources were used to specify the specific sub-skills to be included in the developed domain models. Extensive literature reviews, empirical data and various databases were consulted when selecting the lexical, morphological and syntactic constructions for inclusion in iRead to ensure that frequent and developmentally appropriate lexical and morphosyntactic constructions are targeted. The models also provide information about how the various subskills are interlinked, the level of difficulty involved in developing the particular skills, and the progression of the development of each skill; i.e., the requirements for the establishment of new knowledge.

Domain models include similar, but not identical, linguistic phenomena. Greek for instance (similar to German and Spanish), unlike English, is a highly inflected language, where several inflectional morphemes are used to express a complex system of grammatical categories, such as agreement (first, second, and third person, singular and plural number), tense (past, non-past), aspect (perfective, imperfective), voice (active, non-active), and mood (imperative, non-imperative) (Holton, Mackridge, & Philippaki-Warburton, 1997; Philippaki-Warburton, 1973). Therefore, some domain models are more analytic than others, but all of them consist of the same linguistic levels and store the same properties for their language categories and features. Taking beginning readers of English as an illustrative example, we specified the morphosyntactic phenomena included based on an extensive literature review on typical and atypical reading acquisition. We also collected frequency information derived from the British National corpus[4], the national curriculum in England[5] (for primary public schools), the Wisdom's e-book corpus (an iRead project partner), and the children's printed word

[4] http://www.natcorp.ox.ac.uk/
[5] https://www.gov.uk/government/collections/national-curriculum

database[6] comprising books read by children aged 5-9. For second language learners, we additionally took into account frequency information obtained from learners' corpora. In the case of Greek, we selected the morphosyntactic phenomena included based on (a) an extensive literature review on typical and atypical reading acquisition, (b) empirical evidence derived from online experiments that we designed and delivered (see section 5.1.2.2.) to tap into reading processes in children aged 6-12, and (c) frequency counts derived from the online ILSP written Greek Corpus[7] (Hatzigeorgiou et al., 2000) and in a corpus of school books that are available online[8].

Thus, although the models are organized in a similar overlapping way, significant differences are also apparent given that resources for each language differ. Additionally, the notion of progression between categories differs across domain models. Typically developing children who acquire their reading skills naturally move from phonology to morphology and syntax. However, progression schemes (as defined by prerequisites) differ for EFL learners and children with dyslexia because EFL readers and children with dyslexia have more limited linguistic knowledge than first language readers when they start reading (Grabe, 2009).

[6] https://www1.essex.ac.uk/psychology/cpwd/
[7] http://hnc.ilsp.gr/
[8] http://ebooks.edu.gr/new/allcourses.php

3 General methodological approach to building Domain Models in iRead

3.1 Steps followed

Driven by the knowledge we obtained from the previous iLearnRW project, we built our domain models for both English and Greek language by retaining and revising the existing phonological and morphological classification. The new models developed included new language areas that were not targeted in the previous models, such as syntax as well as its interface with morphology, namely morphosyntax.

Two basic steps were followed in building domain models: (a) selecting language categories and features, and (b) defining the relative difficulty of features within each category based on the relevant literature. Different methodologies were used to ensure that we have included the most relevant phenomena per category. To collect empirical data for typical and atypical reading development and to get an indication on which constructions are harder than others, we (a) investigated the relevant literature, (b) collected empirical data, and (c) used text corpora available across languages where this was possible. These methods are discussed in greater detail in the Sections 3.1.1 and 3.1.2 and are instantiated in each domain model in different ways.

3.1.1 Selecting language categories and features

In order to reliably select the appropriate linguistic phenomena for the new models, we first used evidence from published studies for the populations we are interested in the literature. Specifically, for both the English and Greek domain models for dyslexia, we conducted extensive literature reviews not only to define the relevant language areas for reading development, but also to define which language areas are harder to develop while learning to read or while acquiring language in general (or while acquiring a second language, which can be used for the EFL models).

Literature reviews included studies that investigate typical reading development (which is relevant to both the novice and EFL readers' domain models) for two reasons. Firstly, because it is necessary to define the underlying processes and language areas that are involved in learning to read, and secondly because researchers seem to agree that linguistic areas that are harder during language acquisition also tend to cause difficulty to learners of a second/foreign language as well as speakers with compromised linguistic skills (White, 1985;

Schwartz, 1998; Bley-Vroman, 1989; 2009). Additionally, literature reviews included studies that investigate atypical reading development in children with dyslexia or poor readers in general (relevant to the dyslexia and EFL readers' domain models). This kind of evidence can be used to inform the classification of language features with respect to difficulty when no relevant information on these particular features can be found in the L1 reading or EFL/ESL literature. After the initial compilation of lists of language categories and features, all entries were checked for frequency in relevant corpora of children's texts (depending on availability for each language).

3.1.2 Defining difficulty

Within each language category, features were classified according to their *difficulty/complexity*, which directly relates to both WP5 (content classification), WP6 and WP7 (development of games and e-reader). For instance, each domain model within iRead project is designed to contain several features of the target language (phonological, morphological, syntactic) and the level of difficulty each of these features may cause to students' receptive skills. As a consequence, a text that is rich in words/structures that are sensitive to the "difficulties" of a specific user (captured in the user's model) is likely to cause a barrier in reading. Therefore, less complex or easier constructions are selected/taught first within each language category, while more complex ones follow, a path that is defined by setting learning *prerequisites* for each feature (see below). It is important to note that by definition the notion of "complexity" necessarily predicts difficulties not only in poor readers/readers with dyslexia or EFL readers, but also in novice readers who also have to deal with more complex constructions.

To define which linguistic features are easier or harder, previous findings were drawn from the relevant literature in order to formulate an initial taxonomy of linguistic features per category. This initial difficulty rating will then be empirically tested either through experimentation (see 5.1.2.2), or by collecting empirical data from relevant populations (see 4.1.1.2 and 4.1.2.2). By conducting empirical work, we were able to record and analyze actual data from children with dyslexia and reliably define aspects of language that are more difficult across young readers.

The crucial role of the difficulty levels stored in domain models is to ensure students' progress in reading skills, by specifying the typical progress in learning to read. Specifically, the model was built based on the typical process of reading acquisition across categories, as well as the difficulty levels of features within categories, and denotes prerequisites for the establishment of new knowledge. The prerequisites of a particular feature are linguistic material (features), either within the same category or from other categories that need to have been acquired or mastered before the feature is addressed in the iRead learning apps, i.e. the Games or e-Reader. Difficulty feeds the formation of prerequisites in that all within-category prerequisites of a feature should be of lower or equal difficulty (not higher) to the feature itself. This information will be used for game and e-reader development to sequence features in each student's learning journey. For example, a child learning to read in English cannot proceed to the feature "*ss pronounced as /s/*", if he/she has not mastered the feature "*s pronounced as /s/*" in the category of Grapheme-Phoneme Correspondence (level of Phonology).

It is important to note though that children who struggle with reading due to dyslexia or children who are learning a foreign language do not progress at equal pace as typically beginning readers. In these cases, the adaptivity component of the system will interact with the child's user model to determine their learning journeys through the domain model.

4 English Domain Model for children with dyslexia

4.1 Methodology

This section describes the development of the English Domain Model for children with dyslexia. The model includes phenomena across the linguistic levels we discussed earlier, namely phonology, morphology, morphosyntax, syntax and orthography. Given the very different nature of the levels relating to language sounds and letters (phonology and orthography) on one hand, and levels relating to inner language structure (morphology, syntax, morphosyntax) on the other in terms of linguistic material, the construction of the two parts of the model is described separately in 4.1.1 and 4.1.2 respectively.

4.1.1 Phonology, orthography [confusing letters], word recognition

4.1.1.1 Selecting language categories and features

The iRead domain models for dyslexia build upon the FP7 project iLearnRW, which included the levels for phonology, and orthography. These initial categories developed were refined in iRead looking at the literature on typical and atypical reading development. This led to the identification of a new category for word recognition, which was not addressed in iLearnRW and is critical for learning to read a non-transparent language such as English.

Once the initial categories were chosen, we undertook an evaluation of features that had not been carried out in iLearnRW, namely considering feature frequency in children's books and reading materials to determine feature inclusion. To appraise this, we used the children's printed word database (CPWD) (Masterson et al., 2011). The CPWD is a database of words derived from children's books and reading schemes designed for the first four years of reading instruction. These materials were identified in a survey of representative Primary schools throughout the UK (Masterson et al, 2010). Books from eight publishers were included to construct the CPWD word database totalling 12,193 words. The database has been tagged for sound-letter correspondences and is searchable online at https://www1.essex.ac.uk/psychology/cpwd/.

In addition to this, we triangulated the original iLearnRW domain model features with those included in literacy programmes such as Letters and Sounds[9], Jolly Phonics[10], Progression

[9] https://www.gov.uk/government/publications/letters-and-sounds

in Phonics[11], THRASS charts[12], ReadWrite Inc[13] (top down) and also by searching for words in the CPWD to inductively determine feature position (bottom up). This allowed us to determine if a feature's position affected difficulty and to capture this in the domain model. Given the iterative nature of our domain model design, these methodological decisions will be discussed in Section 4.2 when justifying domain model categories and features.

4.1.1.2 Defining difficulty

The research literature was reviewed to determine whether there is an evidence-base for the order that the features would typically be taught (i.e., feature difficulty). This literature review and conversations with other academics in the field highlighted the lack of existing research in this area. Therefore, literacy and and phonic programmes (Jolly Phonics, Progression in Phonics, Letters and Sounds, ReadWrite Inc) as well as the CPWD were referred to when defining feature difficulty. Across all categories, feature difficulty was defined on a scale of 1 to 4, with Level 1 denoting very easy, and Level 4 very difficult. Based on this, all single letters (phonemes) were grouped under Level 1, as these were considered to be the entry level to phonics training.

For example, within the linguistic level phonology, in a similar progression to the literacy programme Letters and Sounds, Level 2 includes common vowel digraphs and double consonants at the end of a word (e.g., /ss/ as in mess). Level 3 includes the less common vowel/consonant combinations and double consonants in the middle of a word (e.g., /ss/ as in assistant). Again, with reference to existing literacy programmes, Level 4 contains consonant clusters (e.g., /bl/, /cl/, etc), which should be taught once the initial stages have been mastered.

Given the lack of an empirical work to guide us in constructing difficulty levels, it is necessary to revise the difficulty levels. This will be achieved in the second and third years of the project once the domain model is implemented in the games. Obtaining children's game performance on different features will be analysed to empirically validate and update feature difficulty.

---

[10] http://jollylearning.co.uk/overview-about-jolly-phonics/
[11] http://dera.ioe.ac.uk/4419/
[12] http://www.thrass.co.uk
[13] https://global.oup.com/education/content/primary/series/rwi/?region=uk

4.1.2 Morphology, syntax, morphosyntax

This section provides a description of the methodology used for the selection of the language categories and features on the levels of Morphology, Syntax and Morphosyntax.

4.1.2.1 Selecting language categories and features

The initial selection of the language categories and features that were included in the English Domain Model was based on previous research findings on the difficulties encountered by children with dyslexia. Specifically, the literature search focused on the acquisition of morphosyntactic phenomena as well as the processing of these categories by English children with dyslexia. When the dyslexia-relevant literature did not provide specific information on the acquisition or processing of morphology, syntax and morphosyntax, findings relevant to children with developmental language disorders were used, as well as information on typical language development.

After compiling an initial list of linguistic structures and specific features, a thorough corpus search was performed to obtain frequency counts for features of the English domain model (the same search was used for the English domain model for Novice Readers and EFL, see Deliverable 4.2). The search was based on a corpus of children's literature containing 1,518,339 tokens (running words) in 552 texts. The corpus was grammatically annotated (part-of-speech tagged) and searched for the target linguistic features. Grammatical annotation was added using Tree-Tagger (Schmid 1994)[14] and a multi-dimensional analysis tool MAT (Nini 2014). The corpus was searched using #LancsBox (Brezina et al. 2015), a corpus analysis tool developed at Lancaster university. The measurements that were produced through the corpus search were then normalised to frequency rates that correspond to the number of appearances per 10,000 words. The frequency rates were used to eliminate very infrequent features from the domain model. Specifically, a cut-off point of 10 appearances per 10,000 words was used and 5 language features were removed.

4.1.2.2 Defining difficulty

In order to determine the relative difficulty of language features included in the English domain model, an initial literature search was performed. This search led to a

[14] More info: http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/Penn-Treebank-Tagset.pdf

preliminary scaling of features with respect to difficulty (presented in Appendix I), which will be revised through the collection of empirical data. Similar to the approach taken for phonology and orthography (section 4.1.1.2), data will be collected through administering the iRead games to a large sample of children with dyslexia and recording their performance in the learning tasks involved in the games. Following data collection, a Rasch analysis[15] will be performed on the data to determine the exact relative difficulty of the linguistic features included, which will then revise the difficulty scaling that was formed based on the literature. This process is planned to take place in spring 2018.

4.2 Language categories included in English Domain Model for children with dyslexia

As described above, we started with the domain model developed in iLearnRW. This domain model was limited to the following language levels, categories and features:

- Phonology Level (GPC, Syllables, Blends and Letter Clusters Categories; 262 features)
- Visual processing Level (Confusing Letters Category; 15 features)

These two levels had a total of 273 features. We initially scrutinised this previous work, by asking: do the language levels and categories follow terminology consistent with the reading and linguistics literature? Is the categorisation and description of a feature useful in how it is intended to be used by the technology? This led to a first revision in the ways in which levels and categories were re-labelled. For example, the level 'visual processing' was relabelled to 'orthography/confusing letters', since orthography constitutes the level of language that relates spoken language to its written form and includes all skills associated with handling this relation, such as graphophonemic awareness (matching graphemes to phonemes) and visual processing and recognition of words. Once this refinement was complete, we revisited the rationale of features and their inclusion under each language category. In what follows we motivate the importance of each category and then explain our refinements. A summary of the linguistic levels and language categories selected is given in Table 2 and a detailed analysis and description follows.

[15] https://www.rasch.org/rmt/rmt213d.htm

Table 2 Summary of English domain model

Linguistic Level	Language Category	N of Features
Phonology	Grapheme-Phoneme Correspondence	160
	Consonant clusters	39
	Syllabification	5
Word recognition		11
Orthography	Confusing letters	2
Morphology	Prefixes	9
	Derivational suffixes	10
Morphosyntax	Agreement	5
	Inflectional suffixes: Tense (regular)	3
	Inflectional suffixes: Number (regular)	2
	Inflectional suffixes: Aspect	5
	Adjectives: Comparative/Superlative (regular)	6
Morphosyntax (lexical formation)	Irregular past forms	1
	Irregular past participles	1
	Irregular plural	1
	Irregular formation of comparative/superlative	1
Syntax	Adjectives (modifiers)	3
	Discourse anaphors	8
	Binding	3
	Quantifiers	7
	Passive	2
	Coordination	2
	Embedding	11
	Wh- questions	6
	Yes/No questions	1
	Prepositional phrases	3
	Modal verbs	3

Phonology: GPC and Consonant Clusters Categories

The Simple View of Reading (SVoR; Hoover & Gough, 1990) identifies word recognition and comprehension abilities as two important skills for successful reading. To tap into the development of children's word recognition ability, the domain model has an explicit focus on supporting decoding. Models of reading are useful to identify processes involved in skilled reading and this in turn can help us to understand why children may struggle to learn to read. Models of early reading argue that learners must acquire the alphabetic principle, that is to say that they must develop letter-sound knowledge (Ehri, 2005). A popular research-informed model of visual word recognition, the Dual Route Cascaded model (DRC; Coltheart, Rastle, Perry, Langden & Ziegler, 2001), pinpoints two approaches that work in parallel to read text - the lexical versus the non-lexical route. The lexical route requires the learner to access their bank of stored sight words, drawing on semantic information (this route is assessed below in the 'word recognition' section); while the non-lexical route is the process in which learners analyse the phonological features of the word. The non-lexical route suits the decoding of words or letter strings that follow regular spelling-sound rules, as readers will take a step-by-step approach to identify (segment) the grapheme-phoneme correspondence (GPC) and then to blend the sounds together to read the whole word. Thus, phonological awareness (i.e., the ability to identify sounds within words) and phonics knowledge (single letter, vowel digraphs, consonant clusters) contribute to the non-lexical route of reading and are important skills for decoding unfamiliar words. Further, Share's (1995) self-teaching hypothesis argues that once skilled readers have gained letter knowledge and decoding skills, they are well-equipped to become more independent readers and to use prior knowledge and inference to support the development of fluency when reading.

Reinforcing the need to teach phonological skills, research on children with dyslexia - a population with known reading difficulties - has consistently found characteristic deficits in phonological awareness. Specifically, the Rose (2009) definition identifies that children with dyslexia present with significant problems with reading accuracy and fluency, as well as poor phonological awareness, verbal memory and processing speed. There is strong support for the phonological deficit hypothesis as a causal explanation of the reading difficulties by individuals with dyslexia (Vellutino et al., 2004). Moreover, intervention research has consistently shown that the teaching of phonics and phonological awareness supports the development of reading accuracy and fluency for children with and without literacy difficulties (Bowyer-Crane et al., 2008; Rose, 2009).

The domain model follows a synthetic and analytic phonics rationale that recognises the smaller sound representation in words and also the bigger letter units in teaching children how to decode words. There are some similarities between the two approaches. When using analytic phonics children tend to begin with learning some high frequency sight words and learn letter sounds in the beginning, middle and end positions in words. This is followed by children using the sounds for blending. With synthetic phonics, children are taught phonics in a certain order and learn a few new phonemes at a time and synthesise the sounds from very early on to build words (Johnson & Watson, 2014). Both approaches to teaching phonics are incorporated within the domain model, allowing a more varied and motivating instruction in the games.

Within the original iLearnRW domain model there were GPC features characterised by a number of letter groups including both consonants and vowels. In probing the words in which the features appeared, it became evident that these features were parts of words. Examples included to**ngue** and rese**arch**. These were excluded because they were considered to be too fine-grained and children would be able to segment these larger chunks further using their existing GPC knowledge (i.e., phoneme knowledge - separating 'n' from 'gue' in tongue). This decision was validated in two ways: first, we reviewed existing literacy schemes and second, we held two workshops with ten practitioner teachers from UCL's SpLD Masters programme. This confirmed that our approach was suitable and that learners should be taught the most frequent GPC features to avoid over complicating the learning process.

After this initial filtering, we applied additional inclusion criteria for the remaining GPC features based on how frequently they appeared the children's printed word database (CPWD). GPC features that appeared in less than five words within the CPWD, but were yet present in frequent words, were moved to the Word Recognition language level (see below) allowing children to learn these irregular constructions in whole words. Example words included **gn**ome (/n/-gn), de**bt** (/t/-bt), and **wh**ole (/h/-wh). In total, 56 irregular GPC rules were infrequent and were thus moved to the Word Recognition level (forming a new feature, feature 10).

In a final step, we reviewed the position of GPC features and Consonant Cluster features in a word as originally specified in the iLearnRW domain model. We did this by consulting literacy programmes such as Letters and Sounds, Jolly Phonics, Progression in Phonics, THRASS charts, ReadWrite Inc (top down) and also by searching for words in the CPWD to inductively determine feature position (bottom up). This process identified a few instances where a GPC feature (e.g. /t/-tt) required definition both at the end and the middle of the word, the former being easier than the latter, and thus important to represent in the domain model.

This stepwise evaluation process reduced and refined our original feature set from iLearnRW domain model to a final set of **166 GPC and consonant clusters features** represented with both phoneme (in the form of IPA) and grapheme. For example, features /s/-s and /s/-ss share the same sound (IPA) but different graphemes.

Word Recognition

The original domain model from iLearnRW did not have a word recognition language level. Yet, the inclusion of word recognition items was deemed important particularly for English language learners due to the opaque/deep orthography. English has a number of words that cannot be decoded phonetically; these are often labelled as ‘irregular’ or ‘exception’ words (e.g., ‘yacht’; Ziegler et al, 2010). This means that they must be read as whole words and, in the context of the dual-route cascade model (Coltheart et al., 2001; see above), learners would be using the direct lexical route to meaning and pronunciation. In addition, it is considered important that children become familiar with high frequency words early on in the school system (Ehri, 2005). Developing a sight vocabulary alongside developing phonological recoding processes are both essential to becoming a skilled reader. Good oral language provides the foundation that enables children to become successful readers.

The lexical route is considered to be a faster way of reading, as it relies on sight word knowledge rather than step-by-step phonological decoding. The development of a bank of sight words read from memory supports the development of accurate and fluent reading. Sight word reading occurs automatically and is of key importance when learning to read in English.

This language level was designed to comprise of 10 features, each of which hosts a word list. The first ten features cover the 400 most common words, which are typically taught at schools as sight words. We constructed this word list by triangulating sight word lists from common reading-based interventions/materials used in UK (e.g. Jolly phonics, Letters and Sounds, Dolch word reading lists, Fry word lists, and Progression in Phonics) with word frequency metrics retrieved from the entire word database of CPWD. After collecting the 400 words we organised them in features (each feature contained a list of words) based on the order they tend to be taught, and their word length. For example, words such as ‘and’, ‘go’, ‘had’ appeared in feature 1 whereas words such as ‘world’, ‘anything’, ‘chapter’ appear in later features 6 and 7. The 10th feature includes the infrequent GPCs moved from Phonology/GPC.

Using the CPWD, we derived words for each of these features and then in turn chose those that did not have more complex morphological features. The reason behind this was that children must learn the ‘root’ word first of all and morphology is being covered separately in the domain model, which will be at a more advanced stage. A manageable number of words were chosen using the advanced search tool from the CPWD containing the letters identified and the phoneme in question. e.g the word ‘castle’ contains the letters ‘st’ represented by the phoneme ‘s’ as in ‘some’. Other words chosen to form part of this group were ‘bustle, gristle, fasten, listen, and moisten’ and demonstrate that ‘st’ represented by /s/ only seems to occur within the middle of a word. Words were chosen that had a higher frequency per million and matched the IPA symbol in question. Another consideration was where the grapheme appears in the word and some examples were chosen to reflect the variety of positioning within words.

Syllabification

A syllable is a unit of pronunciation. From an early age, prior to reading instruction, children are taught to identify syllables from oral presentation to develop their phonological awareness skills (Hatcher, Duff & Hulme, 2014). Syllabification can be an important strategy for learners because it enables children to work with larger units than phonemes when reading, which may in turn increase fluency although research is still needed in this area. The ability to identify syllables within words is often emphasised in reading intervention programmes (e.g., Toe by Toe, Sound Linkage), with the idea being that this awareness of word structure is a useful strategy to help break down multisyllabic words into more manageable chunks (Diliberto, Beattie, Flowers & Algozzine, 2008), albeit still larger than at the phoneme level.

All syllables must include one vowel, which may be represented by one or more allograph(s). Thus, when reading independently, vowel sound knowledge is crucial to successful syllabification (Lane, Pullen, Eiselle & Jordan, 2002). It follows, therefore, that children must have secure phonic knowledge to be able to read/chunk syllables.

For this category, the original iLearnRW domain model included five features for syllable numbers (1-syllable to 5-syllables) and eight syllable rules indicating where the child should split a word depending on its structure. Rules in particular were used in some of the iLearnRW games where a child was required to split a word with the game dictating where the correct split was. In English there are numerous rules for syllabification, which are complicated to remember, are often ambiguous, and have a limited evidence-base (Diliberto et al., 2008; Kahn, 2015). For example, children may be taught that when there is one consonant between two vowels in a word, they must divide the syllables after the first vowel (e.g., 'bo/nus'). However, if following the previous rule does not make a recognizable word, the child must divide the syllables after the consonant that comes between the vowels (e.g., 'doz/en').

One consistent rule however (mentioned above), is that a single syllable should include a vowel. To this end, it could be argued that the explicit teaching of syllabification rules is less useful than an approach that allows more flexibility for learners to choose the syllable boundary (e.g., splitting as 'ra/bbit' versus 'rab/bit'). In research by Bhattacharya and Ehri (2004), it was shown that syllabification supports the decoding process of struggling readers but only when a flexible approach to syllable boundary is used. The notion to reduce the number of rules that need to be learned also aligns with key teaching principles for specialist groups (e.g., reducing demands on memory and avoiding learning confusion, Carroll et al., 2011). Given this literature, the syllabification features were revised to focus on children's ability to identify the number of syllables only.

Orthography (Confusing letters)

Children that experience difficulty with reading have been reported to confuse visually-similar letters, which has been proposed to stem from problems with visual and phonological decoding (Lachmann & Geyer, 2003). There is limited evidence in this area and the existing research has focused on letter formation (i.e., handwriting) rather than confusing letters while reading. There were 14 pairs of confusing letters in the iLearnRW domain model. However, given that the majority of combinations would likely be confusing when writing, and not necessarily during word recognition, the decision was taken to focus on the 4 letters most commonly confused. This decision was informed by research by Brendler and Lachmann (2001) who identified /b/d/ and /p/q/ as the most commonly reversed letters in the English language due to their shape and sound.

(Morpho)Syntax

Based on extensive literature reviews and frequency data derived from text corpora, we selected the (morpho)syntactic structures to include. Within morphosyntax domain, we tried to cover a broad range of syntactic phenomena to ensure that (a) the level of constructions will range from easy to more difficult, and (b) materials simulate phenomena used in texts. We also tried to be consistent as for the categories included within models to ensure that domain models do not differ significantly across languages. All language categories included across linguistic levels are presented in Appendix I.

In texts, the average sentence often combines at least three clauses (Scott and Balthazar, 2015). These multiclausal sentences are harder to read and process and require from the reader not only to discern the main proposition from the others, but also to recall the logical meaning and relationships among syntactic elements. Thus, *linear* constructions with coordination, where clauses are connected with coordinating conjunctions (e.g., *and, but, so*) are typically easier compared to more *hierarchical* constructions with subordination (e.g., *because, when, if*) (Creaghead and Donnelly, 1982; Macaruso, Bar-Shalom, Crain, and Shankweiler, 1989). Taking this into account, within syntax, we included constructions with coordination (see e.g. 1 below), as well as several types of subordination (embedded constructions), such as adverbial (e.g. 2), relative (e.g. 3) and complement clauses (e.g. 4), with different features. For instance, within adverbial clauses we included features that denote sequence (e.g., *He arrived home when the rain stopped*), causation (e.g., *He died because/due a terrible illness*) or conditionals (e.g., *If you are careful, you can get at the top*), since these sentence types are very frequently used (Fang, 2006; Fang and Cao, 2015).

(1) *The man hit the boy **and** the girl kissed her mother.*
(2) *The man hit the boy **as soon as** he kissed the woman.*
(3) *The man saw the doctor **that** hit the athlete yesterday.*
(4) *The man believes that the athlete kissed the woman.*

Cross-linguistic data on language acquisition has also shown that, up to the age of six, typically developing children have selective difficulties interpreting structures such as *wh*-questions and relative clauses. Children with dyslexia often experience difficulties with these constructions as well (Bar-Shalom, Crain, & Shankweiler, 1993; Byrne, 1981; Casalis, Leuwers, & Hilton, 2013; Leikin & Assayag- Bouskila, 2004; Mann, Shankweiler, & Smith, 1984; Smith, Macaruso, Shankweiler, & Crain, 1989; Stein, Cairns, & Zurif, 1984). Therefore, within relative clauses we included features that take into account word order so we included both subject (e.g. 5a and 6a) and object extracted constructions (e.g. 5b and 6b), given that structures with canonical order (such as subject extracted constructions) have proved to be easier for children to acquire (Correa, 1995; Friedmann and Novogrodsky, 2004; de Villiers, 1979; Friedmann, Belletti and Rizzi, 2009; Guasti, Branchini, Arosio and Vernice, 2012). Additionally, the variable of subject-object reference of the pronoun within the relative clause was combined with alternating positions of the relative clause with respect to the main clause, so we included center embedded (CE) relative clauses (e.g. 5a and b) and structures of right branching (RB) (e.g. 6a and b), given that the former has proved to be harder to understand (Just and Carpenter, 1992; Gibson, 1998).

(5)	a.	*The girl that is chasing the boy is blond.*	(Centre embedded, Subject RC)
	b.	*The girl that the boy is chasing is blond.*	(Centre embedded, Object RC)
(6)	a.	*I know the girl that is chasing the boy.*	(Right branching, Subject RC)
	b.	*I know the girl that the boy is chasing.*	(Right branching, Object RC)

With respect to questions, we included simple yes/no questions (e.g., *Did you kiss him?),* as well as adjunct questions (e.g., *When did you kiss him?*), referential (e.g., *Which girl did you kiss?*) and non-referential argument questions with anímate (e.g., *Who did you kiss?*) or inanimate agents (e.g. *What did you kick?*), given that asymmetries have been revealed in children's' performance. Specifically, better performance has been observed in (a) yes/no-questions compared to *wh*-questions (Friedmann, 2002; Ruigendijk, Kouwenberg, and Friedmann, 2004; Eriksson, 2014); (b) argument compared to adjunct questions (Tutunjian and Boland, 2008); (c) non-referential (*who*) questions compared to referential (*which-NP*) ones (Avrutin, 2000; Hickok and Avrutin, 1996; Friedmann et al., 2009); (d) referential subject questions compared to referential object ones (Hickok and Avrutin, 1996; De Vincenzi, Arduino, Ciccarelli and Job, 1999; Friedmann, Belletti and Rizzi, 2009; Tyack and Ingram, 1977); (e) non-referential subject questions compared to non-referential object ones (Garaffa, 2007; Garaffa and Grillo, 2008; Friedmann et al., 2009); and (f) *what* questions compared to *who* questions (5) (Garaffa, 2007; Garaffa and Grillo, 2008).

Given that passive constructions (e.g., *The girl was kissed by the boy*) have also been considered hard to acquire (Leikin & Assayag-Bouskila, 2004; Reggiani, 2010; Oliveira et al., 2012; Stein et al., 1984), we included them in the model by making a distinguish between short (e.g., *The cat was kicked*) and long passives (e.g., *The cat was kicked by the boy*), since the former constructions are easier to understand (Fox and Grodzinsky, 1998; but see Hirsch and Wexler, 2006).

Moreover, we included both affirmative and negative sentences (e.g., *I had not had dinner when they came in*), driven by the fact that children have more difficulties while interpreting negation (Vender and Delfitto, 2010; Scappini, 2015). Prepositional phrases were also included in the model due to their complex intrinsic properties. On the syntactic level of analysis, prepositions are classified as being either semantically coloured with a primary lexical and spatial value, or semantically weak and marking syntactic function within the prepositional phrase (Morgenstern and Sekali, 2009). Given that asymmetries have been attested in children's' performance (Abramovici, 1983; Morton and Patterson, 1980; Flores d'Arcais, 1984), we included lexical prepositions (*down, in, in front of, on, above, behind, under, between, besides, upstairs, near, out, over, outside, inside, through, next to, up*), functional prepositions (*of, about, for, from, at, with, by*) and semi-lexical prepositions (*to, before, after, until*).

There is also evidence that reference assignment is particularly problematic during acquisition. For instance, a proper name (such as in e.g. 7) is likely to refer to the same referent, whereas pronouns (*him, her*) (e.g. 8) are more fluid and their resolution depends on shifts in discourse focus. Thus, there are several studies that report that while the interpretation of reflexive pronouns (*himself*, *herself*) is well preserved (e.g. 9), the interpretation of personal pronouns (*him*, *her*) is impaired (e.g. 8), since children often erroneously allow personal pronouns to co-refer with an inter-clausal antecedent (Chien and Wexler, 1990; Grodzinsky and Reinhart, 1993; Avrutin et al., 1999; Grodzinsky, Wexler, Chien, Marakovitz, and Solomon, 1993). Evidence from studies on reading confirms that reference assignment is a complex procedure, since less skilled readers are less successful at pronominal resolution compared to skilled ones (Waltzman & Cairns, 2000; Oakhill and Yuill, 1986; Ehrlich, Remond, and Tardieu, 1999; Packenham, 1980). This led us include in the model both pronouns (personal and possessive), and anaphors (reflexives, as well as reciprocal pronouns (e.g., *The kids are washing each other*)).

(7) Lisa washes Mary.
(8) Lisa is washing her.
(9) Lisa is washing herself.

Moreover, conflicting results have been reported for the interpretation of constructions with quantifiers. While there are studies reporting no problems in dyslexia (Arosio et al., 2016), others report significant impairments (see Fiorin, 2010), even at the ages between 8 and 10 years (although by this age, according to the most known studies of quantifier-spreading (see Philip 1995; Crain et al. 1996), children should have reached adult performance.

Problems with productive and receptive (morpho-) syntactic skills have also been attested in children with dyslexia (Scarborough, 1990, 1991; Lyytinen et al., 2001). In these children, speech perception deficits apparently interfere with learning to read as well as with the acquisition of other aspects of language, including morphology (Bishop, 1997b; Kamhi & Catts, 1986; Joanisse & Seidenberg, 1998; Shankweiler et al., 1995; Vogel, 1977; Wiig, Semel, & Crouse, 1973). Specifically, several studies report that even 8-year-old poor comprehenders, with average for age nonword reading scores, have difficulties with past tense formation (see Nation, Snowling, and Clarke, 2005; Joanisse et al., 2000), while others report additional difficulties with verb agreement marking (Casalis et al., 2012; Cantiani et al., 2013; Joanisse, Manis, Keating, & Seidenberg, 2000; Rispens & Been, 2007; Rispens, Roeleven, & Koster, 2004). To investigate the interphase between morphology and syntax, we included morphosyntactic aspects such as Subject-Verb agreement (e.g., *She has/ She does*) as well as inflectional suffixes for regular (-ed forms) and irregular tense marking, number (e.g., *match/ matches*) and aspect (perfect forms (e.g., *have eaten*) and imperfective forms (e.g., *I am eating*)).

Derivational processes seem also to be difficult for poor comprehenders, given that derivations generally involve relatively more complex formations (Carlisle, 2003; Marslen-Wilson, 2001). Morphologically derived words make up 40% of unfamiliar words children encounter in text in their late elementary school years (Nagy & Anderson, 1984; Nagy et al., 1993). Research with typically developing readers has found that inflected words are easier to learn than derived ones (Carlisle, 1995). This may be because the morphological changes of derived words are less predictable and reliable, compared to inflected words, which might hinder poor comprehenders' progress in learning derived forms. Within the domain model, thus, we included several derivational prefixes and suffixes.

5. Greek Domain Model for children with dyslexia

This section describes the development of the Greek Domain Model for children with dyslexia. The Greek domain model was based on the domain model included in the ILearnRW system, a digital application of assistive learning for children with Dyslexia that was funded by the EU under the FP7 Call and constitutes the precursor to iRead. The iLearnRW domain model addressed language areas related to decoding, while no mention was made to sentence structure. Therefore, the language categories in the levels of Phonology and Orthography of the iLearnRW domain model were used as the basis on which the iRead Domain Model was built. Additionally, relevant literature findings were used to revise the linguistic information that was drawn from iLearnRW as well as to select the categories and features to be included in the levels of Morphology and Syntax. Similar to the English Domain Model for dyslexia, the levels of phonology and orthography and morphology were built using a different methodology to those of syntax and morphosyntax, which are described separately in 5.1.1 and 5.1.2 respectively.

5.1 Methodology

5.1.1 Phonology, orthography, morphology

5.1.1.1 Selecting language categories

The language features that were included in these levels were drawn from the iLearnRW domain model. In order to revise the selection of these items, there was an extensive search in the literature related to the common phonological and morphological problems encountered by Greek children with dyslexia in reading and writing. When the relevant literature was unable to provide sufficient information, findings on populations with developmental language disorders as well as children of typical language development were utilised. Finally, features were checked for frequency in a corpus of school books that are available online[16], so that very infrequent structures or structures that do not appear at all were removed from the model.

[16] http://ebooks.edu.gr/new/allcourses.php

5.1.1.2 Defining difficulty

In order to rate the linguistic features included in the Greek domain model for the levels of Phonology, Morphology and Orthography, an initial literature search was conducted. Features were placed in a scale of relative difficulty based on their order of acquisition and the degree of difficulty they pose to children with dyslexia. This process lead to a preliminary taxonomy of features with respect to difficulty, which will be tested through the collection of empirical data by administering the iRead games to a sample of children with dyslexia. The empirical data collected will be analysed using the Rasch measurement model in order to obtain a relative scaling of features with respect to difficulty, based on children's performance in them.

5.1.2 Syntax

5.1.2.1. Selecting language categories

The selection of syntactic phenomena was based on extensive literature research on dyslexia, typical reading development, and language acquisition. Specifically, we included function words such as prepositions, clitics, and negation markers, as well as phenomena such as embedding, passivization and reference assignment (for details see Appendix II). Due to time limitations, we decided to limit ourselves on including three syntactic phenomena that are well described in the literature, i.e., **embedding, passive morphology and discourse anaphors**. The selected language areas are relevant to the purposes of the present project since they cover a big scale of syntactic phenomena very frequently encountered in texts and everyday communication. All three are defined as "complex" syntactic processes; embedded clauses contain extra syntactic information that the reader needs to keep in memory and incorporate it to the main clause, passives contain morphological information that the reader needs to identify to correctly to understand the structure, and anaphors contain discourse information important for the correct interpretation of the construction (for details see 5.2).

5.1.2.2. Defining difficulty

The difficulty of the chosen structures was investigated using psycholinguistic experiments that assess speaker's processing of syntactic structures. Specifically, the self-paced reading technique was used to assess the way children process a variety of syntactic structures. The technique provides researchers with the opportunity to investigate participants' online processing of visually presented sentences by measuring reading times on different parts of each sentence. To this purpose, sentences are broken into segments and presented to participants one by one on a computer screen, in a speed controlled by the participant by pressing a key in the keyboard. Reading times on each segment of the sentence are recorded and compared across different syntactic structures. Delayed reading indicates higher complexity, which makes the technique very suitable for the purposes of the project. The experiments were run in two phases: a pilot phase, which took place in spring 2017 and provided us with preliminary findings on the relative difficulty of the investigated structures, and an experimental phase, which is taking place in winter 2017-18. Results from the pilot phase are presented here (see section 5.2 and Appendix III), as they were used to formulate a preliminary ranking of features with respect to difficulty, which will be revisited when results from the experimental phase have been obtained and analysed.

Five (5) children with dyslexia and seven (7) typically developing age-matched controls took part in the pilots. The children were aged between 8 and 10 years, they were all monolingual Greek children and were randomly selected from schools at Ioannina to participate in the study. The main experimental phase, which will be completed in spring 2018, has recruited ninety (90) children in total: thirty (30) children with dyslexia aged between 8 and 10 years, thirty (30) age-matched controls and thirty (30) typically developing novice readers, aged between 7 and 9 years.

Materials: Specifically, with respect to embedding, we decided to compare relative, adverbial and complement clauses to test which clause type is harder for the children. Given that object (OVS[17]) constructions are particularly hard for children to comprehend, we included both subject and object extracted constructions (e.g. *The man saw the doctor that hit the athlete* vs. *The man saw the doctor that the athlete hit*) within the relative clause condition to investigate how word order affects complexity. Within the adverbial clauses we tested adverbials clauses introduced by four different temporal connectives *(afu =after, eno = while, otan =when, prin =before)*, to investigate whether the adverbial type will cause differences in children's performance.

To investigate how passive morphology is processed by children, we included active and passive constructions with or without an agent *by phrase*, as well as verbs with passive and active morphology with either passive, reflexive (*shave*), unaccusative (*break*) or reciprocal (*hug*) meaning.

With respect to anaphors, we included constructions with definite and indefinite noun phrases, proper names, personal pronouns and null subjects, in subject and object position. Given that discourse reference /anaphoric resolution requires extra processing resources due

[17] OVS = Object-Verb-Subject word order.

to ambiguity they cause, the results will be very informative on the underlying difficulties children encounter.

Experiments contained 470 sentences in total (200 assessing passive constructions, 90 assessing anaphors, 180 assessing embedding). Sentences were segmented in five reading areas: the "NP1"[18], "Intervening PP"[19], "Verb," "Critical region," "Extra region" as illustrated in (10) and (11) below. Slashes indicate segments. Each segment appeared in the middle of the screen and was replaced by the following segment after the participant pressed the spacebar. The last segment appeared with a full stop to indicate the end of the sentence. The experimental items were matched on syllable length, while verbs were additionally matched on frequency (ILSP written Greek Corpus (Hatzigeorgiou et al., 2000)).

(10) O antras / me ta mavra jalja/ sinelifθi / apo tin astinomia / to vraδi./ (passive with by phrase)
The man / with the black glasses / was arrested$_{PASS}$ / by the police / in the evening.
=The man with the black glasses was arrested by the police in the evening.

(11) O antras / me ta mavra jalja/ sinelifθi / apo traɣiko laθos / to vraδi./ (passive with PP)
The man/ with the black glasses / was arrested$_{PASS}$ / by a tragic mistake / in the evening.
= *The man with the black glasses was arrested by a tragic mistake in the evening*.

Sentences within each condition were similar but differed only in the critical region. Plausibility rates were also collected to ensure sentences' acceptability. Specifically, 100 graduate students were instructed to evaluate our stimuli in a scale of 1–5, with 1 being the score for an unacceptable sentence in Greek and 5 for a fully acceptable one. Sentences with scores lower than 1 SD[20] from the mean were modified and a second norming took place. All experimental materials were divided into 4 lists in a Latin Square design, creating four different experimental sets, each containing one version of the experimental sentences. All items were randomised and presented in a set order. Participants were presented with all four lists, in separate days.

Procedure: An online self-paced reading (for details see Jegerski, 2014) psycholinguistic experiment was designed (addressing all three linguistic phenomena) to evaluate in real time the difficulties children encounter while reading. All testing was done individually, in a quiet room at primary schools, with children seated at a comfortable distance from a computer screen. The self-paced reading task was run on a laptop using E-prime 2.0 professional software (Psychology Software Tools, Pittsburgh, PA), which collected and recorded response

[18] NP = Noun Phrase
[19] PP = Prepositional Phrase
[20] SD = Standard Deviation

time and accuracy data. Initially, participants were given detailed instructions about the experiment and 10 practice trials were used as model for the participants to familiarize themselves with the task and ask clarification questions. All stimuli were visually presented to the participants in a centered noncumulative display, meaning that every segment appeared in the center of the display screen and only one segment was visible at a time. Specifically, participants read a sentence word-by-word or phrase-by-phrase, pressing a button to display consecutive segments of text in the center of the screen. The time taken to press the button provides an indication of the processing difficulty at each segment. Response latencies were collected for each segment and were measured from the onset of the target word. Often, this task is used to investigate how people interpret difficult parts within a sentence, such as ambiguous words. At the end of each sentence, participants were asked to answer as quickly and as accurately as possible a simple post-stimulus distractor question, by choosing between two options. Participants had to press one of two pre-specified color-coded buttons (either the answer on the LEFT side “a” or on the RIGHT side “l” key), of the QWERTY keyboard. Post-stimulus questions were included to insure that participants would pay attention to the presented materials, and only correct responses were further analysed. Participants could stop the experiment at any time. The average duration of testing was around 15 minutes.

5.2 Language categories included in Greek domain model for children with dyslexia

This section presents a brief account of the language patterns found in the relevant literature on dyslexia, which lead to both the selection of language categories and to ranking them with respect to difficulty. The description begins with a summary of the language levels and categories included in the model, which is given in Table 3.

Table 3 Summary of Greek domain model

Linguistic Level	Language Category	N of Features
Phonology	Syllable patterns	46
	Phonemes	64
	Grapheme-Phoneme Correspondence	109
Orthography	Letter visual similarity	28
	Letter similarity – double consonants	8

Morphology	Prefixes	13
	Suffixes: Inflectional	80
	Suffixes: Derivational	47
Syntax	Function words	41
	Embedding	15
	Discourse anaphors	9
	Voice (passive)	15

The linguistic levels and language categories summarised in Table 3 are analysed in the following paragraphs. The description first analyses the contents of the levels of phonology, morphology and orthography, followed by a description of the findings that led to the selection and ranking of syntactic features with respect to difficulty, with a presentation of the preliminary findings of pilot experiments.

Beginning with phonology and orthography, research has shown that phonological processing skills are necessary in reading comprehension, writing and spelling (Voulgaris, 2010). In order to learn how to read and write, children have to understand that a language is distinguished in phonological units and that these phonemes are written in graphemes (Panteliadou, 2000). Children with dyslexia have often exhibited poor vocal perceptions (Zachos, 1992), confusion of phonemes with visual or acoustic similarity (Kourakis, 1997) and wrong pronunciation of vowels and consonants (Zachos, 1992). Specifically, children with dyslexia exhibit poor skills in phonological awareness tasks like eliminating, substituting or reordering phonemes. These children have major difficulties in grapheme-phoneme correspondence (Zafiropoulou et al., 2004).

Greek is characterised by a transparent (shallow) orthography (Ktori & Pitchford, 2008; Miles, 2000; Whitehurst & Lanigan, 1998) with relatively high degree of grapheme to phoneme correspondence. Yet, there are certain graphemes that do not correspond to a single phoneme but to more (letter ***υ*** for example, represents phonemes /i/, /f/ and /v/). Due to its orthographic transparency, Greek is thought to cause less difficulty to children with dyslexia compared to deep orthographic systems like English (Miles, 2000). However, words with low grapheme-phoneme correspondence often cause difficulties to children with dyslexia and the difficulty of matching graphemes to phonemes is often highlighted in the literature (Pavlidis & Giannouli, 2003).

Letter similarities often cause difficulties in reading, writing, spelling, handwriting (Bright Solutions for Dyslexia, 2010). In relation to specific graphemes, Greek students with dyslexia have difficulty with digraphs, that is, grapheme combinations that correspond to single phonemes, like *μπ*, which corresponds to the phoneme /b/ (Chliounaki & Bryant, 2003; Aidinis, 1998). Particular difficulty is also observed in letters with visual similarity like *α-ο*, *ε-ω*, *β-Φ*, *β-ϑ*, *6-ϑ, ζ-ξ, η-ω, κ-χ, ρ-9, κ-γκ* and *π-μπ* (Karakirgiou, 2017). Additionally, Papadopoulou & Zafiropoulou (2005) report that the most common initial consonant errors are for letters δ, θ and ν, which are typically pronounced as /ð/ and /θ/ and /n/ respectively. There is a mutual confusion between /v/ and /ð/, as well as /θ/ and /ð/, which are phonemically quite different but retain a certain visuo-motor association and the most frequently misread initial letters are *χ*, *μ*, *π* and *ρ*. In their research, the rates of confusion were affected by the position of the letter in the word.

Syllabification was also found to be the most common type of mistakes in children with learning difficulties (Sakelariou et al., 2007). Generally, there is difficulty in polysyllabic and unfamiliar words (Anastasiou 1998). Syllable tokens being CV[21], followed by V (17.3%), CCV (12.5%), CVC (10.1%), CCVC (2.1%), VC (1.5%), CCCV (.5%) and rare more complex structures (less than 0.1% each) (Protopapas et al., 2012).

As far as morphology is concerned, morphological awareness strongly affects the process of learning to read (Carlisle, 2003). Researchers (Carlisle, 1987; Tyler, 1986) have highlighted that students face problems when attempting to understand suffixes and the relations between morphologically related words. Words are easier to decode when suffixes are frequent and highly productive (Mann & Singson, 2003; Cole et al., 2011). In Greek, inflectional suffixes are frequent and apply to large classes of words, while derivational affixes are not as broadly applicable or as frequent as inflectional ones but their spelling is every bit as arbitrary and lexically unmotivated as that of inflections (Protopapas, 2014).

Based on the information described so far, it was decided to include the categories of Syllable patterns and Phonemes in the level of Phonology, Suffixing and Prefixing in Morphology, and Grapheme-Phoneme correspondence, Letter similarity (visual) and Letter similarity -double consonants in Orthography.

For each language category, a number of features that instantiate this language area are listed. These instances are listed based on the following criteria:

- those instances that are most crucial to reading development, including instances that are basic and are normally taught/developed first, as well as instances that are more challenging for developing readers,
- those instances that appear in texts addressed to children of our target groups (we can omit instances that would never be encountered, if there are any). As part of this task we must construct a relevant corpus of children's books for the learner group considered to derive the frequencies.

These categories were selected based on what is found to be more difficult and most relevant to reading in general as a language skill and linguistic process. According to the literature, they are basic skills that need to be developed when learning to read. Some of them are most

---

[21] CV: C=Consonant, V=Vowel

frequent and clear-cut or are somewhat less frequent but still found in children's books and are considered more complex.

Specifically, with respect to the category of Syllable Division, a variety of combinations of syllable structures were included in the Domain Model, such as Open-Open (like CV-CV), Open-Closed (like CV-CVC), and Closed-Closed (like CVC-CVC) (see Appendix II for the complete list of features). In the category of Phonemes, all Greek phonemes were included in two subcategories: (a) Consonants and (b) Vowels. Consonants were entered in relation to their position in the word (initial or internal). Phoneme categories included include plosives (/p/, /t/, /d/ etc.), nasals (/m/, /n/), fricatives (/f/, /θ/, /v/ etc.), sibilant (/s/, /z/), liquid (/l/, /r/). Consonant clusters with plosive-plosive (/kt/), plosive-fricative (/χt/) and fricative-fricative (/fθ/) combinations were also included in both initial and internal position. As for vowels, the Greek phonemic repertoire contains 5 vowels, which do not normally pose any difficulties to children with dyslexia. Therefore, only vowel combinations were included in the model. Finally, we included Grapheme - phoneme correspondence and especially regular consonant clusters, regular digraphs, irregular: 'υ' as consonant and as vowel, irregular digraph vowels with diacritics, 'i' as glide and Letter similarity (visual) including double consonants (see Appendix II for complete list).

With respect to the level of Morphology, a category of Prefixes and a category of Suffixes were included, containing both derivational and inflectional suffixes. For derivational suffixes, the most common Greek derivational processes were included, such as the formation of diminutives, deverbal nouns showing profession or human capacity, deverbal nouns referring to instruments etc. In the category of Inflectional suffixes we included suffixes of frequent and less frequent noun and adjective classes in nominative, genitive and accusative in singular and plural for masculine, feminine and neuter nouns. As for Verbs, we included suffixes that denote voice (active, passive), tense and aspect (past/non-past, perfective/imperfective). As for Prefixing, we included prefixes in Nouns, Adjectives and verbs (see Appendix II for complete list).

Finally, with respect to orthography, letter visual similarity included combinations of letters with similar shapes that often cause confusion to children with dyslexia (e.g. α and ο, φ-β-θ etc.), while Double consonants were also added to include ββ, ρρ, νν etc.

Moving to syntax, as stated earlier, four main categories were included in the domain model, one of which was part of the iLearnRW domain model (i.e. Function Words), while the remaining three were addressed in the pilot experiments: voice, embedding and discourse anaphors. Preliminary findings obtained from pilots revealed the following patterns:

With respect to voice, although no differences were revealed between sentences with agent or non-agent *by phrase* sentences (e.g., *The man with the glasses was arrested by the police at the protest* vs. *The man with the glasses was arrested by accident at the protest*), constructions with animate agents were easier to interpret compare to sentences with inanimate agents. No differences were yielded when comparing active and passive voice constructions with non-agent *by phrases* (e.g., T*he man with the glasses was arrested by accident at the protest* vs*. The man with the glasses arrested the homeless man by accident),* although a slightly delayed performance was occurred while reading active constructions. In structures with reflexive verbs (*shave*), we compared constructions with or without a disambiguating *apo-phrase* (e.g. T*he man with the glasses shaved yesterday by himself* vs*. The man with the glasses shaved again after a long time*), and constructions with an agent *apo-phase* (e.g., *The man with the glasses was shaved by the barber yesterday).* Constructions without a disambiguating *apo-phrase* (corresponding to the English *by*-phrase) were the hardest for children with dyslexia, while structures with a disambiguating *apo-phrase* were the easiest for both typically developing and children with dyslexia. In structures with unaccusative verbs (*break*), when comparing sentences with or without a *by phrase (e.g., The sail of the boat was torn by the air yesterday* vs*. The sail of the boat was torn into pieces yesterday),* we found that although typically developing children do not differentiate between the two, children with dyslexia have more difficulties in constructions containing a *by phrase*. In structures with reciprocal verbs (*hug*), we compared structures with or without a prepositional phrase (e.g., *The children with the blond hair were hugged by the teacher in the afternoon* vs*. The children with the blonde hair hugged each other in the afternoon)* and structures with transitive use (e.g., *The children with the blonde hair hugged the teacher in the afternoon)*. Both structures with reciprocal verbs in transitive contexts and those without a prepositional phrase were easier compared to those with a *by phrase*. In structures with deponent verbs (*sleep*), transitive and intransitive contexts were included (e.g., *The student with the blonde hair fell in love with the actor immediately* vs. *The student with the blonde hair slept at the porch immediately*); although no difference was attested in typically developing children, children with dyslexia have more difficulties with the transitive use of these verbs. A table summarizing our findings is provided in Appendix III).

With respect to embedding, we compared relative, adverbial and complement clauses; overall, adverbial clauses were the hardest for both typically developing and children with dyslexia. Regarding complement clauses, typical developing children performed slightly better on complement clauses introduced by 'pu' (*that*) (see e.g. 12)*,* compared to 'an' (*if*) (e.g. 13) and 'oti' (*that*) clauses (e.g. 14), whereas children with dyslexia showed a different preference, with 'an' and 'pu' clauses to be the easiest compared to 'oti' clauses.

(12) O adras apori **an** o athlitis filise ti jineka.
The_$_{NOM}$ man_$_{NOM}$ wonders if the_$_{NOM}$ athlete_$_{NOM}$ kissed_$_{3SG.PAST.PERF}$ the_$_{ACC}$ woman_$_{ACC}$.
= The man wonders if the athlete kissed the woman.

(13) O adras herete **pu** o athlitis filise ti jineka.
The$_{_NOM}$ man$_{_NOM}$ is glad that the$_{_NOM}$ athlete$_{_NOM}$ kissed$_{_3SG.PAST.PERF}$ the$_{_ACC}$ woman$_{_ACC}$.
= The man believes that the athlete kissed the woman.

(14) O adras pistevi **oti** o athlitis filise ti jineka.
The$_{_NOM}$ man$_{_NOM}$ believes that the$_{_NOM}$ athlete$_{_NOM}$ kissed$_{_3SG.PAST.PERF}$ the$_{_ACC}$ woman$_{_ACC}$.
= The man believes that the athlete kissed the woman.

Regarding adverbial clauses, no significant differences were attested between groups and categories (we tested adverbials clauses introduced by four different temporal connectives *afu =after, en =while, otan =when, prin =before)*, although children with dyslexia performed slightly worse on adverbial clauses introduced by the adverbial 'eno' (*while*) (e.g., *The man hit the boy while kissing the woman*).

Regarding relative clauses, center embedded (CE) constructions (e.g., *The man that hit the athlete is old*) were harder compared to right-branching (RB) (e.g., *The man saw the doctor that hit the athlete yesterday*) for both groups, although no differences were observed between subject and object-extracted constructions (e.g. *The man saw the doctor that hit the athlete* vs. *The man saw the doctor that the athlete hit*) for either CE or RB clauses. However, the relative pronoun seems to affect children's performance, since longer reaction times were attested in both CE and RB clauses, when 'o opjos' (*who*) relative pronoun (which in Greek is specified for gender, number and case) was used compared to 'pu' (*that*) complementizer which is indeclinable.

With respect to anaphors, we compared constructions with definite and indefinite noun phrases, proper names, personal pronouns and null subjects, in subject and object position. Our results revealed that both groups had greater difficulties with pronouns; however, while typically developing children performed better on constructions with null subjects, children with dyslexia performed better on constructions with proper names.

6 Supportive actions in domain model development and implementation

6.1 Representation of Domain Models as Graphs

Domain models are stored within the iRead infrastructure as *weighted directed acyclic graphs* (DAGs for short). A *directed graph* can be defined as a set of *vertices* and a set of pairs of vertices, called *edges*. The first vertex defined by an edge is its *source*, and the second vertex its *target*. We say that a graph is *acyclic* if there is no directed cycle. Furthermore, all edges are bind with a non-negative value, called *weight*, thus resulting to a *weighted graph*.

For the representation of a domain model, vertices of the graph correspond to the language features while directed edges indicate prerequisites. Naturally, in such a representation, there can be no directed cycle, while the weight on the edges allows to express the fact that two different language features can be both prerequisites of a third language feature, but one of them is more important than the other. In the case where a language feature has prerequisites, the sum of the weights of its incoming edges must be equal to 1.

Using weighted DAGs allows to efficiently instantiate user models and store a user's progress. User models are instantiations of domain models, which are personalised to each learner via their interactions with the games and the e-Reader. User models are also represented as weighted DAGs (in fact a copy of the corresponding domain model graph), where each vertex is equipped with an additional attribute, called *competence*. Competence is a non-negative integer value showing the current mastering of the underlying language feature. When a user exercises a particular language feature, the value of the competence can increase or decrease, depending on her performance. The value of competence cannot vary arbitrarily: for every language feature we define its *minimum* and *maximum value*, in order to define the allowed values for competence. By default, and unless otherwise specified within the domain model, we consider the minimum value to be equal to zero, and the maximum value equal to 10.

For a particular user, not all language features are "available" for practicing. It is clear that a user can practice a language feature if she has mastered the feature's prerequisites. In order to express the "availability" of a language feature, we equipped both vertices and edges of the graph with the *unlock_value* attribute. The availability status of a vertex (or edge) can be either "locked" or "unlocked", i.e. available or non-available respectively. An edge becomes "unlocked" if the competence of its source is higher than the unlock_value of the edge, otherwise it is "locked". Vertices can be "unlocked" in two ways: (i) if the sum of the weights of their incoming "unlocked" edges is higher than the unlock_value of the vertex, or (ii) if the competence of the user is higher than a *threshold value*. As an example, consider the configurations shown in Figure 1: language features are drawn as rectangular shapes and their label shows the language feature and the competence of the user (e.g. the bottom vertex has competence 0 in the first two configurations and competence 5 in the third one). For the sake of simplicity we assume that the unlock_value of all vertices is equal to 0.75 and the threshold value for unlocking is equal to 5. The unlock_value for the edges is denoted within a yellow circle, while the weight of each edge is also present. An immediate consequence of the above unlocking procedure is that a language feature can become active even if the user's competence on the prerequisites cannot unlock incoming edges. Hence, the subgraph of the active vertices and edges, is not necessary a connected graph. This property can be useful in the case where the initialization of a user's model is based on an undertaken test; the domain model contains many language features, and it is almost impossible to test on every feature. Still, the available vertices will allow the user to practice different areas of the domain model in parallel.

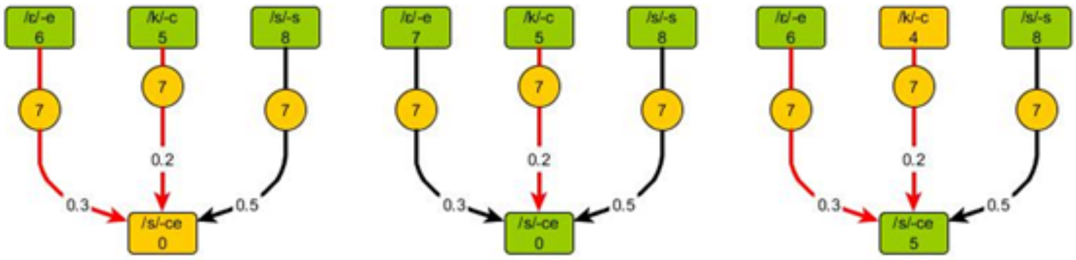


Figure 1 Active language features are highlighted in green, inactive in yellow. Active edges are drawn black, inactive red. We assume that the unlock_value of all vertices is equal to 0.75 and the threshold unlock value is 5. In the leftmost case the bottom language feature is not available as the sum of the active incoming edges is 0.5<0.75. In the middle configuration the corresponding sum is 0.8>0.75 and becomes active, while in the rightmost configuration, although the sum of the active incoming edges is less than the unlock_value of 0.75, the competence of the feature is 5, equal to the threshold value. Thus the feature becomes active.

Since the weighted DAG of a domain model might not be connected, we add a root vertex that is connected with edges to the language features that have no prerequisites. When a user's model is initialized, these language features (that have no prerequisites) are always available for practice. In order to satisfy this, the root vertex has minimum and maximum value equal to zero. Also its unlock_value and threshold value is zero, which implies that the root vertex is always unlocked. The edges that emanate from the root vertex have zero weight and unlock_value; hence these edges are always unlocked.

6.2 Json Representation of Domain Models

In order to be able to insert and store the domain models within the iRead infrastructure, we create a machine-readable representation using the json format[22]. Json format consists of attribute–value pairs and array data types. In the following we describe the structure of the json used for representing the domain models.

Language features

In the json created for a language feature we include two “types” of attributes, namely *linguistic attributes* that are related to the information provided by the linguistic experts, and graph attributes that are related to the weighted DAG representation of the domain model.

Linguistic attributes (as they appear in the json)	
linguistic_level	The name of the **linguistic level** of the language feature, see Section 2.4
category	The name of the **language category** of the language feature, see Section 2.4
difficulty_level_index	an integer value indicating the **difficulty classification** of the language feature within its category, see Section 2.4
type	a string indicating which developed function-routine is used for checking whether a word (or sentence) contains the language feature
description	the human-readable description of the language feature
examples	examples for the language feature
frequency_in_child_text	indicates whether the language feature occurs frequently in child appropriate texts. Accepted values are “UNSET”, “LOW”, “MEDIUM” and “HIGH”
exception	exceptions related to the language feature

[22] https://en.wikipedia.org/wiki/JSON

National_curriculum_place	the place of the language feature in the national curriculum
Related_word_difficulty	an integer value indicating the proposed difficulty of words for practice[23].
Graph attributes (as they appear in the json)	
min_value	an integer denoting the minimum value for a user's competence
max_value	an integer denoting the maximum value for a user's competence
unlock_value	a positive value indicating the threshold for "unlocking" the language feature. Its value is between 0 and 1. If the user's competence satisfies , the language feature is unlocked
threshold	A positive value between 0 and 1. Indicates the amount of flow necessary for becoming active

As an example, we use the first language feature of the GPC category of the English domain model.

```
{
        "id": 2,
        "linguistic_level": "Phonology",
        "category": "GPC",
        "difficulty_level_index": 1,
        "type": "GPC(/s/-s)",
        "description": "/s/ as snake",
        "examples": "sad, sit, sip",
        "frequency_in_child_text": "UNSET",
```

[23] In the developed language dictionaries, each word is associated with a positive integer value, indicating the "word's difficulty". This classification differs across languages, and is a combination of several attributes, such as the number of syllables or the frequency of a word in child-texts. For example, in English, words with difficulty 1 are monosyllabic words, words with difficulty 2 are 2-syllable words with specific CV-patterns, etc. Language features of a domain model may indicate a specific difficulty of words for practice.

"exception": "",
"national_curriculum_place": "Y1",
"related_word_difficulty": 1,
"unlock_value": 0.75,
"min_value": 0,
"max_value": 10,
"threshold": 0.5
}

The root is a special vertex with the following attributes:

{
"id": 1,
"unlock_value": 0,
"min_value": 0,
"max_value": 0,
"threshold": 0
}

Prerequisites

In our graph modelling, edges are always weighted and directed, indicating that the language feature represented by the source-vertex is a prerequisite for the language feature represented by the target-vertex. All edges have the following attributes:

source_id	the id of the source-vertex language feature
target_id	the id of the target-vertex language feature
weight	a value between 0 and 1 indicating the weight of the edge
unlock_value	a value between 0 and 1 indicating the condition for the edge to unlock

Note that the unlock_value of an edge is expressed as a positive value between 0 and 1; the formal condition for the edge to unlock is , where competence, min_value and max_value are the corresponding attributes of the source-vertex of the edge.

As an example, in the English domain model, the first language feature of the GPC category is a prerequisite of the second language feature of the same category:

{
"source_id": 2,
"target_id": 3,
"weight": 1.0,
"unlock_value": 0.75
}

A condition that always holds for any language feature with incoming edges , is that , i.e. the sum of the weights of the incoming edges is equal to one. The only exception to this rule, are the incoming edges of those features connected to the root-vertex. Recall that these edges were added in order to make the graph connected. Hence for such edges we have:

```
{
    "source_id": 1,
    "target_id": 2,
    "weight": 0,
    "unlock_value": 0
}
```

where the root-vertex has id=1 and the target-vertex is the first language feature of the GPC category of the English domain model.

Unless otherwise specified, for any language feature with prerequisites, the incoming edges are equal-weighted. We expect the weights to be updated by an adaptation procedure which will take into account users' performance-data from the trials and decide if and how weights should be redistributed.

Domain Model

Language features and prerequisites are bind together to a single json for the representation of a domain model. In addition, in the domain model, several language features might have been grouped (see for example the Greek domain model). This grouping indicates that the language features of a particular group are highly related and should be practiced together; for example literacy games could present content from all language features of a group in the same game, or, when practicing a language feature of a group, the other language features could be used as distractors in the literacy games. An additional requirement for some group, could be that its language features become active simultaneously. This can be achieved by defining the same prerequisites for all the vertices of the group. As an alternative, a "dummy" vertex could be added with incoming edges from the prerequisites of the group and outgoing edges towards the vertices of the group. Such a vertex would assure that even if the weights of the edges are changed, the language features of the group always become active at the same time.

Each group is described by a name and a list of the language features it contains.

name	the unique name of the group
items	the ids of the language features that belong to the group

As an example, consider the first group defined in the Greek domain model. Its name is "group A", and it contains language features #3-5 of the Syllable Patterns category. These language features have assigned id 4, 5 and 6 respectively in their json representation.

{

```
    "name": "group A",
    "items": [4, 5, 6]
}
```

Combining all the above, the json for the domain model contains a total of three attributes:

features	a json array that contains information about the language features of the domain model. Each element of the array is a json representation of a language feature
edges	a json array that contains information about the prerequisites described in the domain model. Each element of the array is a json representation of a prerequisite
groups	a json array that contains information about the groups defined in the domain model. Each element of the array is a json representation of a group of language features.

The general structure of the json representation of the domain model is the following:

```
{
    "features":[{root-vertex}, {language feature#1}, {languagefeature#2}, …],
    "edges": [{prerequisite#1}, {prerequisite#2},…],
    "groups": [{group#1}, {group#2},…]
}
```

6.3 Tool for Editing Domain Models as Graphs (Editor)

In this section we give a brief description of a desktop application developed in Java for editing the domain models as graphs. The motivation for this tool, is the fact that commonly used programs for storing the domain models (e.g. excel) do not allow the definition of prerequisites. Our aim is to visualize the underlying graph, allow the user to insert vertices (language features), add or remove edges (prerequisites), focus on subgraphs e.t.c., while the graph is always acyclic.

The initial screen is disabled allowing the user to make an action using the menu; see Figure 2. The available options are to create a new domain model from scratch or open an existing file.

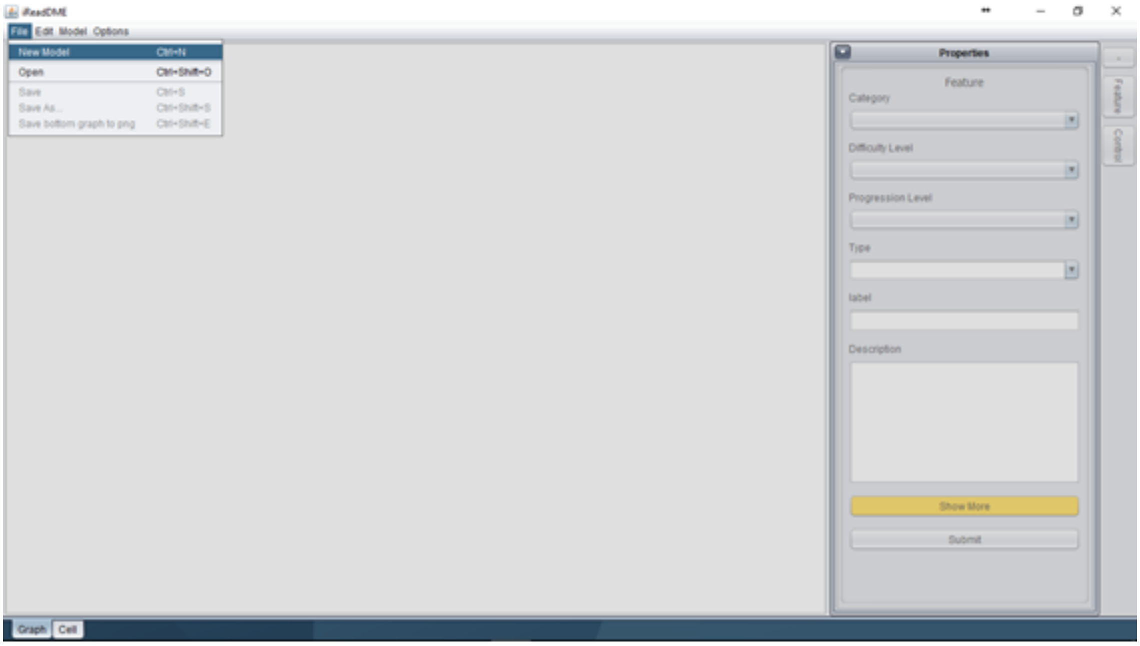


Figure 2 Welcome screen.

Supported file-extensions are:

- **.csv (comma separated):** each row of the file corresponds to a single language feature and its attributes are given in a predefined order. In order to create unique keys for the language features, we impose the following two restrictions: (i) the type of each feature cannot be empty, and (ii) the description of each feature is unique. The input from a csv-file does not support information about prerequisites.
- **.ireadme:** this file extension is supported only by the editor. Practically this is a json formatted file that stores information about the language features, prerequisites and the details of the view.

When a file is opened, a default layout is used for the vertices. For each language feature a vertex is created and features are grouped based on their difficulty within their language category; we refer to these group vertices as “difficulty-vertices”. Difficulty-vertices that belong to the same language category have the same y-coordinate, and their x-coordinate increases along with the difficulty. On the other hand difficulty-vertices that belong to different language categories are sorted vertically based on the lexicographical ordering of their category.

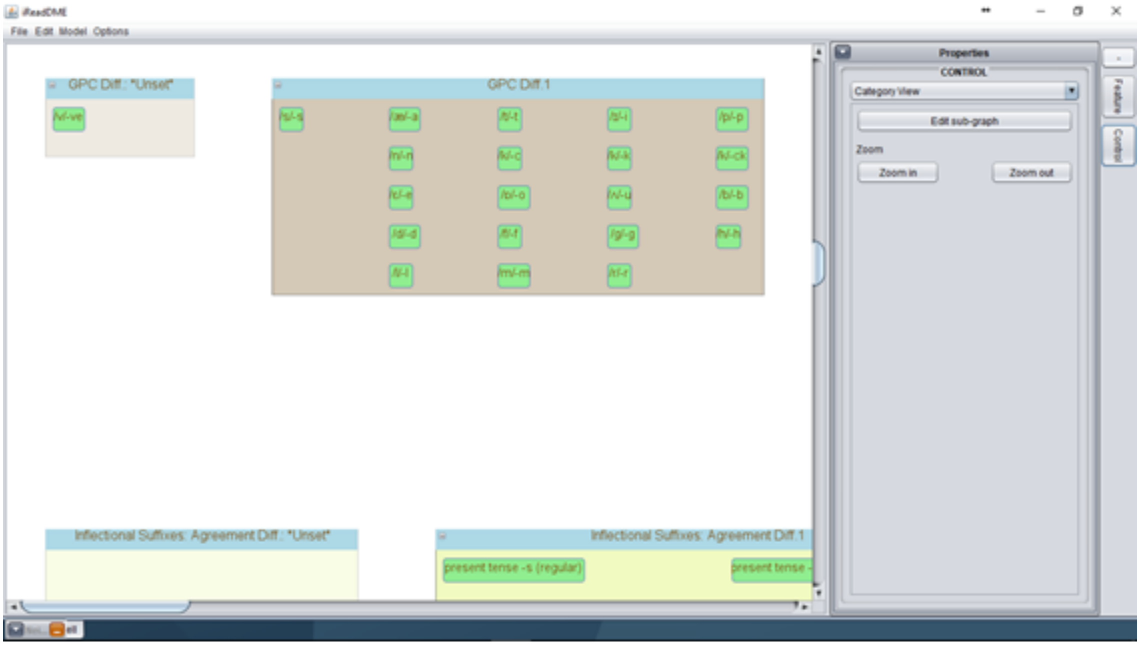


Figure 3 The csv-file of the English domain model is opened. Language features are "grouped" based on their difficulty. The "graph view" to the right allows for zooming and focusing to subgraphs.

Note that there can be difficulty-vertices that do not contain any language feature. Also, for every language category, a special difficulty-vertex is added as leftmost: it contains those language features for which the difficulty is not yet defined. For example in Figure 3, there exists a language feature of category GPC that has unset difficulty. This special difficulty-vertex allows the user to allocate the language features with undefined difficulty. In addition, difficulty-vertices of the same category have the same background colour, allowing for the user to visually distinguish them.

The interaction offered on the graph is somehow limited, due to the size of the whole domain model. The user can:

- add a directed edge: the directed edge can have as endpoints either a language feature or an entire difficulty-vertex. The latter option is useful in the case where the user wants to declare, for example, an entire difficulty-vertex to be prerequisite for some language feature. Instead of creating all the edges, it suffices to draw an edge from the difficulty-vertex to the language feature; the editor calculates the prerequisites in a background process. In Figure 4, we can see such an edge connecting two difficulty-vertices of the Syllables category.
- move selected vertices to a difficulty-vertex. When a user moves elements, the default behaviour is that all edges incident to these elements are removed from the graph. The reason for this choice, is that directed cycles are not allowed, therefore if a cycle is to be created when a language feature is moved, the entire action is cancelled by the editor. However, this behaviour can be changed through the options menu.

Additional functionality is offered with a right-click popup menu; refer to Figure 4. The user can select one of the following options:

- delete the selected elements (edges or vertices)
- move the selected elements (an appropriate menu is shown for choosing the destination for the moving operation)
- create a new language feature (by adding the feature's attributes in the form shown in Figure 5)
- mark the selected vertices as "sources" or "targets". This functionality allows the user to add a bunch of edges at the same time. When vertices are marked as "sources" ("targets") and the user marks his selection as "targets" ("sources" respectively) then edges are added from all the "sources" to all the "targets".
- clear the marked "sources" and "targets"

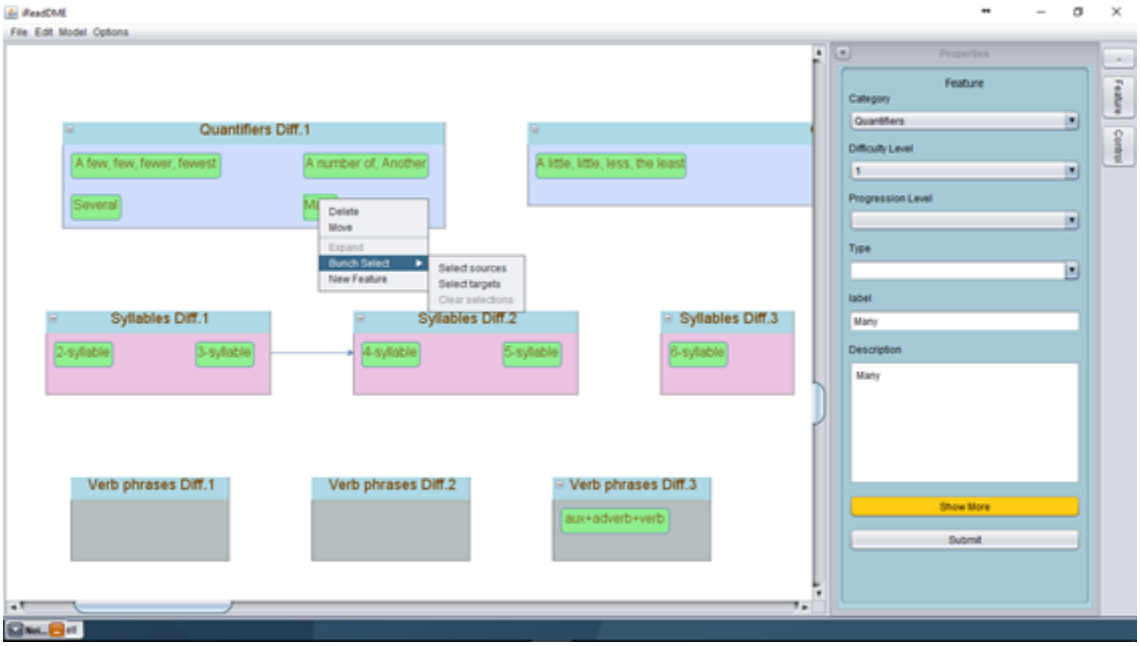


Figure 4 the right-click menu on the entire graph.

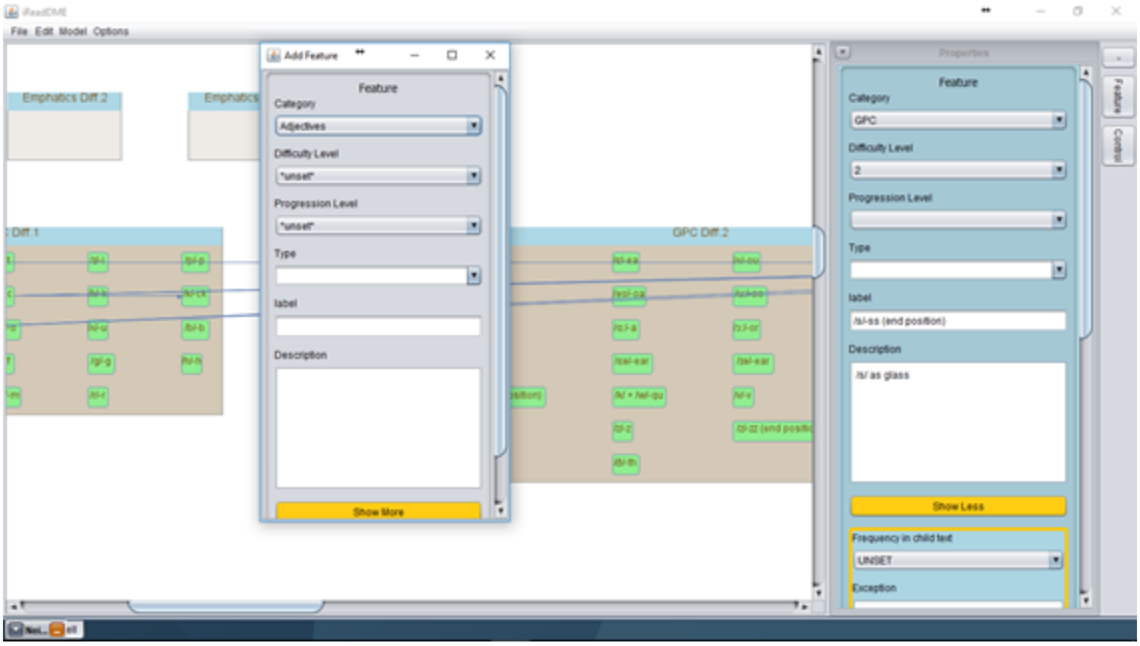


Figure 5 The form for creating a new language feature

The right panel of the view is the "control panel". The control panel can be collapsed or expanded, and it has two different views: the "graph view" and the "feature-properties view". The "graph view" (shown in Figure 3), allows the user to zoom in or out the current graph or to focus to a subgraph of the whole domain model, as we shall shortly see. The "feature-properties view" (shown in Figure 6 and Figure 7), can function in two different ways. If no language feature is selected (Figure 6), its background colour is grey, and the user can define a new language feature by inserting data in the corresponding entries. Since there exist many attributes for each feature, the more "important" ones, i.e. those that cannot be empty are shown first. The remaining attributes are shown if the user hits the "Show More" button, highlighted in yellow color; refer to Figure 8. The "Submit" button adds the language feature in the domain model and updates the view.

The second functionality of the "feature-properties view" occurs when a user selects a specific vertex from the graph. The background of the "feature-properties-view" becomes light-blue and the attributes of the selected language features are shown. The user can edit these values and update the domain model and the view with the "Submit" button; see Figure 7.

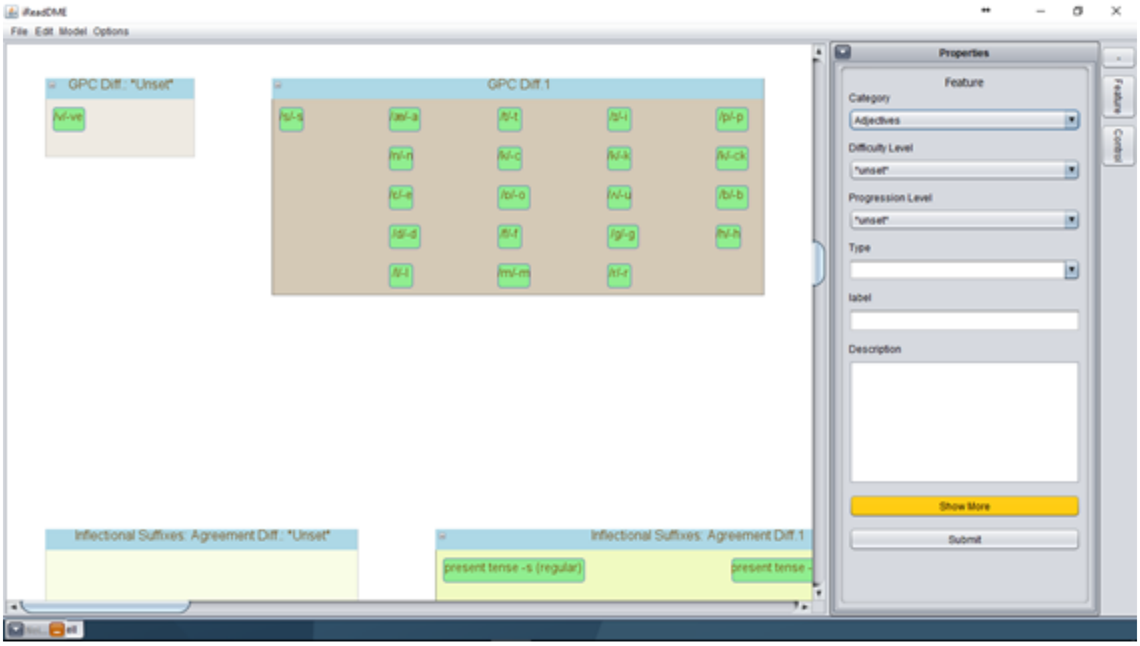


Figure 6 The "feature-properties view" to the right allows the user to define new language features.

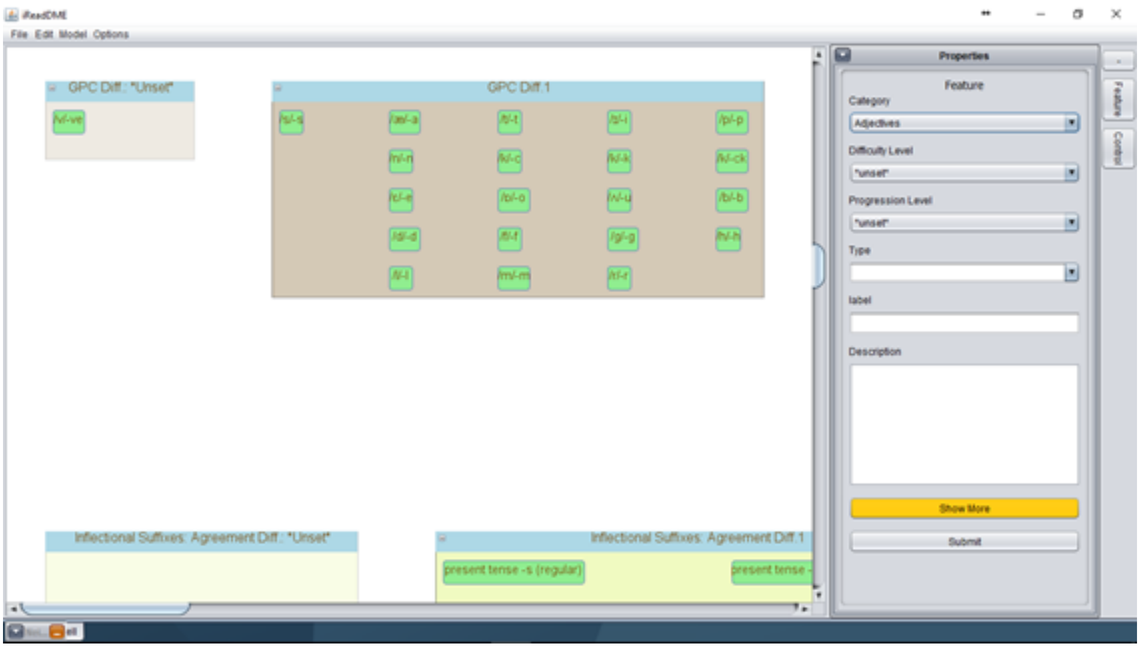


Figure 7 The "feature-properties view" to the right, allows the user to change the attributes of the selected language feature. The neighbourhood popup window shows the selected feature and its neighbouring vertices (if any).

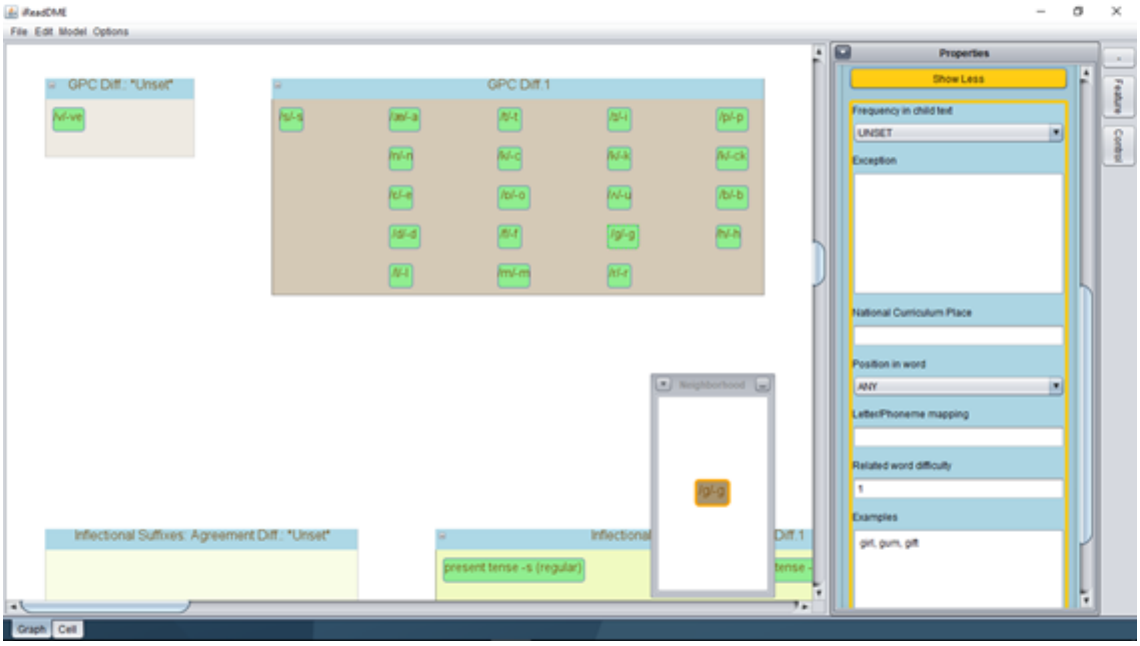


Figure 8 Additional attributes for the selected language feature are shown on the "feature-properties-view".

As already mentioned, the editor allows the user to focus on a subgraph of the entire domain model. This can be accomplished by using the "graph view" of the control panel to the right. When the user clicks the "Edit subgraph" button the view changes; see Figure 9.

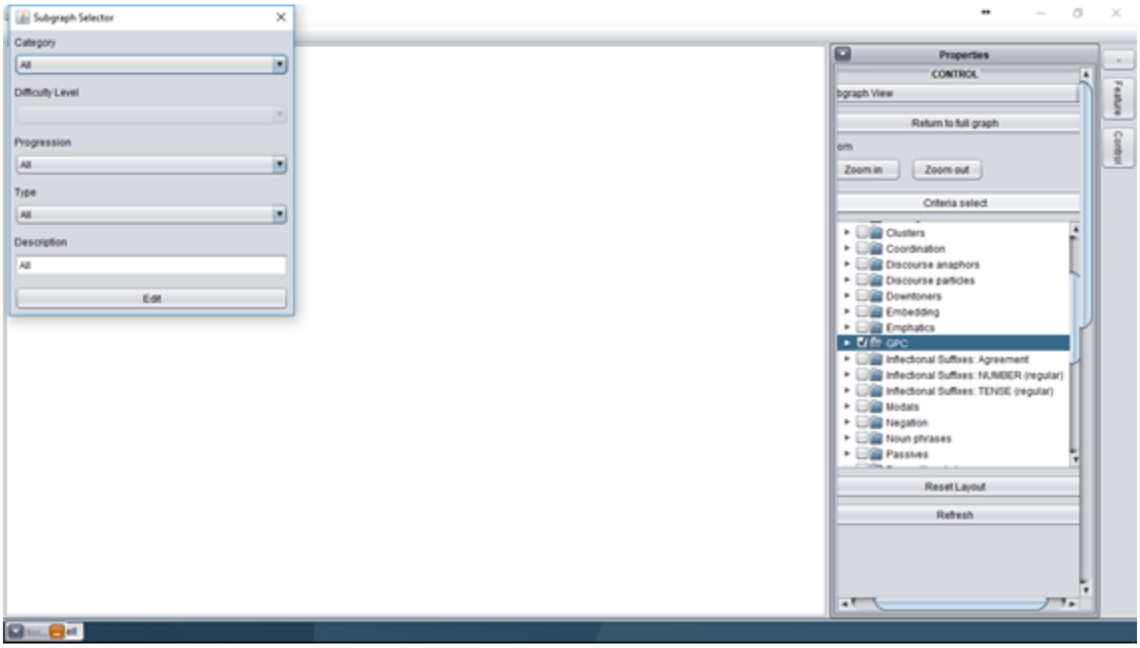


Figure 9 The user can either select features using the check-list on the "graph-view" to the right, or by using the popup window for additional filtering.

Now a check-list appears in the "graph-view" to the right where the user can select either entire language categories or individual language features. Another option for selecting language features is provided by the "critieria select" button. A popup window appers, which allows the user to filter the language features using more criteria.

Once language features are selected the user can click the "Refresh" button, so as to view the subgraph created by the selected vertices. In Figure 10, we can see the subgraph where only language features of the GPC category are selected. The subgraph view is less restrictive than the original view of the entire graph. Here the user can freely move features within their difficulty-vertex, move or resize difficulty-vertices. Edges are added the same way as when the entire graph is shown. Here difficulty-vertices are only shown if at least one of their language features belongs to the subgraph. Furthermore, if a difficulty-vertex has language features that do not belong to the selected subgraph, their border is highlighted in red (see Figure 10). This highlight is an implicit warning for the user in the case she adds an edge with source (or target) such a difficulty-vertex, since such an edge actually has an impact on all the language features of the difficulty-vertex even if they do not belong to the subgraph.

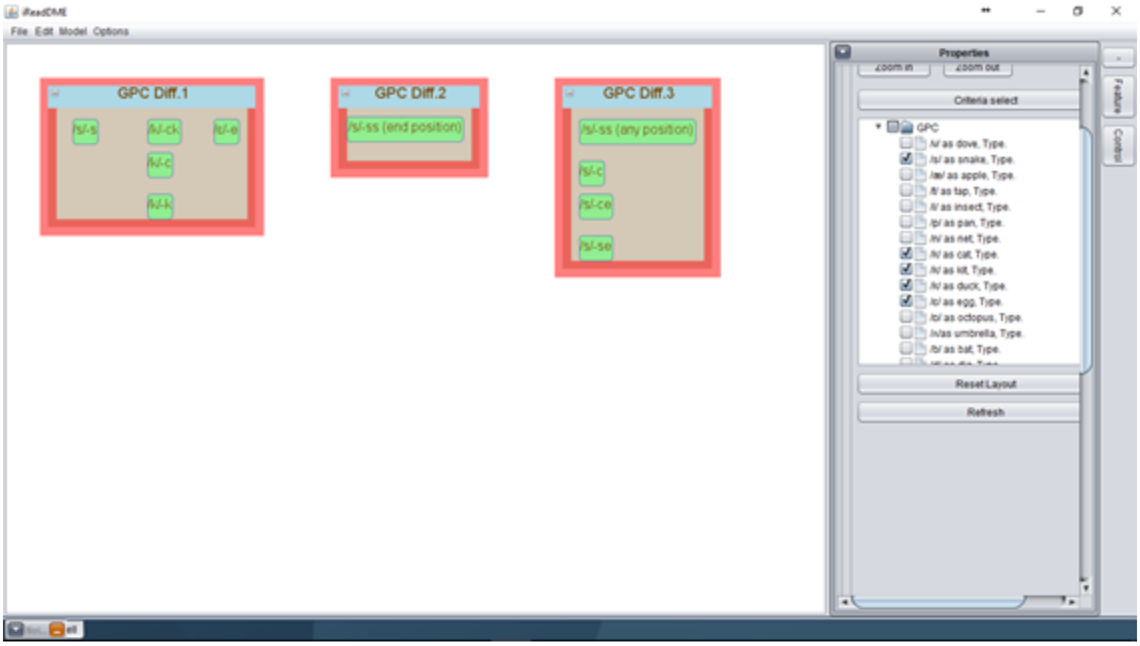


Figure 10 Subgraph of the selected features.

In Figure 11, we can see the selected subgraph and the edges denoting the prerequisites described in the English domain model for these language features. Whenever the user selects a language feature, its “neighborhood graph” pops-up, showing its incoming and outgoing edges. The color of the neighboring vertices is color used for the difficulty-vertices of their category.

In the case where a directed cycle is detected, a warning message appears as in Figure 12, and the action that triggered it is automatically undone.

The editor supports an undo-redo manager for the user’s actions that can be used with the standard ctrl+z/ctrl+y shortcuts or via the “edit” menu. The “model” menu supports editing of the domain model. Here the user can add, edit or delete linguistic levels or categories (refer to Figure 14). Such actions reset the undo-redo manager. The last menu, the “options” menu, allows the user to change most of the layout properties of the graphs, such as the distance between two language features of the same difficulty-vertex, etc.

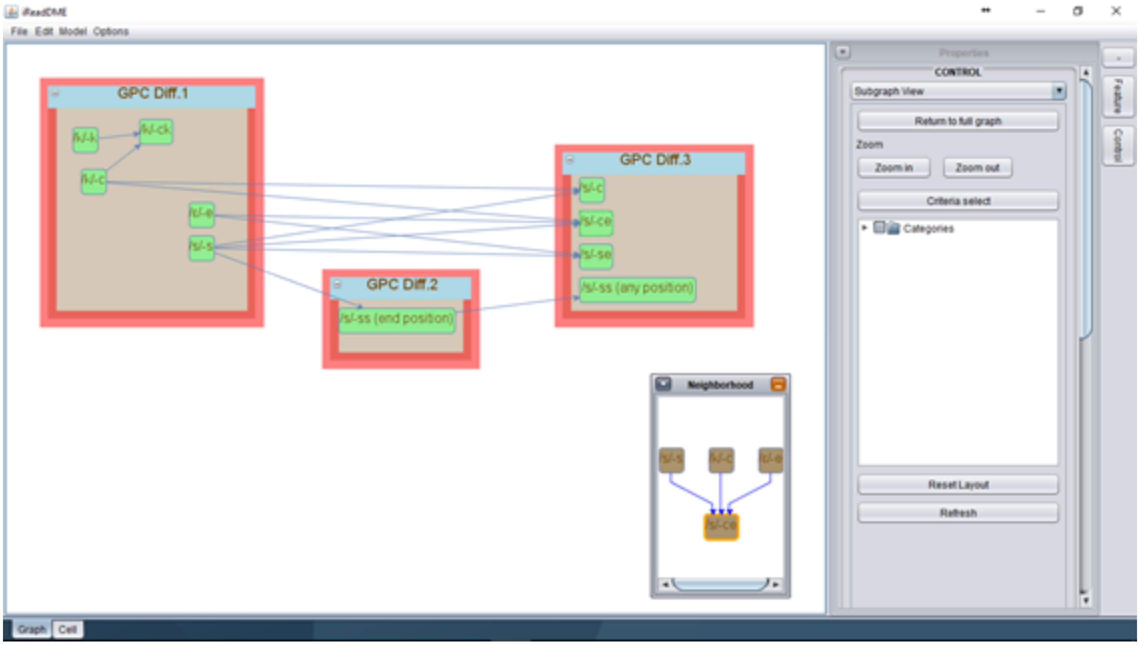


Figure 11 Subgraph with prerequisite edges. The neighborhood subgraph shows incoming and outgoing edges of the selected vertex.

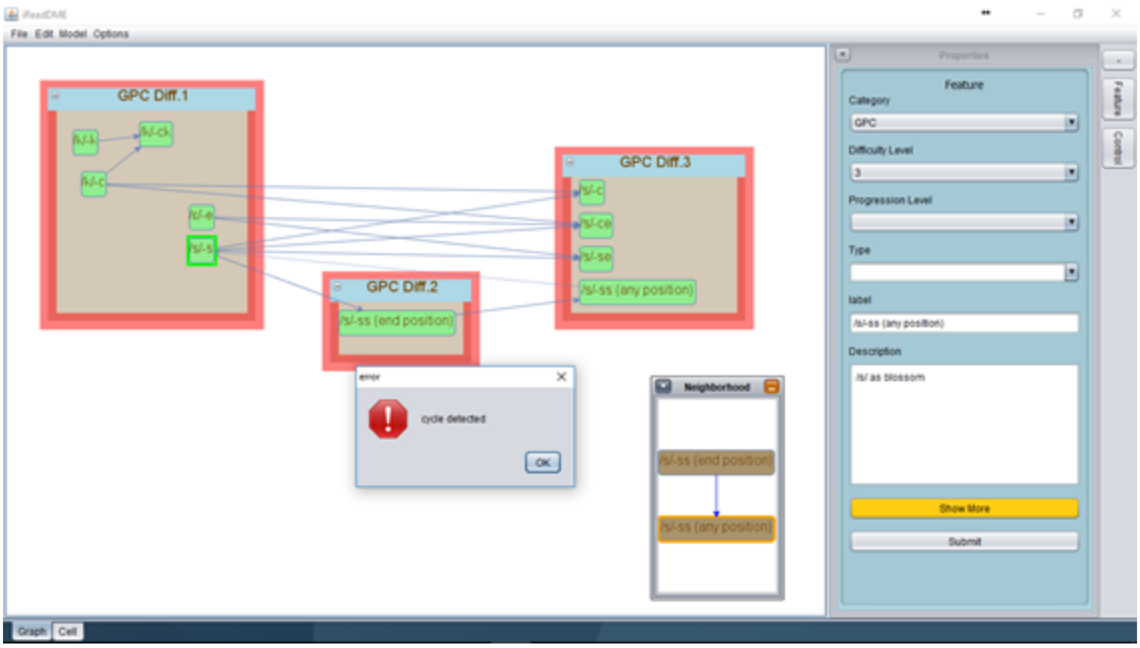


Figure 12 Directed cycles are not allowed.

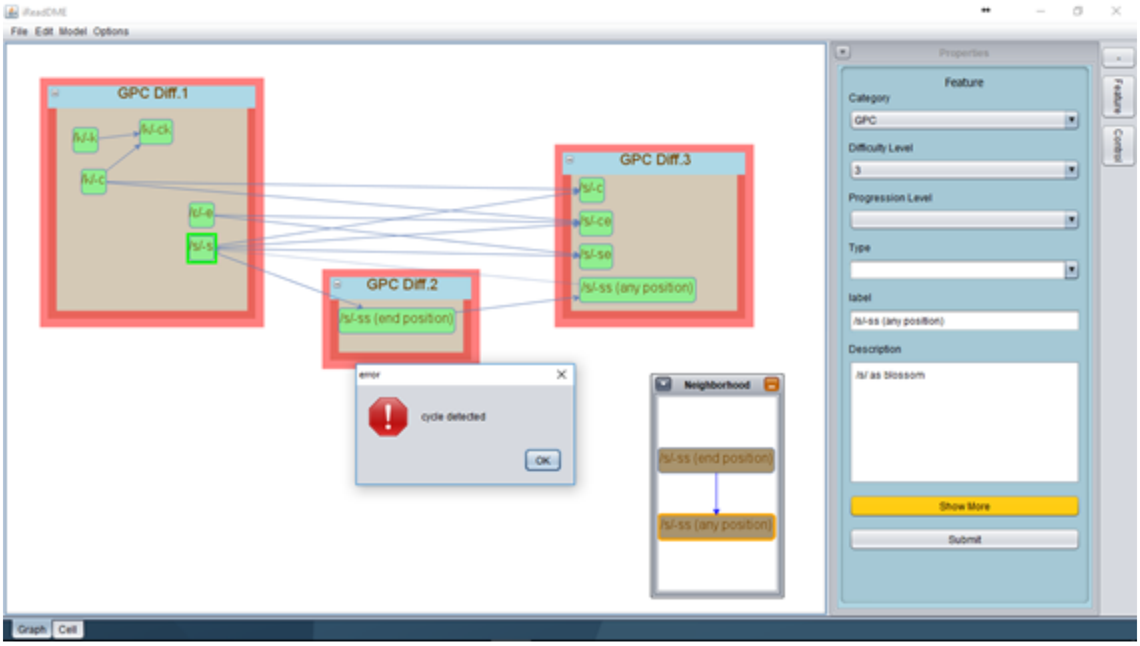
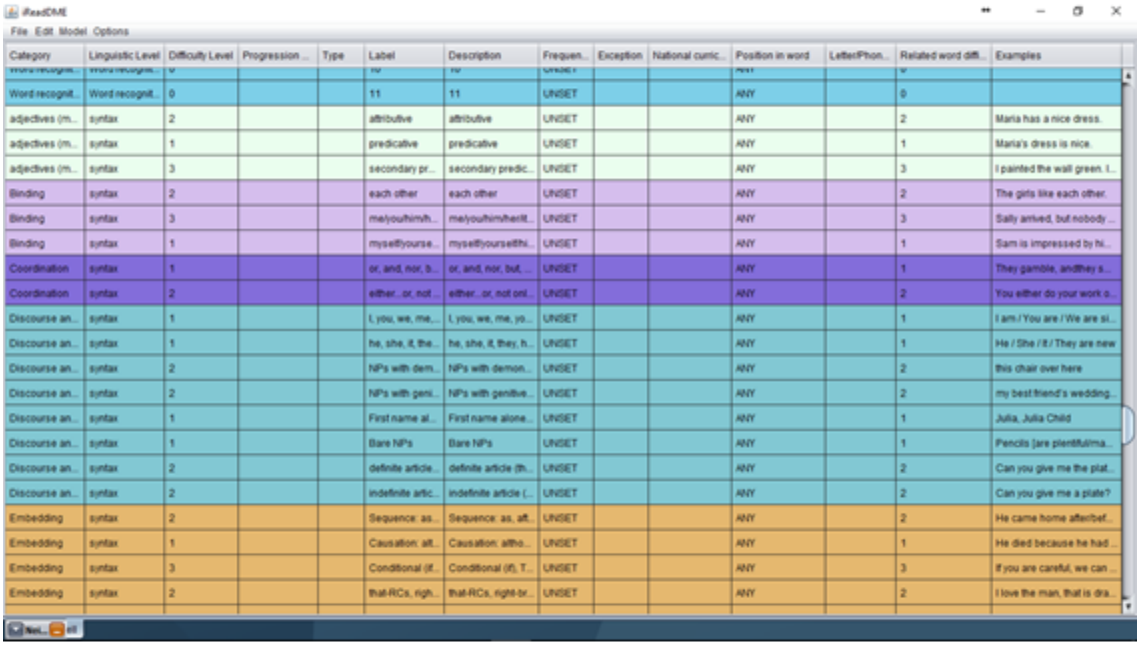


Figure 13 Add, edit or delete a linguistic level.

Finally, a table view is also offered, where the attributes of the language features can be viewed, as in Figure 14. Language features are grouped by category and highlighted with their category's color. The current version does not allow the user to change the table entries, since, changing the difficulty of a language feature for example, could result in changes of prerequisites that the user cannot visually detect.



Category	Linguistic Level	Difficulty Level	Progression …	Type	Label	Description	Frequen…	Exception	National curric…	Position in word	Letter/Phon…	Related word diff…	Examples
Word recognit…	Word recognit…	0			11	11	UNSET			ANY		0	
adjectives (m…	syntax	2			attributive	attributive	UNSET			ANY		2	Maria has a nice dress.
adjectives (m…	syntax	1			predicative	predicative	UNSET			ANY		1	Maria's dress is nice.
adjectives (m…	syntax	3			secondary pr…	secondary predic…	UNSET			ANY		3	I painted the wall green. I…
Binding	syntax	2			each other	each other	UNSET			ANY		2	The girls like each other.
Binding	syntax	3			me/you/him/h…	me/you/him/her/it…	UNSET			ANY		3	Sally arrived, but nobody …
Binding	syntax	1			myself/yourse…	myself/yourself/hi…	UNSET			ANY		1	Sam is impressed by hi…
Coordination	syntax	1			or, and, nor, b…	or, and, nor, but, …	UNSET			ANY		1	They gamble, andthey s…
Coordination	syntax	2			either…or, not …	either…or, not onl…	UNSET			ANY		2	You either do your work o…
Discourse an…	syntax	1			I, you, we, me,…	I, you, we, me, yo…	UNSET			ANY		1	I am / You are / We are si…
Discourse an…	syntax	1			he, she, it, the…	he, she, it, they, h…	UNSET			ANY		1	He / She / It / They are new
Discourse an…	syntax	2			NPs with dem…	NPs with demon…	UNSET			ANY		2	this chair over here
Discourse an…	syntax	2			NPs with geni…	NPs with genitive…	UNSET			ANY		2	my best friend's wedding…
Discourse an…	syntax	1			First name al…	First name alone…	UNSET			ANY		1	Julia, Julia Child
Discourse an…	syntax	1			Bare NPs	Bare NPs	UNSET			ANY		1	Pencils [are plentiful/ma…
Discourse an…	syntax	2			definite article…	definite article (th…	UNSET			ANY		2	Can you give me the plat…
Discourse an…	syntax	2			indefinite artic…	indefinite article (…	UNSET			ANY		2	Can you give me a plate?
Embedding	syntax	2			Sequence: as…	Sequence: as, aft…	UNSET			ANY		2	He came home after/bef…
Embedding	syntax	1			Causation: alt…	Causation: altho…	UNSET			ANY		1	He died because he had …
Embedding	syntax	3			Conditional (if…	Conditional (if), T…	UNSET			ANY		3	If you are careful, we can …
Embedding	syntax	2			that-RCs, righ…	that-RCs, right-br…	UNSET			ANY		2	I love the man, that is dra…

Figure 14 Table view of the language features.

6.4 Tool for Generating Grapheme-Phoneme-Correspondence

English, as a language, is considered to be a rather "difficulty" case, especially when referring to rules that map sounds to letters. In contrast, Greek is a more transparent language, since the pronunciation of words is straightforward. In other words, the rules for mapping graphemes to phonemes can be easily specified. However this is not the case with English words, since the same combination of letters (graphemes) can be mapped to several phonemes, there exist split digraphs and silent letters.

In the English domain model there exist several language features (approximately 120) that concern grapheme-phoneme correspondence (GPC). In order to check whether a word has a GPC language feature, the GPC of a word needs to be computed. Since no such tool is available, the challenge remains to derive the GPC for a word based on the graphemes and its phonetic transcription (in the International Phonetic Alphabet[24], IPA for short).

We focused on two goals: (i) given a set of GPC-rules, derive the GPC of a word, and, (ii) create a tool for the linguistics that will allow them to make decisions on the GPC-rules that will be used.

[24] https://en.wikipedia.org/wiki/International_Phonetic_Alphabet

Generating Grapheme Phoneme Correspondence

Our starting point is a set of GPC-rules, i.e. a mapping of phonemes to graphemes, where a phoneme can be mapped to several graphemes. For example the phoneme /eɪ/ has several mappings, such as 'a' (in the word ancient - /eɪnʃənt/) or 'a_e' (in the word bake-/beɪk/). An algorithm was created in order to derive all possible GPCs for a word based on these rules. The procedure "reads" the phonetic transcription of the word from left to right, and step by step. At each step, the algorithm reads a phoneme and checks which of its mapped graphemes can be used so that the word created by the concatenation of the graphemes matches the given word. At the end, a list of accepted GPCs is produced; recall that it can be the case that no GPC is produced for a word, or that multiple GPCs exist. The set of partial GPCs that were rejected at any step of the algorithm are also part of the output.

The following filtering is done on the initial output of the algorithm:

- If a word contains a double consonant (e.g. 'bb') or a double 'ee', then the produced GPC should not split them into two pairs of grapheme-phoneme. For example, for the word steel-/stiːl/, both GPCs /s/-s, /t/-t, /iː/-ee, /l/-l and /s/-s, /t/-te, /iː/-e, /l/-l are part of the initial output. The second one is excluded by the aforementioned rule, since it split the double 'ee' to two pairs. The rule is ignored only in the case where all produced GPCs fail; for example the word accent-/æksənt/, a unique GPC was produced /æ/-a, /k/-c, /s/-c, /ə/-e, /n/-n, /t/-t, where the double 'cc' is split into two pairs.
- Any GPC that contains a silent letter is removed, unless all produced GPCs contain silent letters. In this case, only those with the fewer silent letters are kept. As an example, silent letters can be letters 'e', 'h', 'l', 'r' or 'w'. Word above-/əbʌv/ has the following two initially computed GPCs: /ə/-a, /b/-b, /ʌ/-o, /v/-ve and /ə/-a, /b/-b, /ʌ/-o, /v/-v, //-e. The second one contains 'e' as a silent letter and is therefore rejected.
- Several times it occurs that a combination of graphemes can be expressed with different pairs. For example, in the word abbreviation-/əbriːvɪeɪʃən/, the combination 'tion' can be expressed as one pair: /ʃən/-tion, or as three pairs: /ʃ/-ti, /ə/-o, /n/-n. In such cases we keep the expression with the fewer pairs and reject the other.
- Finally, some GPC-rules are used only if they appear at the end of a word. In particular, these rules are: /m/ -'me', /n/-'ne', /t/-'te', /dʒ/-'dge', /t/-'tte', /d/-'de', /v/-'ve', /l/-'lle' and /ɪ/-'ie'. For example the word called-/kɔːld has two GPCs: /k/-c, /ɔː/-a, /l/-lle, /d/-d and /k/-c, /ɔː/-a, /l/-ll, /d/-ed. The first one is excluded since the pair /l/-lle does not appear at the end of the produced GPC. This rule is not applied if all produced GPCs violate it. For example, the word somehow/sʌmhaʊ/ has only one produced GPC: /s/-s, /ʌ/-o, /m/-me, /h/-h, /aʊ/-ow and it is therefore kept.

Tool for Grapheme-Phoneme Correspondence

A desktop application was developed in java in order to facilitate the linguist experts with the task of defining the GPC-rules that will be used by the algorithm previously described. By default the application loads a list of about 10.000 words and a list of GPC-rules. The words were selected by the linguist experts so that they are child-appropriate and will be the words that a user will encounter when interacting with the iRead literacy games.

In the main window there are three basic views as tabs:

dictionary view: The user can see the list of loaded words (to the left; see Figure 15) with their phonetic transcription. When selecting on a word, the information about the computed GPCs of the word appears to the right. The list of accepted GPCs (if any) and the list of rejected or partial GPCs (i.e. those that were either rejected during the filtering step, or that could not generate the entire word) is shown. This allows the user to practically look at all the computational steps. If for example, there is no accepted GPC computed, it is possible to understand where the computation halted and possibly add an additional GPC-rule. It occurred for several words, that some of the produced accepted GPCs contained split digraphs, while others did not. A toggle button (bottom right) rejects those GPCs that do not contain split digraphs, only for those words that have at least an accepted GPC with split digraphs.

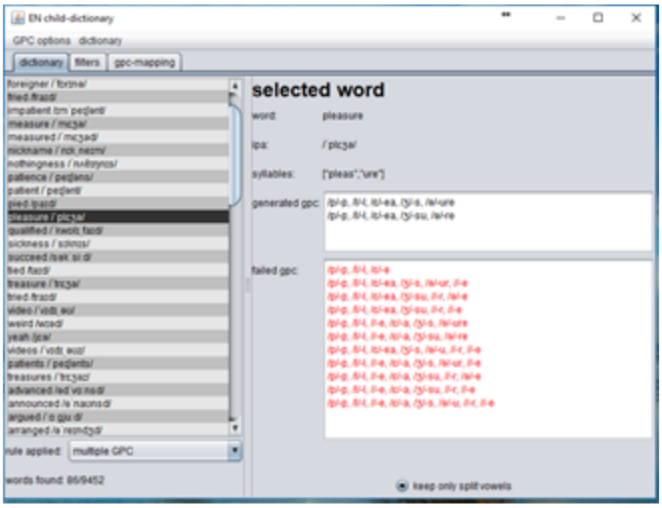


Figure 15 Selecting a word shows the valid and the rejected computed GPCs

gpc-mapping view: Offers a view of the current GPC-rules; see Figure 16. The user can click on an existing GPC-rule in order to edit or delete it. Also, she can add new rules with the "Add" button. The popup dialog shown in Figure 17 is used to enter the grapheme and the phoneme for the GPC-rule. Since phonemes use special characters, an embedded keyboard with these characters can be used. Whenever there is a change in the GPC-rules, all words are again processed for computing their GPCs according to the new GPC-rules.

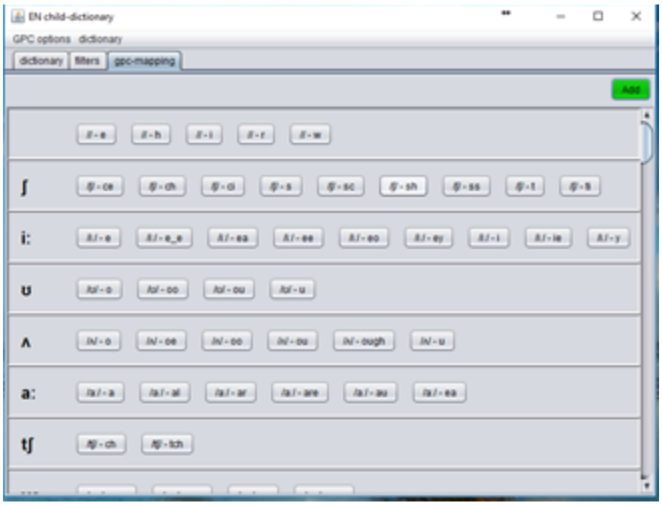


Figure 16 the gpc-mapping view

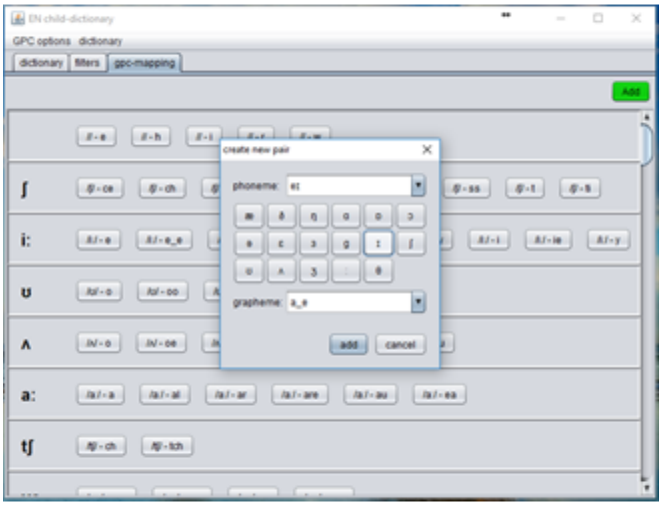


Figure 17 adding a new GPC-rule. An embedded keyboard can be used for the special phonetic characters

filters' view: Here the user can create filters for the list of words. Since the list of words can be arbitrarily large, it is hard to locate those words that are of interest (e.g. words that have no accepted GPC, words that contain silent letters, etc.). The list of available filters is shown to the left, and information about a selected filter is displayed to the right, as shown in Figure 18.

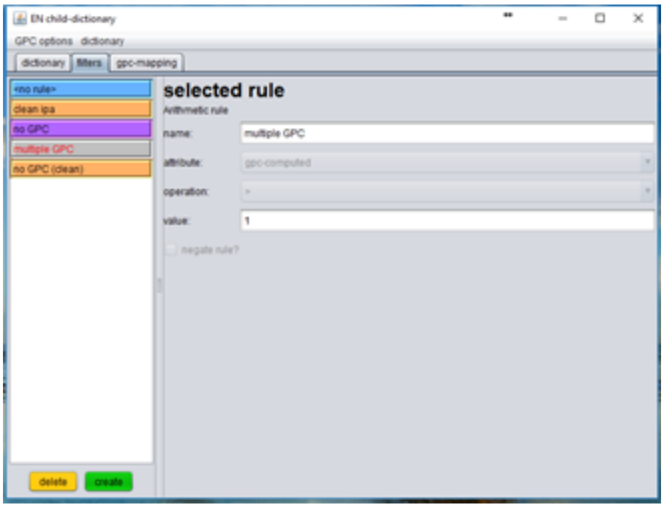


Figure 18 the filters' view

There are four types of filters that can be defined. All of them must have a unique name as identifier. Also, their display-color depends on their type.

- *Arithmetic filters:* these filters apply to the number of characters of a word, the number of characters of its phonetic transcription, the number of syllables or the number of accepted GPCs. Hence, the user can define filters such as: "number of syllables greater than 3", "number of accepted GPCs equal to 1", etc. Each filter can also be negated, translating into "number of syllables at most 3", or "number of accepted GPCs not equal to 1" respectively. For example in Figure 19, a filter for words that have only one accepted GPC is shown.

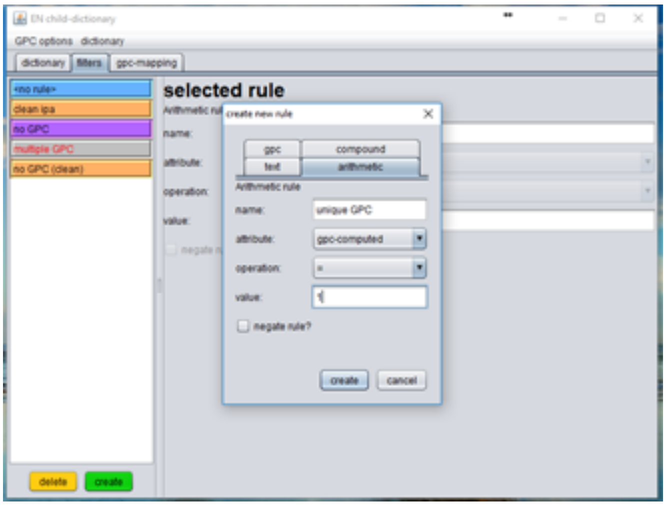


Figure 19 creating arithmetic filter "words whose accepted GPCs is equal to 1"

- *gpc filters:* the user can create filters based on specific GPC-rules. Available filters are "words whose accepted GPCs contain a particular GPC-rule", where the GPC-rule is selected from a dropdown list of all the defined GPC-rules. Again the filter can be negated, indicating that only those words that do not contain the selected GPC-rule in any of their accepted GPCs will be selected. In Figure 20 a filter for words whose accepted GPC contains a silent 'e' is created.

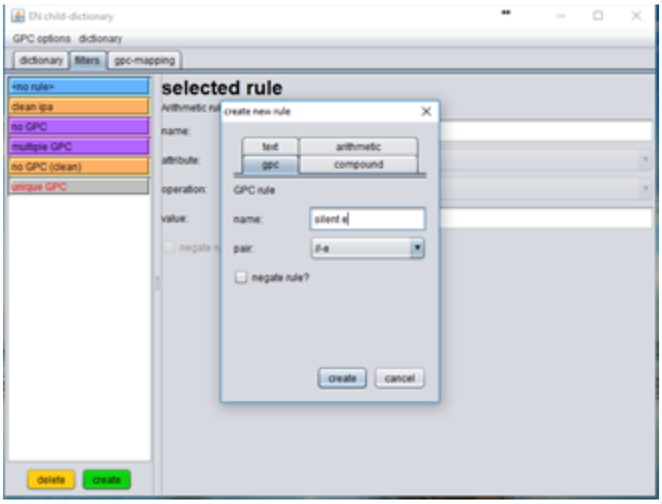


Figure 20 creating gpc filter "words with an accepted GPC that contains the GPC-rule //-e (i.e. silent e)"

- *compound filters:* Here the user can combine two already defined filters using the logical operands OR and AND. For example, in Figure 21, a filter for selected those words that have one accepted GPC and this GPC contains a silent 'e' is created.

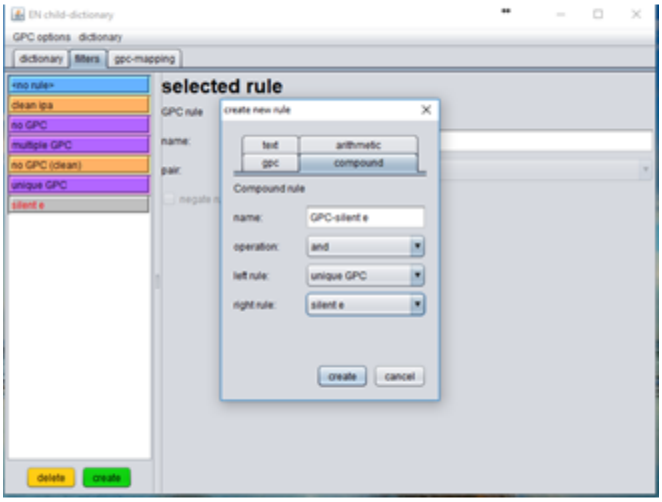


Figure 21 creating compound filter "unique GPC AND silent e"

- *text filters:* The user can search for a substring either at the beginning, the middle, the end or anywhere in the word or its phonetic transcription. In Figure 22, a filter is created for words that end with ‘ee’.

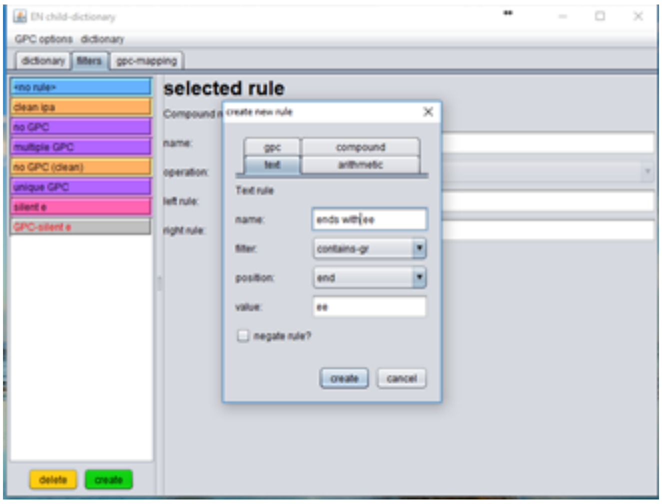


Figure 22 creating text filter "words that end with 'ee'"

After defining filters the user can apply them to the loaded list of words from the dictionary view (see Figure 23). The number of words filtered over the total number of words is also shown.

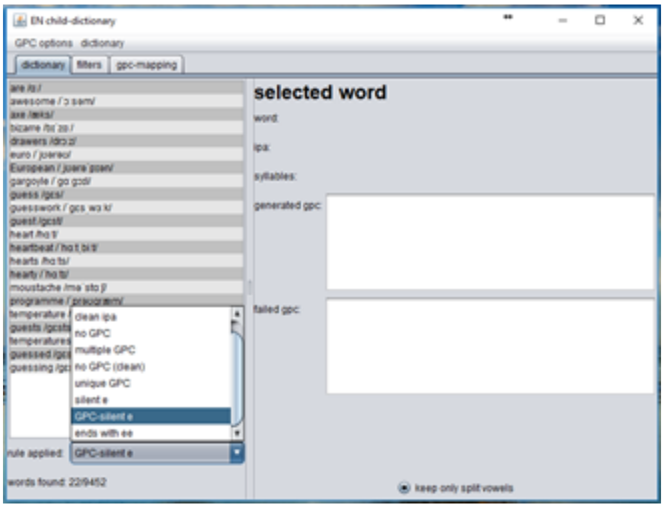


Figure 23 applying a filter to the list of words

The main window offers two menus: “GPC options” and “dictionary”. The “GPC options” allows the user to save the current GPC-rules to a local file, load GPC-rules from a local file or load the default built-in list of GPC-rules. In the same way the user can save the list of filtered words (in csv-format), load words from a local file (also in csv-format) or load the default built-in list of words, by using the “dictionary” menu options.

6.5 Child appropriate dictionaries for Greek and English

In a high-level of description, the iRead system is designed for supporting novice readers in four basic language, including English and Greek. The targeted age-groups are mainly children aged between 6 and 12 . This implies that the content delivered to the children in the iRead literacy games (and the iRead Reader) should be appropriate for their age. Furthermore, one of the attributes defined for the language features in the domain models is the “frequency within child texts”. This gives rise to two needs; a corpus of child texts and appropriate content for the literacy games. In the following we give a brief description of the related resources and the required processing. Full details will be presented in **Deliverable 8.3: “The iRead Integrated System”**.

Corpus of child texts: Specific resources were indicated by the linguist experts, and were processed in order to have a unified format and become available for further processing and statistical analysis.

Child dictionaries: In the domain models there exist two main categories of language features: those that are related to single words, and those related to sentences. The iRead literacy games should draw their content from an approved pool of words and sentences, in order to assure that children using the iRead system encounter appropriate content. Two child-dictionaries for words are created, one for English and one for Greek, while the collection of sentences will be hand-made by the linguist experts. The child-dictionaries contain information that is relevant to the language-features defined in the domain models. In particular for every word we store the following information:

word	the word of the dictionary
phonetic	the phonetic transcription of the word
part of Speech	the part of speech: verb, noun, adjective, adverb, etc.
grapheme-phoneme	the grapheme phoneme correspondence
stem	the word's stem, for example the stem for 'advancing' is 'advance'
syllables	the syllabification of the word
number of syllables	the number of syllables
word type info	information related to the tense and number of the word if it is a verb, or the number if it is a noun, or the degree if it is an adjective or adverb
CV form	the letter-based CV form of the word
number of characters	the number of characters of the word
number of phonemes	the number of pairs in the grapheme phoneme correspondence
number of morphemes	the number of morphemes of the word

suffix	the suffix of the word
prefix	the prefix of the word
suffix type	the type of suffix: add, drop, double, irregular
prefix type	the type of prefix: add, visual, irregular
irregular word (no GPC)	marks if the word is irregular
word frequency in child language	the frequency of the word in the corpus of child-texts
synonyms	a number of words with same or very similar meaning
antonyms	a number of words with opposite meanings
meaning (definition)	
feature_info	Information about the language features the word contains and the position they appear

English child-dictionary: The corpus of child texts was analysed and a list of approximately 10.000 words were selected for the child-dictionary. The phonetic transcription was taken from the English dictionary of previous project iLearnRW. All other information has been generated algorithmically. In particular for the language features of the word, for each language feature defined in the domain model, we mapped a function-routine that checks whether a word contains the specific feature and the positions within the word that it appears.

Greek child dictionary: The child-dictionary for Greek has been almost entirely based on the dictionary developed for iLearnRW, and the word-entries it contains are the same words used by the iLearnRW-games.

7 Conclusions

As shown, a wide range of linguistic instances are included within the developed models across languages, going beyond the skills of decoding of sound-letter correspondence and word recognition to more complex reading skills, including morphological, syntactic and discourse processing of text. Although different methods were used for the selection of the language categories to be included (literature reviews, frequency counts, and empirical data), the two models do not differ significantly across languages. Importantly, ratings of difficulty, indicating each feature's difficulty or complexity relative to the rest of the features within the same category, as well as progression scales, indicating the order of acquisition of those features are included within models. Complexity or difficulty of linguistic features not only predicts difficulties in poor readers/readers with dyslexia or EFL readers, but also in novice readers who also have to deal with more complex constructions. To reliably define which linguistic features are easier or harder, previous findings were drawn from the relevant literature in order to formulate an initial taxonomy of linguistic features per category. Difficulty was also tested experimentally, by collecting empirical data from children with dyslexia and young typically developing children. By conducting empirical work, we were able to record and analyze actual reading data and define aspects of language that are more difficult across readers. At this stage, we used our pilot preliminary data to inform the models. However, the number of participants in the pilot study described here is not sufficient to reliably indicate whether our taxonomy of difficulty is totally accurate. Therefore, empirical data collection will continue in larger groups of participants in order help us reliably indicate which constructions are harder than other and revise the difficulty taxonomy within models at later stages.

REFERENCES

Abramovici, S. (1983). Errors in proofreading: Evidence for syntactic control of letter processing? *Memory & Cognition*, 11, pp. 258-261.

Aidinis, A. 1998. Phonemes, morphemes and literacy development: Evidence from Greek. At. http://eprints.ioe.ac.uk/18894/7/313848_Redacted.pdf. 17.06.2017.

Anastasíou, D. (1998). *Dyslexia. Theoria kai erevna: opseis Praktikis*. [Dyslexia. Theory and Research: Aspects of Practice] Volume 1. Athens: Atrapos.

Arosio, F., Pagliarini, E., Perugini, M., Barbieri, L. & Guasti M. T. (2016). Morphosyntax and logical abilities in Italian poor readers: the problem of SLI misidentification. *First Languag*e. DOI: 10.1177/0142723716639501.

Avrutin, S. (2000). Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. *Language and the Brain: Representation and Processing*. San Diego: Academic Press, pp. 295-313.

Avrutin, S., Lubarsky, S. & Greene, J. (1999), Comprehension of contrastive stress by agrammatic Broca's aphasics. *Brain and Language,* 70, pp. 163-186.

Bar-Shalom, E., Crain, S. & Shankweiler, D. (1993). A comparison of comprehension and production abilities of good and poor readers. *Applied Psycholinguistics*, 14, pp. 197–227.

Bhattacharya, A. & Ehri, L. C. (2004). Graphosyllabic analysis helps struggling adolescent readers read and spell words. *Journal of Learning Disabilities*. Volume 37, pp. 331–348.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? *Linguistic perspectives on second language acquisition.* New York: Cambridge University Press, pp. 41-68

Bley-Vroman, R. (2009). The Evolving Context of the Fundamental Difference Hypothesis. *Studies in Second Language Acquisition* , 31 ,pp. 175– 198.

Bishop, D. V. M. (1997b). Un*common understanding: Development and disorders of language comprehension in children.* Hove, East Sussex, UK: Psychology Press.

Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, *20*(2), 139-173.

Bright Solutions for Dyslexia. (2010). What is dyslexia? Retrieved December 10, 2017 from: http://www.dys-add.com/define.html.

Brown, G.D.A. & Deavers, R.P. (1999). Units of analysis in nonwords reading: Evidence from children and adults. *Journal of Experimental Child Psychology 73*, 208-242.
Byrnes, B. (1981). Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics*, 2, pp. 201–212.

Cantiani, C., Lorusso, M.L., Guasti, M.T., Sabisch, B. & Männel, C. (2013). Characterizing the morphosyntactic processing deficit and its relationship to phonology in developmental dyslexia? *Neuropsychologia* 51 (8), pp. 1595–1607.

Carlisle, J. F. (1987). The use of morphological knowledge in spelling derived forms by learning-disabled and normal students. *Annals of Dyslexia* 37, pp. 90–108.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed), *Morphological aspects of language processing*. Hillsdale, NJ: Erlbaum, pp. 189–209.

Carlisle, J.F. (2003). Morphological matters in learning to read: a commentary. *Reading Psychology*, 24, pp. 291-322.

Casalis**,** S., Leuwers, C. & Hilton, H. (2013). Syntactic comprehension in reading and listening: a study in French dyslexic children. *Journal of Learning Disabilities, 46,* pp. 201-219.

Chien, Y.-C. & Wexler, K. (1990). Children s knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics *.* *Language Acquisition* 1, pp. 225-295.

Chliounaki, K. & Bryant, P. (2003). Choosing the right spelling in Greek: morphology helps. In *Revue française de linguistique appliquée* (Vol. VIII), pp. 35 - 45.

Correa, L.M.S. (1995). An alternative assessment of children's comprehension of relative clauses. *Journal of Psycholinguistic Research* 24 (3), pp. 183–203.

Cole, P.M., Tan, P.Z., Hall, S.E., Zhang, Y., Crnic, K.A., Blair, C.B., & Li, R. (2011). Developmental changes in anger expression and attention focus: Learning to wait. *Developmental Psychology* 47:1078–1089. doi: 10.1037/a0023813.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.

Crain, S., Thornton, Th., Boster, C., Conway, L., Lillo-Martin, D. & Woodams, W. (1996). Quantification without qualification. *Language Acquisition,* 5, pp. 83-153.

Creaghead, N. A. & Donnelly, K. G. (1982). Comprehension of superordinate and subordinate information by good and poor readers. *Language, Speech, and Hearing Services in the Schools,* 13, pp. 177-186.

de Vincenzi, M., Arduino, L. S., Ciccarelli, L. & Job, R. (1999). Parsing strategies in children comprehension of interrogative sentences. *Proceedings of European Conference on Cognitive Science*. Rome: Istituto di Psicologia del CNR, pp. 301–308.

Ehri, L. (2004). Teaching phonemic awareness and phonics. *The voice of evidence*. Baltimore, NY: Paul Brookes, pp. 153-186.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In: P. Gough, L. Ehri & R. Treiman (Eds.), *Reading Acquisition* (pp. 107-143). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Ehrlich, M.-F., Rémond, M. & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: differences in metacognitive monitoring. *Reading and Writing* 11(1), pp. 29–63. http://dx.doi.org/10.1023/A:1007996502372.

Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5), pp. 491–520.

Fiorin, G. (2010). *Meaning and dyslexia: a study on pronouns, aspect and quantification*, Ph.D. dissertation, Utrecht Institute of Linguistics, OTS. LOT series.

Fox, D. & Grodzinsky, Y. (1998). Children's passive: A view from the by-phrase. *Linguistic Inquiry,* 29, pp.311-332.

Friedmann, N. (2002). Syntactic tree pruning and question production in agrammatism. *Brain and Language*, 83, pp. 117-120.

Friedmann, N., Belletti, A. & Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119, pp. 67–88.

Friedmann, N. & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, 31, pp. 661–681.

Garraffa, M. (2007). *Impoverishment of grammatical features in a non-fluent aphasic speaker: The grammatical nature of minimal structures*. Ph.D. dissertation. University of Siena.

Garraffa, M. & Grillo, N. (2008). Canonicity effects as a grammatical phenomenon. *Journal of Neurolinguistics,* 21, 177-197.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, pp. 1–76.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Grodzinsky, Y. & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, 24, pp. 69-102.

Grodzinsky, Y., Wexler, K., Chien, Y., Marakovitz, S. & Solomon, J. (1993). The breakdown of binding relations. *Brain and Language*, 45, pp. 396-422.

Guasti, M.T., Branchini, C., Arosio, F. & Vernice M. (2012) A developmental study of subject and object relatives in Italian. *Revue Roumaine de Linguistique,* 72, 105-116.

Harley, T. (2001). *The psychology of language: From data to theory*. Psychology Press.

Harm, M.W. & Seidenberg, M.S. (2001). Are there orthographic impairments in phonological dyslexia? *Cognitive Neuropsychology 18,* 71-92.

Hatzigeorgiou, N., Gavrilidou, M., Piperidis, S., Carayannis, G., Papakostopoulou, A., Spiliotopoulou, A. et al. (2000). Design and implementation of the online ILSP corpus. *Proceedings of the Second International Conference of Language Resources and Evaluation (LREC)* 3, pp. 1737-1740.

Hickok, G. & Avrutin, S. (1996). Comprehension of wh-questions in two Broca's aphasics. *Brain and Language,* 52, pp. 314-327.

Hirsch, Ch. & Wexler, K. (2006). Acquiring verbal passives: Evidence for a maturational account. *Annual Meeting of the Linguistics Society of America*, Albuquerque, NM.

Holton, D., Mackridge P. & Philippaki-Warburton, I. (1997*). Greek: A Comprehensive Grammar of the Modern Language*. London: Routledge.

ILEARNRW: "Integrated Intelligent Learning Environment for Reading and Writing", 2012-2015, Funding FP7 / Contract no.: 318803, http://www.ilearnrw.eu/.

Jegerski, J. (2014). Self-paced reading. In *Research Methods in Second Language Psycholinguistics*. New York, NY: Routledge, pp. 20–49.

Joanisse, M. F., Manis, F. R., Keating, P. & Seidenberg, M. S. (2000). Language deficits in dyslexic children: speech perception, phonology, and morphology. *Journal of Experimental Child Psychology,* 77(1), pp. 30-60.

Joanisse, M. F. & Seidenberg, M. S. (1998). Specific language impairment: A deficit in language or processing? *Trends in Cognitive Sciences,* 2, pp. 240-247.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review,* 98, pp. 122–149.

Kamhi, A. G. & Catts, H. W. (1986). Toward an understanding of developmental language and reading disorders. *Journal of Speech and Hearing Disorders,* 51(4), pp. 337-347.

Karakirgiou, M. (2017). *Katagrafi mathisiakon dyskolion (dyslexia) kai tropos diexagogis tis exetasis* [learning difficulties (dyslexia) and methods of test]. Retrieved October 2, 2017 from http://docplayer.gr/38239695-Ypoyrgeio-politismoy-paideias-thriskeymaton-kentro-ellinikis-glossas.html.

Kourakis, I.E. (1997). A*nichnefsi ston kosmo ton mathisiakon dyskolion.* [Detecting the learning difficulties]. Athens: Ellin.

Ktori, M. & Pitchford, N. J. (2008). Effect of orthographic transparency on letter position encoding: A comparison of Greek and English monoscriptal and biscriptal readers*. Language & Cognitive Processes,* 23, pp. 258-281.

Leikin M. & Assayag-Bouskila, O. (2004). Expression of syntactic complexity in sentence comprehension: A comparison between dyslexic and regular readers*. Reading and Writing,* 17 (8), pp. 01–822.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T. K., Laakso, M.-L., Leinonen, S., Leppänen, P. H. T., Lyytinen, P., Poikkeus, A.-M., Puolakanaho, A., Richardson, U. & Viholainen, H. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology,* 20, pp. 535-554.

Macaruso, P., Bar-Shalom E., Crain S. & Shankweiler D. (1989). Comprehension of temporal terms by good and poor readers. *Language and Speech,* 32 (1), pp. 45–67.

Mann V. A., Shankweiler D. & Smith S. T. (1984). The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language,* 11, pp. 627–643.

Mann, V. & Singson, M. (2003). Linking morphological knowledge to English decoding ability: large effects of little suffixes. In *Reading Complex Words: Cross-Language Studies*. New York: Kluwer Academic/Plenum Publishers, pp. 1–25.

Marslen-Wilson, WD. (2001). Access to lexical representations: Cross-linguistic issues. *Language Cognitive Processes,* 16, pp. 699-708.

Morgenstern, A. & Sekali, M. (2009). What can child language tell us about prepositions? A contrastive corpus-based study of cognitive and social-pragmatic factors. *Studies in Language and Cognition*, Cambridge Scholars Publishing, pp. 261-275.

Morton, J. & Patterson, K.E. (1980). A new attempt at interpretation, or, an attempt at a new interpretation. *Deep dyslexia*. London: Routledge and Kegan Paul, pp. 91-118.

Nagy, W.E, & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly,* 19, pp.304-330.

Nagy, W. E., Garcia, G. E., Durgunoglu, A. & Hancin-Bhatt, B. (1993). Spanish–English bilingual children's use and recognition of cognates in English reading. *Journal of Reading Behavior,* 25, pp. 241–259.

Nation, K., Snowling, M. J. & Clarke, P. J. (2005). Production of the English past tense by children with language comprehension impairments, *Journal of Child Language,* 32, pp. 117-137.

Oakhill, J. & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech,* 29, pp. 25-37.

Oliveira, L. E. S., Moro, C. M. C., Souza, A. C., Nohama, P. & Cancian, P. S. (2012). Identificacao de continuidade de cuidado em sumarios de alta hospitalar. In XIII Congresso Brasileiro em Informatica em Saude. Retrieved May 2, 2005 from: http://docplayer.com.br/1917815-Identificacao-de-continuidade-de-cuidado-em-sumarios-de-alta-hospitalar-continuity-of-care-identification-in-discharge-summaries.html

Packenham, K. J. (1980). *The resolution of lexically-reinforced anaphoric reference and its relation to reading in English as a second language.* Ph.D. Dissertation, University of Pittsburgh, Russia.

Pandeliadou, S. (2011). *Mathisiakes dyskolies kai ekpaideftiki praxi, ti & giati.* [Learning difficulties and educational practice, what & why]. Athens: Pedio Publications.

Papadopoulou, M. & Zafiropoulou, M. (2005). Comparison of reading errors made by first grade Greek students and English-speaking students with reading problems. In: *Australian Journal of Learning Disabilities,* 10:1, pp. 25-33, DOI: 10.1080/19404150509546783

Philip, W. (1995). *Event quantification in the acquisition of universal quantification.* Ph.D. Dissertation, University of Massachusetts, Amherst.

Pavlidis, G.Th. & Giannouli, V. (2003). Spelling Errors Accurately Differentiate USA-English Speakers from Greek Dyslexics: Implications for Causality and Treatment. *Literacy Acquisition: The Role of Phonology, Morphology and Orthography*. Washington, IOS Press.

Philippaki-Warburton, I. (1973). Modern Greek verb conjugation: Inflectional morphology in a Transformational Grammar. *Lingua,* 32, pp. 193-226.

Protopapas, A., Simos, P. G., Sideridis, G. D. & Mouzaki, A. 2012. The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology,* 33, pp. 217–240.

Protopapas, A. (2014). Learning to read Greek. To appear in: *Reading acquisition across languages and writing systems: An international handbook.* Cambridge University Press. Retrieved August 2, 2017 from:
http://users.uoa.gr/~aprotopapas/CV/pdf/Protopapas_NIASChapterGreek_draft5.pdf

Reggiani, D. (2010). *Dyslexia and the acquisition of syntax: Passive and control*, Ph.D. Dissertation, Università degli Studi di Verona, Italy.

Rispens, J. & Been, P. (2007). Subject-verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): a closer look. *International Journal of Language and Communication Disorders,* 42(3), pp. 293-305.

Rispens, J., Roeleven, S. & Koster, C. (2004). Sensitivity to subject-verb agreement in spoken language in children with developmental dyslexia. *Journal of Neurolinguistics,* 17, pp. 333-347.

Ruigendijk, E., Kouwenberg, M. & Friedmann, N. (2004). Question production in Dutch agrammatism. *Brain and Language,* 91, pp. 116-117.

Sakellaríou, G. (2007). I axiologisi ton glossikon diskolion tou graptou logou se paidia kai efivous os meros mias dievrymenis diadikasias prosengisis tou thematos stin kliniki logopediki praktiki. [Assessment of the language difficulties in written speech in children and adolescents as part of an extended process of approaching the topic in clinical therapist practice] In: *Glossikes Dyskolíes kai Graptos logos [language difficulties and Written Language]*. Athens: ed. Grigori.

Scappini, M. (2015). *On the processing of sentential negation in dyslexic adults: An ERP study,* Ph.D. Dissertation, University of Verona, Italy.

Scarborough, H. (1990). Very early language deficits in dyslexic children. *Child Development,* 61, pp. 1728- 1743.

Scarborough, H. S. (1991). Early syntactic development of dyslexic children. *Annals of Dyslexia*, 41, pp. 207-220.

Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. In *Proceedings of International Conference on New Methods in Language Processing*, Manchester, UK.

Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J. M., Stuebing, K. K., Shaywitz, S. E. & Shaywitz, B. A. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology and syntax. *Psychological Science,* 6(3), pp. 149-156.

Smith S., Macaruso P., Shankweiler D. & Crain S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics,* 10, pp. 429–454.

Stein, C. L., Cairns, H. S. & Zurif, E. B. (1984). Sentence comprehension limitations related to syntactic deficits in reading-disabled children, *Applied Psycholinguistics,* 5, pp. 305–322.

Tutunjian, D. & Boland, J. E. (2008). Do we need a distinction between arguments and adjuncts? Evidence from psycholinguistic studies of comprehension. *Language and Linguistic Compass,* 2, pp. 631–646.

Tyack, D. & Ingram, D. (1977). Children production and comprehension of *wh*-questions. *Journal of child language,* 4, pp. 211-224.

Tyler, A. E. (1986). *Acquisition and use of English derivational morphology: An experimental investigatio*n. Ph.D. Dissertation, University of Iowa, Iowa City.

Vender, M. & Delfitto, D. (2010). Towards a pragmatic of negation: The interpretation of negative sentences in developmental dyslexia. *Generative Grammar at Geneva*, 6, pp. 1–27.

Vogel, S. A. (1977). Morphological ability in normal and dyslexic children. *Learning Disorders,* 10(1), pp. 41-49.

Voulgaris, D., Ikanotites fonologikis Epexergasias [Phonological Processing Capacities]. Retrieved October 2, 2017 from http://repository.edulll.gr/edulll/retrieve/3612/1058.pdf

Waltzman, D. & Cairns, H. (2000). Grammatical knowledge of third grade good and poor readers. *Applied Psycholinguistics,* 21, 263–284.

Whitehurst G. J. & Lonigan C. J. (1998). Child development and emergent literacy. *Child development,* 69, pp. 848 – 872 . Doi: 10.1111/j.1467–8624.1998.tb06247.x

Wiig, E., Semel, E. & Crouse, M. (1973). The use of morphology by high risk and learning disabled children*. Journal of Learning Disabilities,* 6, pp. 457-465.

Zachos, D.H. (1992). *Anagnosi-grafi. Psychoglossologiki prosengisi*.[Reading-writing. Psycholinguistic approach]. Athens: Kentro Psychologikon Meleton.

Zafiropoulou, M. et al. (2004). *Prosarmogi sto scholeio. Prolipsi kai antimetopisi dyskolion.* [Adaptation to school. Prevention and confrontation of learning disabilities]. Athens: Ellinika grammata.

APPENDICES

Appendix I. The English domain model for dyslexia

Linguistic Level	**Category**	**Feature**		**Examples**	**Difficulty level of feature**	**Prerequisites**
Phonology	GPC	/s/	s	sad, sit, sip	1	None
Phonology	GPC	/s/	ss (end position)	mess, pass, hiss	2	/s/-s
Phonology	GPC	/s/	ss (any position)	assistant, blossom, grasshopper, message	3	/s/-ss (end position)
Phonology	GPC	/s/	c	city, bicycle, pencil	3	/s/-s /k/-c
Phonology	GPC	/s/	ce	ice, race, palace	3	/s/-s /k/-c /ɛ/-e
Phonology	GPC	/s/	se	house, mouse, chase	3	/s/-s /ɛ/-e
Phonology	GPC	/æ/	a	add, ant, act	1	None

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/t/	t	tap, tin, tip	1	None
Phonology	GPC	/t/	tt	better, butter, otter	3	/t/-t
Phonology	GPC	/ɪ/	i	ink, pig, ring	1	None
Phonology	GPC	/p/	p	pan, pet, pop	1	None
Phonology	GPC	/p/	pp	puppet, apple, floppy	3	/p/-p
Phonology	GPC	/n/	n	net, nip, nap	1	None
Phonology	GPC	/n/	nn	bonnet, dinner, winner	3	/n/-n
Phonology	GPC	/n/	kn	knot, knee, knob	3	/k/-k /n/-n
Phonology	GPC	/ŋ/	ng	long, song, ring	2	/n/-n /g/-g
Phonology	GPC	/k/	c	cat, cap, can	1	/k/-k
Phonology	GPC	/k/	k	kit, kid, king	1	None
Phonology	GPC	/k/	ck	duck, back, clock, sock	1	/k/-k /k/-c
Phonology	GPC	/ɛ/	e	egg, end, empty, ten	1	None
Phonology	GPC	/ɛ/	ea	head, bread, feather	2	/ɛ/e /æ/a
Phonology	GPC	/ɒ/	o	odd, dog, octopus	1	None

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/ʌ/	u	up, under, bus	1	None
Phonology	GPC	/ʌ/	ou	touch, young, double	2	/ɒ/-o /ʌ/-u
Phonology	GPC	/eɪ/	ai	aim, wait, paint, train, rain	2	/æ/-a /ɪ/-i
Phonology	GPC	/eɪ/	ay	say, day, pay, play	3	/æ/-a /i/-y
Phonology	GPC	/eɪ/	a	table, alien, acorn	3	/æ/-a
Phonology	GPC	/eɪ/	a_e	cake, game, plane, cave, safe	3	/æ/-a /ɛ/-e
Phonology	GPC	/i:/	ee	eel, bee, feed, tree, seed, peel	2	/ɛ/-e
Phonology	GPC	/i/	ie	movie, relief, thief	3	/ɛ/-e /ɪ/-i
Phonology	GPC	/i:/	ea	eat, sea, tea, meat	3	/æ/-a /ɛ/-e
Phonology	GPC	/i:/	e	be, she, he	3	/ɛ/-e
Phonology	GPC	/i:/	e_e	eve, these, extreme, theme	3	/ɛ/-e
Phonology	GPC	/i/	ey	key, monkey, donkey	3	/ɛ/-e /i/-y
Phonology	GPC	/i/	y	sunny, happy, baby, any	3	None

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/aɪ/	igh	sigh, high, night, light	2	/ɪ/-i /g/-g /h/-h
Phonology	GPC	/aɪ/	ie	tie, lie, pie	3	/ɪ/-i /ɛ/-e
Phonology	GPC	/aɪ/	i	behind, idea, item	3	/ɪ/-i
Phonology	GPC	/aɪ/	y	fly, cry, shy	3	/i/-y
Phonology	GPC	/aɪ/	i_e	bike, fine, like, kite	3	/ɪ/-i /ɛ/-e
Phonology	GPC	/əʊ/	oa	boat, oak, float, coat	2	/ɒ/-o /æ/-a
Phonology	GPC	/əʊ/	ow	blow, snow, window, show	3	/ɒ/-o /w/-w
Phonology	GPC	/əʊ/	o	go, no, yo-yo	3	/ɒ/-o
Phonology	GPC	/əʊ/	oe	oboe, toe, video	3	/ɒ/-o /ɛ/-e
Phonology	GPC	/əʊ/	o_e	rope, hope, bone, home	3	/ɒ/-o /ɛ/-e
Phonology	GPC	/ʊ/	oo	book, room, cook	3	/ɒ/-o
Phonology	GPC	/ʊ/	ou	should, would, could	3	/ɒ/-o

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
						/ʌ/-u
Phonology	GPC	/ʊ/	u	put, bull, bush	3	/ʌ/-u
Phonology	GPC	/u:/	oo	moon, spoon, food	2	/ɒ/-o
Phonology	GPC	/u:/	ue	blue, clue, statue	3	/ʌ/-u /ɛ/-e
Phonology	GPC	/u:/	u_e	cube, tube, flute	3	/ʌ/-u /ɛ/-e
Phonology	GPC	/u:/	ew	new, grew, flew	3	/ɛ/-e /w/-w
Phonology	GPC	/u:/	ou	soup, group, route	3	/ɒ/-o /ʌ/-u
Phonology	GPC	/u:/	u	super, tuna, rude	3	/ʌ/-u
Phonology	GPC	/ju:/	ue	statue, due,	3	/ʌ/-u /ɛ/-e
Phonology	GPC	/ju:/	u	unicorn, uniform, use	3	/ʌ/-u
Phonology	GPC	/ɔɪ/	oi	coin, avoid, join	2	/ɒ/-o /ɪ/-i
Phonology	GPC	/ɔɪ/	oy	toy, boy, joy	3	/ɒ/-o /i/-y
Phonology	GPC	/aʊ/	ou	shout, about, count, mouth	3	/ɒ/-o

Linguistic Level	**Category**	**Feature**	**Examples**	**Difficulty level of feature**	**Prerequisites**
					/ʌ/-u
Phonology	GPC	/aʊ/ ow	owl, brown, how	2	/ɒ/-o /w/-w
Phonology	GPC	/ɑː/ ar	artist, alarm, barn	2	/æ/-a /r/-r
Phonology	GPC	/ɑː/ a	ask, bath, fast, father	2	/æ/-a
Phonology	GPC	/ɔː/ or	fork, sort, torn	2	/ɒ/-o /r/-r
Phonology	GPC	/ɔː/ aw	dawn, draw, paw	3	/æ/-a /w/-w
Phonology	GPC	/ɔː/ au	sauce, author, pause	3	/æ/-a /ʌ/-u
Phonology	GPC	/ɜː/ er	her, alert, nerve	2	/ɛ/-e /r/-r
Phonology	GPC	/ɜː/ ir	birthday, girl, firm	3	/ɪ/-i /r/-r
Phonology	GPC	/ɜː/ ur	nurse, hurt, turn	2	/ʌ/-u /r/-r
Phonology	GPC	/ɜː/ or	worm, work	3	/ɒ/-o /r/-r
Phonology	GPC	/ɜː/ ear	earth, learn, pearl	3	/æ/-a

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
					/ɛ/-e /r/-r
Phonology	GPC	/eə/ air	hair, pair, chair	2	/æ/-a /ı/-i /r/-r
Phonology	GPC	/eə/ are	care, square, dare	3	/æ/-a /ɛ/-e /r/-r
Phonology	GPC	/eə/ ear*	bear, pear, wear	2	/æ/-a /ɛ/-e /r/-r
Phonology	GPC	/eə/ ere	where, there, anywhere	3	/ɛ/-e /r/-r
Phonology	GPC	/ıə/ eer	deer, peer, cheer	3	/ɛ/-e /r/-r
Phonology	GPC	/ıə/ ear	ear, near, tear	2	/ɛ/-e /æ/-a /r/-r
Phonology	GPC	/ıə/ ere	interfere, adhere, here	3	/ɛ/-e /r/-r
Phonology	GPC	/b/ b	bat, bed, bug	1	None

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/b/	bb	rabbit, pebble, rubber	3	/b/-b
Phonology	GPC	/d/	d	dog, did, dad	1	None
Phonology	GPC	/d/	dd	puddle, teddy, daddy	3	/d/-d
Phonology	GPC	/f/	f	fan, fig, fog	1	None
Phonology	GPC	/f/	ff (end position)	cliff, sniff, off	2	/f/-f
Phonology	GPC	/f/	ff (middle position)	giraffe, affect, toffee	3	/f/-ff (end position)
Phonology	GPC	/f/	ph	photo, dolphin, phone	3	/f/-f
Phonology	GPC	/g/	g	girl, gum, gift	1	None
Phonology	GPC	/g/	gg	juggle, doggy, egg	3	/g/-g
Phonology	GPC	/h/	h	hat, hug, hop	1	None
Phonology	GPC	/dʒ/	j	jug, jam, jet	2	None
Phonology	GPC	/dʒ/	g	giant, age, village	3	/dʒ/-j /g/-g
Phonology	GPC	/dʒ/	ge	cabbage, page, stage	3	/dʒ/-j /ɛ/-e
Phonology	GPC	/dʒ/	dge	fridge, lodge, edge	3	/g/-g /dʒ/-g /d/-d

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
						/ɛ/-e
Phonology	GPC	/l/	l	left, lid, leg	1	None
Phonology	GPC	/l/	ll (end position)	ball, will, doll	2	/l/-l
Phonology	GPC	/l/	ll (any position)	alligator, million, yellow	3	/l/-ll (end position)
Phonology	GPC	/m/	m	map, mix, man	1	None
Phonology	GPC	/m/	mm	hammer, mummy, swimmer	3	/m/-m
Phonology	GPC	/kw/	qu	queen, quack, quest	2	/ʌ/-u
Phonology	GPC	/r/	r	rat, run, rob	1	None
Phonology	GPC	/r/	rr	arrow, carrot, parrot	3	/r/-r
Phonology	GPC	/r/	wr	write, wrap, wreck	3	/r/-r /w/-w
Phonology	GPC	/v/	v	van, visit, vet	2	None
Phonology	GPC	/v/	ve	dove, love, live		/v/-v /ɛ/-e
Phonology	GPC	/w/	w	web, won, war wall	2	None
Phonology	GPC	/w/	wh	wheel, whale, whip	3	/w/-w /h/-h
Phonology	GPC	/ks/	x	fox, six, taxi	2	None
Phonology	GPC	/j/	y	yolk, yes	2	/i/-y

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/z/	z	zip, zoo, zebra, zero	2	None
Phonology	GPC	/z/	zz (end position)	jazz, buzz	2	/z/-z
Phonology	GPC	/z/	zz (any position)	dazzle, sizzle, fuzzy	3	/z/-zz (end position)
Phonology	GPC	/z/	se	cheese, cause, chinese	3	/s/-s /ɛ/-e /z/-z
Phonology	GPC	/z/	ze	breeze, daze, zero	3	/z/-z /ɛ/-e
Phonology	GPC	/tʃ/	ch	chairs, check, chop	2	/h/-h /k/-c
Phonology	GPC	/tʃ/	tch	patch, fetch, pitch	3	/h/-h /k/-c /t/-t
Phonology	GPC	/ʃ/	sh	sheep, shell, shake	2	/s/-s /h/-h
Phonology	GPC	/ʃ/	ti	station, patient	3	/t/-t /ı/-i
Phonology	GPC	/θ/	th	thistle, thirty, thin	2	/t/-t /h/-h
Phonology	GPC	/ð/	th	there, breathe, together	2	/t/-t /h/-h

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	GPC	/ʒ/	s	treasure, occasion, leisure	3	/s/-s
Phonology	GPC	/ʒ/	ge	collage, courgette, strange	3	/ɛ/-e /dʒ/-g
Phonology	Consonant Clusters	/b/ + /l/	bl, initial	black, block, bless	4	/b/-b /l/-l
Phonology	Consonant Clusters	/c/ + /l/	cl, initial	clock, clever, cliff	4	/k/-c /l/-l
Phonology	Consonant Clusters	/f/ + /l/	fl, initial	flag, flood, fly	4	/f/-f /l/-l
Phonology	Consonant Clusters	/g/ + /l/	gl, initial	glad, glove, glitter	4	/g/-g /l/-l
Phonology	Consonant Clusters	/p/ + /l/	pl, initial	plan, plot, plate	4	/p/-p /l/-l
Phonology	Consonant Clusters	/s/ + /l/	sl, initial	slap, slide, slurp	4	/s/-s /l/-l
Phonology	Consonant Clusters	/b/ + /r/	br, initial	brick, bring, brush	4	/b/-b /r/-r
Phonology	Consonant Clusters	/c/ + /r/	cr, initial	crab, crisp, crowd	4	/k/-c /r/-r
Phonology	Consonant Clusters	/d/ + /r/	dr, initial	drop, dress, dragon	4	/d/-d /r/-r

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
Phonology	Consonant Clusters	/f/ + /r/	fr, initial	frog, frozen, fries	4	/f/-f /r/-r
Phonology	Consonant Clusters	/g/ + /r/	gr, initial	grab, grapes, green	4	/g/-g /r/-r
Phonology	Consonant Clusters	/p/ + /r/	pr, initial	pram, problem, proud	4	/p/-p /r/-r
Phonology	Consonant Clusters	/t/ + /r/	tr, initial	train, trick, trouble	4	/t/-t /r/-r
Phonology	Consonant Clusters	/s/ + /c/	sc, initial	scale, scatter, scrap	4	/s/-s /k/-c
Phonology	Consonant Clusters	/s/ + /k/	sk, initial	skate, skip, skateboard	4	/s/-s /k/-k
Phonology	Consonant Clusters	/s/ + /n/	sn, initial	snake, snap, snow	4	/s/-s /n/-n
Phonology	Consonant Clusters	/s/ + /m/	sm, initial	smile, smart, smoky	4	/s/-s /m/-m
Phonology	Consonant Clusters	/s/ + /t/	st, initial	step, start, stick	4	/s/-s /t/-t
Phonology	Consonant Clusters	/s/ + /w/	sw, initial	swam, swallow, swirl	4	/s/-s /w/-w
Phonology	Consonant	/t/ + /w/	tw, initial	twin, twice, twirl	4	/t/-t

Linguistic Level	**Category**	**Feature**		**Examples**	**Difficulty level of feature**	**Prerequisites**
	Clusters					/w/-w
Phonology	Consonant Clusters	/s/ + /c/ + /h/	sch, initial	school, scheme	4	/s/-s /k/-c /h/-h
Phonology	Consonant Clusters	/s/ + /t/ + /r/	str, initial	straw, street, stripe	4	/s/-s /t/-t /r/-r
Phonology	Consonant Clusters	/s/ + /c/ + /r/	scr, initial	scrap, scrub, screw	4	/s/-s /k/-c /r/-r
Phonology	Consonant Clusters	/s/ + /p/ + /r/	spr, initial	spray, spring, spread	4	/s/-s /p/-p /r/-r
Phonology	Consonant Clusters	/s/ + /p/ + /l/	spl, initial	splash, split, splendid	4	/s/-s /p/-p /l/-l
Phonology	Consonant Clusters	/s/ + /q/ + /u/	squ, initial	square, squeak, squid	4	/s/-s /p/-p /h/-h
Phonology	Consonant Clusters	/s/ + /h/ + /r/	shr, initial	shrink, shrug, shriek	4	/s/-s /h/-h

Linguistic Level	Category	Feature		Examples	Difficulty level of feature	Prerequisites
						/r/-r
Phonology	Consonant Clusters	/t/ + /h/ + /r/	thr, initial	throw, thrill, three	4	/t/-t /h/-h /r/-r
Phonology	Consonant Clusters	/c/ + /t/	ct, end	pact, act, connect	4	/k/-c /t/-t
Phonology	Consonant Clusters	/f/ + /t/	ft, end	left, soft, gift	4	/f/-f /t/-t
Phonology	Consonant Clusters	/n/ + /t/	nt, end	bent, cement, want	4	/n/-n /t/-t
Phonology	Consonant Clusters	/l/ + /t/	lt, end	belt, adult, jolt	4	/l/-l /t/-t
Phonology	Consonant Clusters	/s/ + /t/	st, end	quest, cast, ghost	4	/s/-s /t/-t
Phonology	Consonant Clusters	/n/ + /d/	nd, end	hand, wind, almond	4	/n/-n /d/-d
Phonology	Consonant Clusters	/m/ + /p/	mp, end	lamp, chimp, damp	4	/m/-m /p/-p
Phonology	Consonant Clusters	/c/ + /k/	ck, end	back, wreck, snack	4	/k/-c /k/-k
Phonology	Consonant	/l/ + /k/	lk, end*	milk, yolk, talk	4	/l/-l

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
	Clusters				/k/-k
Phonology	Consonant Clusters	/n/ + /k/ nk, end	bank, wink, shrink	4	/n/-n /k/-k
Phonology	Consonant Clusters	/s/ + /k/ sk, end	whisk, mask, ask	4	/s/-s /k/-k
Phonology	Consonant Clusters	/t/ + /c/ + /h/ tch, end	scratch, itch, watch	4	/t/-t /h/-h /k/-c
Phonology	Syllabification	2-syllable		1	
Phonology	Syllabification	3-syllable		1	2-syllable
Phonology	Syllabification	4-syllable		2	3-syllable
Phonology	Syllabification	5-syllable		2	4-syllable
Phonology	Syllabification	6-syllable		3	5-syllable
Word recognition		Feature-1	a, an, and, as, can, back		None
Word recognition		Feature-2	them, this, too, was, we		1
Word recognition		Feature-3	it's, just, like, little, one, out, said		2
Word		Feature-4	called, came, could, day, don't, here		3

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
recognition					
Word recognition		Feature-5	did, even, eyes, good, head, home		4
Word recognition		Feature-6	must, really, school, should, tell, things		5
Word recognition		Feature-7	hard, hear, inside, land, let, might, name		6
Word recognition		Feature-8	animals, became, bed, dragon, each, everything		7
Word recognition		Feature-9	different, duck, eggs, felt, floppy		8
Word recognition		Feature-10 (infrequent GPCs)	magazine, listen, pizza, rough, scene		All GPCs
Orthography		/b/, /d/	bad, dad		/b/-b, /d/-d
Orthography		/p/, /q/	pick, quick		/p/-p,/q/-q
syntax	adjectives (modifiers)	attributive	Maria has a nice dress.	2	predicative
syntax	adjectives (modifiers)	predicative	Maria's dress is nice.	1	none

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
syntax	adjectives (modifiers)	secondary predicate	I painted the wall green. I like my coffee hot. He came home drunk.	3	predicative, attributive,
syntax	Binding	Reciprocal: each other	The girls like each other.	2	myself/yourself/himself/herself/itself/ourselves/yourselves/theirselves in binding structures
syntax	Binding	Personal pronouns: me/you/him/her/it/us/you/them in binding structures	Sally arrived, but nobody saw her.	3	myself/yourself/himself/herself/itself/ourselves/yourselves/theirselves in binding structures, each other
syntax	Binding	Reflexive pronouns: myself/yourself/himself/herself/itself/ourselves/yourselves/theirselves in binding structures	Sam is impressed by himself.	1	none
syntax	Coordination	or, and, nor, but, or, yet, so, and nor, but nor, or nor, neither, no more, only	They gamble, andthey smoke. They do not gambleorsmoke, for they are ascetics. Theydon'tgamble,neitherdo they smoke, They don't gamble,no moredo they smoke, I would go,onlyI don't have time.	1	none
syntax	Coordination	either...or, not only...but (also),	You either do your work or prepare	2	or, and, nor, but, or, yet, so,

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
		neither...nor, both...and, whether...or, just as...so, the...the, as...as, as much...as, no sooner...than, rather..than	for a trip to the office.Neitherthe basketball teamnorthe football team is doing well. Football isas fast ashockey.		and nor, but nor, or nor, neither, no more, only
syntax	Discourse anaphors	Personal pronouns: 1st/2nd person sing/pl: I, you, we, me, you, us	I am / You are / We are sick	1	First name alone, Full proper name
syntax	Discourse anaphors	Personal pronouns: 3rd person sing/pl: he, she, it, they, him, her, it, them	He / She / It / They are new	1	I, you, we, me, you, us
syntax	Discourse anaphors	NPs with demonstrative determiners: This, that	this chair over here	2	Bare NPs
syntax	Discourse anaphors	NPs with genitive NPs as determiner	my best friend's wedding, our house	2	I, you, we, me, you, us - First name alone, Full proper name
syntax	Discourse anaphors	First name alone, Full proper name	Julia, Julia Child	1	none
syntax	Discourse anaphors	Bare NPs	Pencils [are plentiful/made of wood], beauty [is eternal], Children [are crying], snow [was piled high]	1	none
syntax	Discourse anaphors	definite article (the)+N	Can you give me the plate?	2	Bare NPs. indefinite article (a/an)+N

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
syntax	Discourse anaphors	indefinite article (a/an)+N	Can you give me a plate?	2	Bare NPs,
syntax	Embedding	Sequence: as, after, as soon as, at first, once, before, since, then, by the time, until, when, whenever, while,	He came home after/beforeme. I was asleepwhenshe called me.By the timethe firefighters arrived, the building had already burnt down. The girl always sits on the couchwhenshe eats her lunch.	2	Causation: although, because, if, despite, since, though, thus, unless, until, as long as, so that,
syntax	Embedding	Causation: although, because, if, despite, since, though, thus, unless, until, as long as, so that,	He died because he had cancer. She drives a carin order togo home.	1	clause + clause, verb+to+inf
syntax	Embedding	Conditional (if), Temporal (when, while, after, before, once, until, since, as, as soon as, as long as, Causal (because, since, as)	If you are careful, we can get at the top. When you get back here, we have dry hooves.It had certainly changed in the short hour since I had come out.	3	Sequence: as, after, as soon as, at first, once, before, since, then, by the time, until, when, whenever, while, Causation: although, because, if, despite, since, though, thus, unless, until, as long as, so that,
syntax	Embedding	that-RCs, right-branching	I love the man, that is drawing the kid.	2	that-RCs, subject extracted, that-RCs, object extracted

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
syntax	Embedding	that-RCs, center embedded	The man, that is drawing the kid, is tall.	3	that-RCs, subject extracted, that-RCs, object extracted, that-RCs, right-branching
syntax	Embedding	that-RCs, subject extracted	I love the man that is drawing the kid.	1	Coordination, clause + clause, verb+to+inf,
syntax	Embedding	that-RCs, object extracted	I love the manthatthe kid is drawing.	2	that-RCs, subject extracted
syntax	Embedding	RCs with a relative pronoun	I met the man who wasn't very kind on the phone. I love the man who lives in this house.	2	that-RCs, subject extracted, that-RCs, object extracted
syntax	Embedding	Reduced relative clauses	The man I saw yesterday went home.	1	clause + clause, verb+to+inf
syntax	Embedding	Complementizers (that, who, what, whether, empty complementizers)	This is the book that every student should read. This is the book which every student should read.	2	Reduced relative clauses, RCs with a relative pronoun, that -RCs, subject extracted, Sequence, Causation,
syntax	Embedding	Complement clauses (wish, think, understand, recommend, decide, regret,)	I wish that I could join you. She thoughtthat all kids were hungry.	2	Complementizers (that, who, what, whether, empty complementizers)
syntax	Negation	do not, don't, am not, is not, did not, have not, haven't, had not, hadn't, should not, shouldn't, would not, wouldn't, could not,	I did notgo to Moscow. I had not had dinner when they came in. You shouldnotsleep late. Not like this.	1	none

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
		couldn't, not,			
syntax	Noun phrases	more than one adj+N	soft gentle flow	2	None
syntax	Passives	without agent by-phrase (short passives)	The cut was kicked.	1	none
syntax	Passives	with agent by-phrase (long passives)	The woman was kissed by the boy.	2	without agent by-phrase (short passives)
syntax	Prepositional phrases	(down,in, in front of, on, above, behind, under, between, besides, upstairs, near, out, over, outside, inside, through, next to, up)	Put down the cup.	1	none
syntax	Prepositional phrases	(of, about, for, from, at, with, by)	Translation of the book. A man of honour.	3	(down,in, in front of, on, above, behind, under, between, besides, upstairs, near, out, over, outside, inside, through, next to, up - to, before, after, until)
syntax	Prepositional phrases	(to, before, after, until)	Run to the store.	2	(down,in, in front of, on, above, behind, under, between, besides, upstairs, near, out, over, outside, inside, through, next to, up)

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
syntax	Prepositional phrases		The person that I was talking to	3	(down,in, in front of, on, above, behind, under, between, besides, upstairs, near, out, over, outside, inside, through, next to, up - to, before, after, until)
syntax	Quantifiers	A few, few, fewer, fewest	A few frogs are green.	1	none
syntax	Quantifiers	A number of, Another	I have anotherexample.	1	none
syntax	Quantifiers	Several		1	none
syntax	Quantifiers	Many	Manyfrogs are green.	1	none
syntax	Quantifiers	A little, little, less, the least	Can I have a littlesugar please?	2	A few, few, fewer, fewest - A number of, Another - Several - Many
syntax	Quantifiers	A bit of, A great deal of, A large amount of, Much		2	A few, few, fewer, fewest - A number of, Another - Several - Many
syntax	Quantifiers	all, enough, more/most, less/least, no/none, not any, some, any, a lot of, lots of plenty of	I don't have anytime. A lot of people are smart.Allfrogs are green.Most of the people I talked to didn't have a clue who the candidates were.Someof	3	A few, few, fewer, fewest - A number of, Another - Several - Many- A little, little, less, the least - A bit of, A great

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
			the people standing across the river have white armbands.		deal of, A large amount of, Much
syntax	Wh-Questions	non-referential (what) questions, object-extracted	**What is the man chasing?**	1	non-referential (what) questions, subject-extracted
syntax	Wh-Questions	referential (which-NP) Questions	**Which kid is pulling the teacher?Who is pulling the teacher?**	4	non-referential (what) questions, subject-extracted,non-referential (what) questions,object-extracted,non-referential (who) questions, subject-extracted, what, non-referential (what) questions, subject-extracted,non-referential (what) questions,object-extracted,non-referential (who) questions, object-extracted
syntax	Wh-Questions	non-referential (who) questions, subject-extracted	Who is chasing the man?	2	non-referential (what) questions, subject-extracted
syntax	Wh-Questions	non-referential (who) questions,	**Who is the man chasing?**	3	non-referential (what)

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
		object-extracted			questions, subject-extracted,non-referential (what) questions,object-extracted,non-referential (who) questions, subject-extracted
syntax	Wh-Questions	adjunct questions	**Where/When did they meet?**	3	yes/no questions, non-referential (what) questions, subject-extracted, referential (who) questions, subject-extracted
syntax	Wh-Questions	non-referential (what) questions, subject-extracted	**What is chasing the man?**	1	none
syntax	Yes-No Questions	yes/no questions	Do you love him?	1	none
syntax	Modals	will/would/shall	I will leave tomorrow / I would leave the next day / I shall	1	none
syntax	Modals	can/may/might/could	He can/may/might/could pass the exam	2	will/would/shall
syntax	Modals	should/must/(ought to/have to)	I should/must/ought to/have to study	3	will/would/shall, can/may/might/could

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
morphology	prefixes	re (= again)	rewrite a letter	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	semi (= half), mono (= one), multi (= many), super (= big/more), sub (= under/less), mini (= small)	semi-skilled workers, monolingual, a multinational company, a superhuman effort, sub-zero temperature, a minicomputer	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	pre (= before),post (= after), ex (= previously)	prehistoric times, the post- war period , our ex-Director	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	inter (= between), trans (= across)	an international phone call, a heart transplant operation	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	co (= together)	my co-driver	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	over (= too much), under (= too	an overgrown garden, underpaid,	2	-ment/-ion/-tion/-ation/-

Linguistic Level	**Category**	**Feature**	**Examples**	**Difficulty level of feature**	**Prerequisites**
		little), out (= more/better)	outplayed their opponents		ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	pro (= in favour of), anti (= against), mis (= badly/wrongly)	pro-European policies, anti-aircraft guns, a misunderstanding	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	un, in, dis	unhappy, unfair, independent, indirect, dishonest, disunited	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	prefixes	non, de	non-alcoholic drinks, a non-stop flight, defrost a fridge, the depopulation of the countryside	2	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee, able/ible/ing/ed
morphology	derivational suffixes	-ment/-ion/-tion/-ation/-ition/-sion/-ance/-ence/-ing/-er/-or/-ant/-ent/-ee	payment, movement, correct - correction, discuss -discussion, inform - information, add - addition, decide - decision, permit - permission, performance, acceptance, existence,building, my feelings,	1	none

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
			walker, owner, applicant, assistant,employee, payee,		
morphology	derivational suffixes	-ent/-ence/-ty/-ity/-ness	silent - silence , ertainty, royalty, stupidity, happiness, illness, freshness	1	none
morphology	derivational suffixes	Noun + an/ian/ess/an/ian/ess=noun	republican, electrician,waitress, actress,republican, electrician,waitress, actress	1	none
morphology	derivational suffixes	Noun/Verb/Adjective + ist=noun	journalist, motorist	1	none
morphology	derivational suffixes	-ize/-en	modernize, popularize,shorten, widen	1	none
morphology	derivational suffixes	Noun + al/ic=adjective	national, industrial,heroic, artistic	1	none
morphology	derivational suffixes	Verb/Noun + ive=adjective	active, effective	1	none
morphology	derivational suffixes	Noun + ful/ less/ous/y/ly=adjective	careful, hopeful,careless, hopeless,dangerous, luxurious,salty, healthy,friendly, costly	1	(-ly),
morphology	derivational suffixes	able/ible/ing/ed	eatable, manageable,exciting, fascinating,excited, fascinated	1	none
morphology	derivational	(-ly)	quickly	1	none

Linguistic Level	**Category**	**Feature**	**Examples**	**Difficulty level of feature**	**Prerequisites**
	suffixes				
morpho-syntax	Adverbs		more slowly		none
morpho-syntax	Agreement	auxiliary 'be', present and past	am/is/are/was/were	1	none
morpho-syntax	Agreement	auxiliary 'have', present and past	have/has/had	3	auxiliary 'be', present and past, auxiliary 'do', present and past
morpho-syntax	Agreement	auxiliary 'do', present and past	do/does/did	2	auxiliary 'be', present and past
morpho-syntax	Inflectional Suffixes: Agreement	present tense -s (regular)	plays, wants	1	none
morpho-syntax	Inflectional Suffixes: Agreement	present tense -es (after s, ss, sh, ch, z)	messes, pushes, washes, watches	1	none
morpho-syntax	Inflectional Suffixes: NUMBER (regular)	plural number -s, pronounced as /s/ or /z/	dogs, cats	1	none
morpho-	Inflectional	plural number -es, pronounced as	buses, bushes, dishes, matches	2	plural number -s,

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
syntax	Suffixes: NUMBER (regular)	/iz/ (after s, ss, sh, ch, z)			pronounced as /s/ or /z/
morpho-syntax	Inflectional Suffixes: TENSE (regular)	past tense (-ed), pronounced as /t/, /d/ or /id/	watched, judged, wanted	2	none
morpho-syntax	Inflectional Suffixes: TENSE (regular)	past tense (-d), vowel drop	loved	2	past tense (-ed), pronounced as /t/, /d/ or /id/
morpho-syntax	Inflectional Suffixes: TENSE (regular)	past tense (double C + ed, pronounced as /t/ or /d/)	stopped, jogged	3	past tense (-ed), pronounced as /t/, /d/ or /id/, past tense (-d), vowel drop
morpho-syntax	Inflectional Suffixes: TENSE (regular)	imperfective/progressive (be + verb+ing): +ing	calling, playing, watching	4	imperfective/progressive (be + verb+ing): -ie --> y +ing, imperfective/progressive (be + verb+ing): double C + ing
morpho-syntax	Inflectional Suffixes: TENSE (regular)	imperfective/progressive (be + verb+ing): double C + ing	getting, stopping	4	imperfective/progressive (be + verb+ing): +ing,imperfective/progressive (be + verb+ing): -ie --> y +ing,

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
					imperfective/progressive (be + verb+ing): double C + ing
morpho-syntax	Inflectional Suffixes: TENSE (regular)	imperfective/progressive (be + verb+ing): V drop + ing	love->loving, receiving	4	imperfective/progressive (be + verb+ing): +ing,imperfective/progressive (be + verb+ing): -ie --> y +ing, imperfective/progressive (be + verb+ing): double C + ing
morpho-syntax	Inflectional Suffixes: TENSE (regular)	imperfective/progressive (be + verb+ing): -ie --> y +ing	lie->lying	4	imperfective/progressive (be + verb+ing): +ing,imperfective/progressive (be + verb+ing): -ie --> y +ing, imperfective/progressive (be + verb+ing): double C + ing
morpho-syntax	Inflectional Suffixes: TENSE (regular)	perfect forms (have+en/ed)	will have written, had spoken, have spoken,	3	past tense (-ed), pronounced as /t/, /d/ or /id/. past tense (-d), vowel drop. past tense (double C + ed, pronounced as /t/ or /d/)

Linguistic Level	**Category**	**Feature**	**Examples**	**Difficulty level of feature**	**Prerequisites**
Morphosyntax (Lexical)	irregular past participles	<List to be provided>			
morphosyntax	Adjectives	Comparative, <2 syllables, : -er	warmer than	2	Comparative, <2 syllables, exceptions: double C +er, -y-->+ier
morphosyntax	Adjectives	Comparative, <2 syllables, exceptions: double C +er, -y-->+ier	hotter, happier	1	none
morphosyntax	Adjectives	Comparative, >2 syllables: more Adj than	more beautiful than	2	Comparative, <2 syllables, exceptions: double C +er, -y-->+ier, Comparative, <2 syllables, : -er
morphosyntax	Adjectives	Superlative, <2 syllables, : +est	the +est (the warmest)	3	Comparative, <2 syllables, : -er, Comparative, <2 syllables, exceptions: double C +er, -y-->+ier, Comparative, >2 syllables: more Adj than, Superlative, <2 syllables, exceptions: double C +est, -y-->+iest
morphosynta	Adjectives	Superlative, <2 syllables,	hottest, happiest	3	Comparative, <2 syllables, : -

Linguistic Level	Category	Feature	Examples	Difficulty level of feature	Prerequisites
x		exceptions: double C +est, -y-->+iest			er, Comparative, <2 syllables, exceptions: double C +er, -y-->+ier, Comparative, >2 syllables: more Adj than, Superlative, <2 syllables, : +est
morphosyntax	Adjectives	Superlative, >2 syllables: the most +Adj	the most beautiful	4	Comparative, <2 syllables, : -er, Comparative, <2 syllables, exceptions: double C +er, -y-->+ier, Comparative, >2 syllables: more Adj than, Superlative, <2 syllables, : +est, Superlative, <2 syllables, exceptions: double C +est, -y-->+iest
Morphosyntax (Lexical)	irregular comparatives/superlatives	better/best, worse/worst, more/most, less/least, fewer		4	Comparative, <2 syllables, : -er, Comparative, <2 syllables, exceptions: double C +er, -y-->+ier, Comparative, >2 syllables: more Adj than, Superlative, <2 syllables, :

Linguistic Level	**Category**	**Feature**	**Examples**	**Difficulty level of feature**	**Prerequisites**
					+est, Superlative, <2 syllables, exceptions: double C +est, -y-->+iest
Morphosyntax (Lexical)	irregular past forms	<List to be provided>			
Morphosyntax (Lexical)	irregular plural forms	<List to be provided>			

Appendix II. The Greek domain model for dyslexia

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Syllabification	cv-cv	κα-λα		1	none
Phonology	Syllabification	v-cv	α-πο		1	none
Phonology	Syllabification	cv-cv(c)	κα-λό(ς)	A	2	none
Phonology	Syllabification	v-vc	έ-ως		2	cv-cv, v-cv
Phonology	Syllabification	cv-vc	χά-ος		2	cv-cv, v-cv
Phonology	Syllabification	vc-cv(c)	ασ-βοί, ασ-βός	B	3	cv-cv, v-cv, v-vc,cv-cv(c)
Phonology	Syllabification	cvc-cv(c)	καρ-πό(ς)		3	cv-cv (C) , v-vc, cv-vc
Phonology	Syllabification	cv-ccv(c)	μι-κρό(ς)	C	3	none
Phonology	Syllabification	v-ccv(c)	α-σβό(ς)		3	none
Phonology	Syllabification	cv-cccv(c)	κά-στρο	D	4	cv-cv (C) , v-vc, cv-vc
Phonology	Syllabification	v-cccv(c)	ά-σπρο(ς)		4	cv-cv (C) , v-vc, cv-vc, v-ccv(c))
Phonology	Syllabification	ccv-cv(c)	σκά-βει(ς)		3	none
Phonology	Syllabification	cccv-cv(c)	στρα-βό(ς)		4	cv-cv (C) , v-vc, cv-vc,cv-ccv(c), ccv-cv(c)
Phonology	Syllabification	ccvc-cv(c)	σκάρ-το(ς)		4	cv-cv (C) , v-vc, cv-vc,cv-ccv(c), ccv-cv(c)
Phonology	Syllabification	ccv-ccv(c)	σπα-στό(ς)		4	cv-cv (C) , v-vc, cv-vc, cv-ccv(c)
Phonology	Syllabification	ιά	καρ-διά	E	5	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Syllabification	ειά	για-τρει-ά		5	none
Phonology	Syllabification	ιά	καρ-**διά / **αρ-**γί-α**	F	5	none
Phonology	Syllabification	ία	καρ-**διά / **αρ-**γί-α**		5	none
Phonology	Syllabification	αί	**παί**-ζω		6	none
Phonology	Syllabification	αΐ	**πα-ΐ**-δια		6	none
Phonology	Syllabification	εί	προ-**τεί**-νω		6	none
Phonology	Syllabification	εΐ	πρω-**τε-ΐ**-νη		6	none
Phonology	Syllabification	οί	α-**θροί**-ζω		6	none
Phonology	Syllabification	οΐ	**θρο-ΐ**-ζω		6	none
Phonology	Syllabification	αί	μαί-α	G	5	none
Phonology	Syllabification	ά	μά-ι-ος		5	none
Phonology	Syllabification	οί	ρόλο		5	none
Phonology	Syllabification	όι	ρο-λό-ι		5	none
Phonology	Syllabification	άι (diphthongs)	**γάι**-δα-ρος	H	5	none
Phonology	Syllabification	όι (diphthongs)	κο-**ρόι**-δο		5	none
Phonology	Syllabification	αι	**παι**-δί	I	6	none
Phonology	Syllabification	αϊ	**πα-ϊ**-δάκι		6	none
Phonology	Syllabification	οι	**μοί**-ρα		6	none
Phonology	Syllabification	οϊ	**προ-ϊ**-στορία		6	οί, όι
Phonology	Syllabification	ει	α-**στεί**-ο		6	none
Phonology	Syllabification	εϊ	α-**στε-ϊ**-σμός		6	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Syllabification	ου	**κου**-τά-λι	J	6	none
Phonology	Syllabification	οϋ	**προ-ϋ**-πόθεση		6	none
Phonology	Syllabification	ού	με-**δού**-λι		6	ου
Phonology	Syllabification	οΰ	αρα-**χνο-ϋ**-φαντος		6	ου
Phonology	Syllabification	αυ	**αυ**-λή/**α-ϋ**-πνί-α,	K	6	ά
Phonology	Syllabification	αϋ	**αυ**-λή/**α-ϋ**-πνί-α,		6	ά
Phonology	Syllabification	αύ	**αυ**-λή/**α-ϋ**-πνί-α,		6	ά
Phonology	Syllabification	αΰ	**αυ**-λή/**α-ϋ**-πνί-α,		6	ά
Phonology	Syllabification	εύω	κο-ροϊ-**δεύ**-ω		5	none
Phonology	Phonemes	/t/ Initial	*/tino/*	A	2	none
Phonology	Phonemes	/d/ Initial	*/dino/*		2	none
Phonology	Phonemes	/p/ Initial	*/pino/*		2	none
Phonology	Phonemes	/b/ Initial	*/beno/*		2	none
Phonology	Phonemes	/k/ Initial	*/kano/*		1	none
Phonology	Phonemes	/m/ Initial	*/meno/*	B	1	none
Phonology	Phonemes	/n/ Initial	*/nemo/*		1	none
Phonology	Phonemes	/θ/ Initial	*/ϑelo/*	C	3	none
Phonology	Phonemes	/ð/ Initial	*/ðeno/*		3	none
Phonology	Phonemes	/f/ Initial	*/fovos/*		3	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Phonemes	/v/ Initial	*/vazo/*		3	none
Phonology	Phonemes	/χ/ Initial	*/χano/*		3	none
Phonology	Phonemes	/γ/ Initial	*/γoma/*		3	none
Phonology	Phonemes	/s/ Initial	*/soma/*	D	1	none
Phonology	Phonemes	/z/ Initial	*/zoni/*		1	none
Phonology	Phonemes	/l/ Initial	*/lemos/*	E	2	none
Phonology	Phonemes	/r/ Initial	*/roda/*		2	none
Phonology	Phonemes	/kt/ Initial	*/ktinos/*	F	4	/t/ Initial,/k/ Initial
Phonology	Phonemes	/pt/ Initial	*/ptino/*		4	/p/ Initial,/k/ Initial
Phonology	Phonemes	/ks/ Initial	*/ksino/*	G	4	/s/ Initial,/k/ Initial
Phonology	Phonemes	/ps/ Initial	*/psino/*		4	/s/ Initial,/p/ Initial
Phonology	Phonemes	/sk/ Initial	*/skini/*		4	/s/ Initial,/k/ Initial
Phonology	Phonemes	/sp/ Initial	*/sperno/*		4	/s/ Initial,/p/ Initial
Phonology	Phonemes	/ðr/ Initial	*/ðromos/*	H	5	/ð/ Initial, /r/ Initial
Phonology	Phonemes	/θr/ Initial	*/ϑronos/*		5	/θ/ Initial, /r/ Initial

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Phonemes	/fr/ Initial	*/freno/*		5	/f/ Initial, /r/ Initial
Phonology	Phonemes	/χr/ Initial	*/χroma/*		5	/χ/ Initial, /r/ Initial
Phonology	Phonemes	/χθ/ Initial	*/χϑesinos/*	I	5	/χ/ Initial, /θ/ Initial
Phonology	Phonemes	/fθ/ Initial	*/fϑano/*		5	/f/ Initial, /θ/ Initial
Phonology	Phonemes	/t/ Internal	*/kota/*	J	3	none
Phonology	Phonemes	/d/ Internal	*/kodos/*		3	none
Phonology	Phonemes	/p/ Internal	*/kapos/*		3	none
Phonology	Phonemes	/b/ Internal	*/kabos/*		3	none
Phonology	Phonemes	/k/ Internal	*/likos/*		2	/k/ Initial
Phonology	Phonemes	/m/ Internal	*/lemos/*	K	2	/m/ Initial
Phonology	Phonemes	/n/ Internal	*/ponos/*		2	/n/ Initial
Phonology	Phonemes	/θ/ Internal	*/paϑos/*	L	2	none
Phonology	Phonemes	/ð/ Internal	*/roða/*		2	none
Phonology	Phonemes	/f/ Internal	*/lofos/*		2	none
Phonology	Phonemes	/v/ Internal	*/ravo/*		4	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Phonemes	/χ/ Internal	*/paχos/*		4	none
Phonology	Phonemes	/γ/ Internal	*/paγos/*		4	none
Phonology	Phonemes	/s/ Internal	*/mesa/*	M	4	/s/ Initial
Phonology	Phonemes	/z/ Internal	*/vazo/*		4	/z/ Initial
Phonology	Phonemes	/l/ Internal	*/pali/*	N	4	/l/ Initial
Phonology	Phonemes	/r/ Internal	*/pera/*		3	none
Phonology	Phonemes	/kt/ Internal	*/ðiktis/*	O	3	/k/ Initial, /t/ Initial
Phonology	Phonemes	/pt/ Internal	*/lepto/*		3	/p/ Initial, /t/ Initial
Phonology	Phonemes	/ks/ Internal	*/leksi/*	p	3	/k/ Initial, /s/ Initial
Phonology	Phonemes	/ps/ Internal	*/lipsi/*		2	/p/ Initial, /s/ Initial
Phonology	Phonemes	/sk/ Internal	*/voskos/*		2	/k/ Initial, /s/ Initial
Phonology	Phonemes	/sp/ Internal	*/laspi/*		3	/p/ Initial, /s/ Initial
Phonology	Phonemes	/ðr/ Internal	*/foðra/*	Q	3	/ð/ Initial, /r/ Initial
Phonology	Phonemes	/θr/ Internal	*/vaϑro/*		4	/θ/ Initial, /r/ Initial
Phonology	Phonemes	/fr/ Internal	*/afros/*		4	/f/ Initial, /r/ Initial

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	Phonemes	/χr/ Internal	*/oχros/*		4	/χ/ Initial, /r/ Initial
Phonology	Phonemes	/χθ/ Internal	*/moxϑos/*	R	4	/θ/ Initial, /χ/ Initial
Phonology	Phonemes	/fθ/ Internal	*/afϑonos/*		5	/θ/ Initial, /f/ Initial
Phonology	Phonemes	εϋ: /eι/	*Σεϋχέλλες*		7	none
Phonology	Phonemes	οϋ: /oι/	*προϋπόθεση*		7	none
Phonology	Phonemes	αϊ: /aι/	*παϊδάκι*		7	none
Phonology	Phonemes	αη: /aι/	*καημένος*		7	none
Phonology	Phonemes	οΰ: /oi/	*αραχνοΰφαντος*		7	none
Phonology	Phonemes	αΐ: /ai/	*Μαΐου*		7	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -άκι, -άκης, -άκος	*αρκούδα-αρκουδάκι, Γιώργος-Γιωργάκης, δρόμος-δρομάκος*		1	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -ίτσα	*κοπέλα-κοπελίτσα*		1	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -οπούλα,-όπουλο,	*βοσκός-βοσκοπούλα, κότα-κοτόπουλο*	A	1	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -ούδι	*άγγελος-αγγελούδι*		1	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -ούλης/α/ούλικο,	*βάρκα-βαρκούλα, ϑεός-ϑεούλης, μικρούλης/α/ούλικο*		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: deriv.	NOUNS&ADJs: Diminutives: -ούτσικος/η/ούτσικο	*μικρούτσικος/η/ούτσικο*		1	none
Morphology	Suffixes: deriv.	NOUNS&ADJs: Enlargement: -άκλα,-άρα,-αράς,-αρόνα,-αρος,-ούκλα	*κόρη-κοράκλα, κοπέλα-κοπελάρα, σπίτι-σπιταρόνα, ψάρι-ψαρούκλα*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS&ADJs: quality/effect: -ίλα	*άσπρο-ασπρίλα*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N Suffixes -άς, -έας, -ιάς	*κλειδί-κλειδαράς, κουρεύω-*	B	2	Noun Suffix -ος, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			κουρέας, γράφω-γραφιάς			Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes-άρης,-ιάρης,-ιέρης/-ιέρα	*βάρκα-βαρκάρης, σκουπίδι-σκουπιδιάρης, πόρτα-πορτιέρης, κάμαρα-καμαριέρα*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -ίτης,-ιώτης	*τέχνη-τεχνίτης, ταξίδι-ταξιδιώτης*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						-ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -ίστας,-ιστής/-ίστρια	*πιάνο-πιανίστας/πιανίστρια, βιολί-βιολιστής*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -της/-τής/-τρια/-τισσα	*κλέβω-κλέφτης, προπονώ-προπονητής/προπονήτρια, αγρός-αγρότης/αγρότισσα*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -τζής/-τζού	*φορτηγό-φορτηγατζής, καφές-καφετζής/καφετζού*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -τίας,-τορας	*επάγγελμα-επαγγελματίας, εισπράττω-εισπράκτορας*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS: Place: -είο,-ιά,-ία,-	*φάρμακο-φαρμακείο, ποτάμι-*		2	Noun Suffix -ος, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
		ικο,-δικο	*ποταμιά, Γάλλος-Γαλλία, μπακάλης-μπακάλικο, ψαράς-ψαράδικο*			Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS: neuter : -ιλίκι.	*άντρας - αντριλίκι*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS:nationality or place or origin: -ανός, -αίος, -έζος, -ιος	*Μεξικό-Μεξικανός, Ευρώπη-Ευρωπαίος, Κίνα-Κινέζος, Κάρπαθος-Καρπάθιος*		3	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						-ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS: neuter from verbs: -μα	*διαβάζω-διάβασμα*		1	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	NOUNS: feminine : -άδα, -αινα,	*πορτοκάλι-πορτοκαλάδα, λύκος-λύκαινα*		3	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -τήρι,-τήριο,-τρα,-της	*ποτίζω-ποτιστήρι, πλένω-πλυντήριο, κρεμάω-κρεμάστρα, διακόπτω-διακόπτης*	C	2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -ιέρα, -έρα	*τσάι-τσαγέρα, ζάχαρη-ζαχαριέρα*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffix -ί	*κανέλα-κανελί, θάλασσα-*	D	2	Noun Suffix -ος, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			θαλασσί			Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffix -ιά	*κεράσι-κερασιά*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -άριο, -άρι ,-αριό	*βιβλίο-βιβλιάριο, σκουπίδι-σκουπιδαριό*		3	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						-ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -εία,-ειά,-εια	*θεραπεύω-θεραπεία, δουλεύω-δουλειά, προσπαθώ-προσπάθεια,*	E	2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(σ/ξ/ψ)η	*πλένω-πλύση, πλέκω-πλέξη, χωνεύω-χώνεψη*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(σ)ία	*αποτυγχάνω-αποτυχία, εργάζομαι-εργασία*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(σ/ψ/ξ/ματ)ιά	*δάχτυλο-δαχτυλιά, κλέβω-κλεψιά, δαγκώνω-δαγκωματιά*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(σ/ξ/ψ)ιμο	*δένω-δέσιμο, γράφω-*		2	Noun Suffix -ος, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			γράψιμο, τρέχω-τρέξιμο			Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(α/η/ω/σ/γ/)μα, -μός	*ζεσταίνω-ζέσταμα, βοηθώ-βοήθημα, τελειώνω-τελείωμα, βαδίζω-βάδισμα, ανοίγω-άνοιγμα, χάνω-χαμός*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	N suffixes -(η/α/κ/χ/φ/π)το	*βογκάω-βογκητό, κατεβαίνω-κατεβατό, πλέκω-πλεκτό, γράφο-γραφτό/γραπτό*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						-ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ικός/-ική/-ικό	*δήμος-δημοτικός/ή/ό*	F	2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ερός/-ερή/-ερό	*βροχή-βροχερός/ή/ό*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ένιος/-ένια/-ένιο, -ιάρης/-ιάρα/άρικο	*σοκολάτα-σοκολατένιος/α/ο, άρρωστος-αρρωστιάρης/α/άρικο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -τός/-τή/-τό, -άτος/-άτη/-άτο	*φουσκώνω-φουσκωτός/ή/ό, φεύγω-φευγάτος/η/ο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ινος/-ινη/-ινο,	*ξύλο-ξύλινος/η/ο, δίπλα-*		2	Noun Suffix -ος, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
		-ανός/-ανή/-ανό	*διπλανός/ή/ό*			Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ούρης/-ούρα/-ούρικο	*μουρμουράω-μουρμούρης/α/ο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ίστικος/-ίστικη/-ίστικο	*αγόρι-αγορίστικος/η/ο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						-ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ήσιος/-ήσια/-ήσιο, -ίσιος/-ίσια/-ίσιο	*ημερήσιος/α/ο, σπίτι-σπιτήσιος/α/ο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -λέος/-λέα/-λέο,-αίος/-αία/-αίο	*φεύγω-φευγαλέος/α/ο, μήνας-μηνιαίος/α/ο*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
						Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	Adj suffixes -ένιος/α/ο	*μολύβι-μολυβένιος*		2	Noun Suffix -ος, Noun Suffix -ας, Noun Suffix -ης, Noun Suffix -α, Noun Suffix -η, Noun Suffix -ο, Noun Suffix (ending) -ι, Noun Suffix -ο, Noun Suffix -α, Noun Suffix -η, Noun Suffix -οι, Noun Suffix -ες, Noun Suffix -ες, Noun Suffix -εις, Noun Suffix -α,Noun Suffix (ending) -ια
Morphology	Suffixes: deriv.	V suffixes -ίζω, -άζω, -ιάζω	*κακός-κακίζω, στέγη-στεγάζω, ξαφνικά-ξαφνιάζω*		3	none
Morphology	Suffixes: deriv.	V suffixes -αίνω, -ώνω, -ύνω	*βάθος-βαθαίνω, δύναμη-δυναμώνω, μέγεθος-μεγεθύνω*	G	3	none
Morphology	Suffixes: deriv.	V suffixes -εύω, -άρω	*χορός-χορεύω, πρόβα-προβάρω*		3	none
Morphology	Suffixes: deriv.	VERBS: lexical suffixes: -βολώ,-λογώ,-ποιώ	*πέτρα-πετροβολώ, κακός-κακολογώ, δράμα-δραματοποιώ*		4	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: deriv.	ADJs: lexical suffixes: -ειδής/-ειδές, -μελής/-μελές, -ετής/-ετές	*αράχνη-αραχνοειδής, τριμελής/πολυμελής/μονομελής..., πρωτοετής/τριετής... Κλπ*		5	none
Morphology	Suffixes: infl.	Noun Suffix -ος	*άνθρωπος*	A	1	none
Morphology	Suffixes: infl.	Noun Suffix -ας	*ταμίας*		1	none
Morphology	Suffixes: infl.	Noun Suffix -ης	*μαθητής*		1	none
Morphology	Suffixes: infl.	Noun Suffix -α	*πόρτα*		1	none
Morphology	Suffixes: infl.	Noun Suffix -η	*τάξη*		1	none
Morphology	Suffixes: infl.	Noun Suffix -ο	*δένδρο*		1	none
Morphology	Suffixes: infl.	Noun Suffix (ending) -ι	*παιδί*		1	none
Morphology	Suffixes: infl.	Noun Suffix -ο	*άνθρωπο*		1	none
Morphology	Suffixes: infl.	Noun Suffix -α	*ταμία*		1	none
Morphology	Suffixes: infl.	Noun Suffix -η	*μαθητή*		1	none
Morphology	Suffixes: infl.	Noun Suffix -οι	*άνθρωποι*	B	1	none
Morphology	Suffixes: infl.	Noun Suffix -ες	*ταμίες, μαθητές*		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: infl.	Noun Suffix -ες	*πόρτες*		1	none
Morphology	Suffixes: infl.	Noun Suffix -εις	*τάξεις*		1	none
Morphology	Suffixes: infl.	Noun Suffix -α.	*δένδρα*		1	none
Morphology	Suffixes: infl.	Noun Suffix (ending) -ια	*παιδιά*		1	none
Morphology	Suffixes: infl.	Noun Suffix -ου	*ανθρώπου*	C	2	Noun Suffix -ος,Noun Suffix -ο
Morphology	Suffixes: infl.	Noun Suffix -ων/-ών	*ανθρώπων, ταμίων, μαθητών*		2	Noun Suffix -οι, Noun Suffix -ες
Morphology	Suffixes: infl.	Noun Suffix -ας	*πόρτας*		2	Noun Suffix -α
Morphology	Suffixes: infl.	Noun Suffix -ης	*τάξης*		2	Noun Suffix -η
Morphology	Suffixes: infl.	Noun Suffix -ών	*πορτών*		2	Noun Suffix -ες
Morphology	Suffixes: infl.	Noun Suffix -εων	*τάξεων*		2	Noun Suffix -εις
Morphology	Suffixes: infl.	Noun Suffix -ου/-ού	*δένδρου, ματιού*		2	Noun Suffix -ο, Noun Suffix (ending) -ι
Morphology	Suffixes: infl.	Noun Suffix -ων/-ών	*δένδρων, ματιών*		2	Noun Suffix -α., Noun Suffix (ending) -ια
Morphology	Suffixes: infl.	Noun Suffix -ές, -άς, -ούς	*καναπές, παπάς, παππούς*	D	3	none
Morphology	Suffixes: infl.	Noun Suffix -έ, -ά, -ού	*καναπέ, παπά, παππού*		3	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: infl.	Noun Suffix -ού	*αλεπού*		3	none
Morphology	Suffixes: infl.	Noun Suffix -ος	*δάσος*		3	none
Morphology	Suffixes: infl.	Noun Suffix -ας	*κρέας*		3	none
Morphology	Suffixes: infl.	Noun Suffix (ending) -μα	*γράμμα*		3	none
Morphology	Suffixes: infl.	Noun Suffix -υ	*δόρυ*		3	none
Morphology	Suffixes: infl.	Noun Suffix -δες	*καναπέδες, παπάδες, παππούδες*	E	3	none
Morphology	Suffixes: infl.	Noun Suffix -δες	*αλεπούδες*		3	none
Morphology	Suffixes: infl.	Noun Suffix -οντα	*καθήκοντα*		3	none
Morphology	Suffixes: infl.	Noun Suffix -ατα	*κρέατα, γράμματα, δόρατα*		3	none
Morphology	Suffixes: infl.	Noun Suffix -έδων, -άδων, -ούδων	*καναπέδων, παπάδων, παππούδων*	F	4	Noun Suffix -ές, -άς, -ούς, Noun Suffix -έ, -ά, -ού, Noun Suffix -δες
Morphology	Suffixes: infl.	Noun Suffix -ούς	*αλεπούς*		4	Noun Suffix -ού
Morphology	Suffixes: infl.	Noun Suffix -ούδων	*αλεπούδων*		4	Noun Suffix -δες, Noun Suffix -ού
Morphology	Suffixes: infl.	Noun Suffix -ατος	*κρέατος, γράμματος, δόρατος*		4	Noun Suffix -ας,Noun Suffix (ending) -μα, Noun Suffix -υ
Morphology	Suffixes: infl.	Noun Suffix -άτων	*κρεάτων, γραμμάτων,*		4	Noun Suffix -ατα, Noun Suffix -ας,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			δοράτων			Noun Suffix (ending) -μα, Noun Suffix -υ
Morphology	Suffixes: infl.	Adj Suffixes -ος/-η/-ο	*μεγάλος/-η/-ο*	G	2	Noun Suffix -ος, Noun Suffix -η, Noun Suffix -ο
Morphology	Suffixes: infl.	Adj Suffixes -ος/-α/-ο	*πλούσιος/-α/-ο*	G	2	Noun Suffix -ος, Noun Suffix -α, Noun Suffix -ο
Morphology	Suffixes: infl.	Adj Suffixes -ός/-ιά/-ό	*κακός/-ιά/-ό*	G	2	Noun Suffix -ος, Noun Suffix -α, Noun Suffix -ο
Morphology	Suffixes: infl.	Adj Suffixes -ης/-α/-(ικ)ο	*γκρινιάρης/-α/-ικο*	G	2	Noun Suffix -ος, Noun Suffix -α, Noun Suffix -ο,NOUNS&ADJs: Diminutives: -ούλης/α/ούλικο,
Morphology	Suffixes: infl.	Adj Suffixes -άς/-ού/-άδικο	*υπναράς/-ού/-ούδικο*	G	2	Noun Suffix -ας, Noun Suffix -ού, Noun Suffix -ο,NOUNS&ADJs: Diminutives: -ούλης/α/ούλικο,
Morphology	Suffixes: infl.	Adj Suffixes -ύς/-ιά/-ύ	*βαθύς/-ιά/-ύ*	H	4	Noun Suffix -υ, Noun Suffix -α
Morphology	Suffixes: infl.	Adj Suffixes -ής/-ιά/-ί	*δεξής/-ιά-/ί*	H	4	Noun Suffix -ης, Noun Suffix -α, Noun Suffix (ending) -ι
Morphology	Suffixes: infl.	Adj Suffixes -ής/-ές	*διεθνής/-ές*	H	4	Noun Suffix -ης
Morphology	Suffixes: infl.	πολύς/πολλή/πολύ	*πολύς/πολλή/πολύ*	H	4	Noun Suffix -η
Morphology	Suffixes: infl.	Adj Suffixes -οι/-ες/-α	*μεγάλοι/-ες/-α, πλούσιοι/-*	I	2	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			ες/-α, κακοί/-ες/-α			
Morphology	Suffixes: infl.	Adj Suffixes -ηδες/-ες/-ικα	*γκρινιάρηδες/-ες/-ικα*		2	none
Morphology	Suffixes: infl.	Adj Suffixes -άδες/-ούδες/-άδικα	*υπναράδες/-ούδες/-ούδικα*		2	none
Morphology	Suffixes: infl.	Adj Suffixes -ιοί/-ιές/-ιά	*βαθιοί/-ιές/-ιά, δεξιοί/-ιές/-ιά*	J	5	Adj Suffixes -ύς/-ιά/-ύ, Adj Suffixes -ής/-ιά/-ί
Morphology	Suffixes: infl.	Adj Suffixes -είς/-ή	*διεθνείς/-ή*		5	Adj Suffixes -ής/-ές,
Morphology	Suffixes: infl.	πολλοί/-ές/-ά	*πολλοί/πολλές/πολλά*		5	πολύς/πολλή/πολύ
Morphology	Suffixes: infl.	Adj Suffixes -ου/-ης/-ου, -ου/-ας/-ου, -ού/-ιάς/-ού	*μεγάλου/-ης/-ου, πλούσιου/-ας/-ου, κακού/-άς/-ού*	K	3	Adj Suffixes -ος/-η/-ο
Morphology	Suffixes: infl.	Adj Suffixes -η/-ας/-(ικ)ου	*γκρινιάρη/-ας/-ικου*		3	Adj Suffixes -ης/-α/-(ικ)ο
Morphology	Suffixes: infl.	Adj Suffixes -ά/-ούς/-άδικου	*υπναρά/-ούς/-ούδικου*		3	Adj Suffixes -άς/-ού/-άδικο
Morphology	Suffixes: infl.	Adj Suffixes -ων, -ών	*μεγάλων, πλούσιων, κακών*	L	3	Adj Suffixes -οι/-ες/-α
Morphology	Suffixes: infl.	Adj Suffixes -ηδων/-ων/-ικων,	*γκρινιάρηδων/-ων/-ρικων*		3	Adj Suffixes -ηδες/-ες/-ικα
Morphology	Suffixes: infl.	Adj Suffixes -άδων/-ούδων/-ούδικων	*υπναράδων/-ούδων/-ούδικων*		3	Adj Suffixes -άδες/-ούδες/-άδικα

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: infl.	Adj Suffixes -ιού/-ιάς/-ιού	*βαθιού/-ιάς/-ιού, δεξιού/-ιάς/-ιού*	M	4	Adj Suffixes -ύς/-ιά/-ύ, Adj Suffixes -ής/-ιά/-ί
Morphology	Suffixes: infl.	Suffix -ούς	*διεθνούς*		4	Adj Suffixes -ής/-ές
Morphology	Suffixes: infl.	Adj Suffixes -ιών, -ών	*βαθιών, δεξιών, διεθνών*		4	Adj Suffixes -ιοί/-ιές/-ιά
Morphology	Suffixes: infl.	Adj Suffixes -ιού/-ιάς/-ιού/-ιών, -ούς/-ών	*βαθιού/-ιάς/-ιού/-ιών, δεξιού/-ιάς/-ιού/-ιών, διεθνούς/-ών*		4	Adj Suffixes -ύς/-ιά/-ύ,Adj Suffixes -ής/-ιά/-ί, Adj Suffixes -ής/-ές
Morphology	Suffixes: infl.	Verbal Suffixes -ω/-εις/-ει/-ουμε/-ετε/-ουν	*ντύνω/-εις/-ει/-ουμε/-ετε/-ουν*	N	3	none
Morphology	Suffixes: infl.	Verbal Suffixes -ώ/-άς/-ά(ει)/-άμε/-ούμε/-άτε/-ούν(ε)/-άν(ε)/ -ούν	*αγαπώ/-άς/-ά(ει)/-άμε/-άτε/-άν(ε)/-ούν*		3	none
Morphology	Suffixes: infl.	Verbal Suffixes -ώ/-είς/-εί/-ούμε/-είτε/-ούν(ε)	*θεωρώ/-είς/-εί/-ούμε/-είτε/-ούν*		3	none
Morphology	Suffixes: infl.	Verbal Suffixes -α/-ες/-ε/-αμε/-ατε/-αν(ε)	*έντυνα/-ες/-ε/ντύναμε/-ατε/έντυναν, έντυσα/-ες/-ε/-ντύσαμε/-ατε/-αν*	O	4	Verbal Suffixes -ω/-εις/-ει/-ουμε/-ετε/-ουν
Morphology	Suffixes: infl.	Verbal Suffixes -αγα/-αγες/-αγε/-άγαμε/-άγατε/-αγαν(ε)	*αγάπαγα/-αγες/-αγε/-άγαμε/-άγατε/-αγαν*		4	Verbal Suffixes -ώ/-άς/-ά(ει)/-άμε/-ούμε/-άτε/-ούν(ε)/-άν(ε)/ -ούν

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: infl.	Verbal Suffixes -ούσα/-ούσες/-ούσε/-ούσαμε/-ούσατε/-ούσαν(ε)	*αγαπούσα/-ες/-ε/-ούσαμε/-ούσατε/-ούσαν(ε)*		4	Verbal Suffixes -ώ/-άς/-ά(ει)/-άμε/-ούμε/-άτε/-ούν(ε)/-άν(ε)/ -ούν
Morphology	Suffixes: infl.	Verbal Suffixes -ησα/-ησες/-ησε/-ήσαμε/-ήσατε/-ησαν	*αγάπησα/-ες/-ε/-ήσαμε/-ήσατε/αν*		4	Verbal Suffixes -ώ/-άς/-ά(ει)/-άμε/-ούμε/-άτε/-ούν(ε)/-άν(ε)/ -ούν
Morphology	Suffixes: infl.	Verbal Suffixes verbs -ομαι/-εσαι/-εται/-όμαστε/-όσαστε/-εστε/-ονται	*ντύνομαι/-εσαι/-εται/-όμαστε/-όσαστε/-εστε/-ονται*	P	5	Verbal Suffixes -ω/-εις/-ει/-ουμε/-ετε/-ουν
Morphology	Suffixes: infl.	Verbal Suffixes verbs -ιέμαι/-ιέσαι/-ιέται/-ιόμαστε/-ιέστε/-ιούνται/-ιόνται	*αγαπιέμαι/-ιέσαι/-ιέται/-ιόμαστε/-ιέστε/-ιούνται*		5	Verbal Suffixes -ώ/-άς/-ά(ει)/-άμε/-ούμε/-άτε/-ούν(ε)/-άν(ε)/ -ούν
Morphology	Suffixes: infl.	Verbal Suffixes verbs -ιούμαι/-είσαι/-είται/-ούμαστε/-είστε/-ούνται,	*θεωρούμαι/-είσαι/-είται/-ούμαστε/-είστε/-ούνται*		5	Verbal Suffixes -ώ/-είς/-εί/-ούμε/-είτε/-ούν(ε)
Morphology	Suffixes: infl.	Verbal suffixes -όμουν(α)/-όσουν(α)/-όταν(ε)/-όμασταν/-όμαστε/-όσασταν/-όσαστε/-ονταν(ε)/-όντουσαν	*ντυνόμουν/-όσουν/-όταν/-όμασταν(-ε)/-όσασταν(-ε)/-ντύνονταν/-όντανε/-όντουσαν*	Q	6	Verbal Suffixes verbs -ομαι/-εσαι/-εται/-όμαστε/-όσαστε/-εστε/-ονται, Verbal Suffixes -α/-ες/-ε/-αμε/-ατε/-αν(ε)

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Suffixes: infl.	Verbal suffixes -όμουν(α)/-όσουν(α)/-όταν/-όμασταν/-όμαστε/-όσασταν/-όσαστε/-όντουσαν	*αγαπιόμουν(α)/-ιόσουν(α)/-ιόταν(ε)/-ιόμασταν(-ιόσαστε)/-ιόσασταν(-ιόσαστε)/-ιούνταν(ε)/-ιόνταν(ε)/-ιόντουσαν*		6	Verbal Suffixes verbs -ιέμαι/-ιέσαι/-ιέται/-ιόμαστε/-ιέστε/-ιούνται/-ιόνται, Verbal Suffixes -αγα/-αγες/-αγε/-άγαμε/-άγατε/-αγαν(ε), Verbal Suffixes -ούσα/-ούσες/-ούσε/-ούσαμε/-ούσατε/-ούσαν(ε)
Morphology	Suffixes: infl.	Verbal suffixes -ούμουν(α)/-ούσουν(α)/-ούταν/-ούμασταν/-ούμαστε/-ούσασταν/-ούσαστε/-ούνταν(ε)	*θεωρούμουν(α)/-ούσουν(α)/-ούνταν/-ούμασταν(-ούμαστε)/-ούσασταν(-ούσαστε)/-ούνταν(ε)*		6	Verbal Suffixes verbs -ιούμαι/-είσαι/-είται/-ούμαστε/-είστε/-ούνται,
Morphology	Suffixes: infl.	Verbal suffixes -θηκα/-θηκες/-θηκε/-θήκαμε/-θήκατε/-θηκαν	*ντύθηκα/-ες/-ε/..., αγαπήθηκα/-ες/-ε/..., θεωρήθηκα/-ες/-ε/...*		6	Verbal Suffixes verbs -ομαι/-εσαι/-εται/-όμαστε/-όσαστε/-εστε/-ονται
Morphology	Suffixes: infl.	adjectival participles: -ών/-ούσα/-όν, -είς/-είσα/-έν	*παρών/-ούσα/-όν/-όντος/-ούσας/-όντες/-όντων,προαχθείς/-είσα/-έν/-έντος/προαχθέντες/-είσες/-έντα/-έντων/-εισών*		7	none
Morphology	Prefixing	ADJS: Prefix α-	*συνεπής-ασυνεπής*	A	1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Morphology	Prefixing	ADJS: Prefix αντι-	*λαϊκός-αντιλαϊκός*		1	none
Morphology	Prefixing	ADJS: Prefix δυσ-	*εύκολος-δύσκολος*		1	none
Morphology	Prefixing	ADJS&NOUNS: Prefix υπερ-	*υπεραρκετός, υπερκόπωση*	B	2	ADJS: Prefix α-, ADJS: Prefix αντι-, ADJS: Prefix δυσ-
Morphology	Prefixing	ADJS&NOUNS: Prefix υπο-	*υπογλυκαιμία, υποτονικός*		2	ADJS: Prefix α-, ADJS: Prefix αντι-, ADJS: Prefix δυσ-
Morphology	Prefixing	ADJS&NOUNS: Prefix κατα-	*κατάμαυρος, καταλάθος,*		2	ADJS: Prefix α-, ADJS: Prefix αντι-, ADJS: Prefix δυσ-
Morphology	Prefixing	VERBS: Prefix υπέρ-	*υπερβάλλω*	C	2	ADJS&NOUNS: Prefix υπερ-
Morphology	Prefixing	VERBS: Prefix υπό-	*υποτάσσω*		2	ADJS&NOUNS: Prefix υπο-
Morphology	Prefixing	VERBS: Prefix κατα-	*καταβάλλω*		2	ADJS&NOUNS: Prefix κατα-
Morphology	Prefixing	Lexical prefixes δι-/τρι-/τετρα-	*δίτροχος, τετράτροχος*	D	2	none
Morphology	Prefixing	Lexical prefixes πρωτο-,αυτο-,πολυ-,μικρο-,ψιλο-,ημι-	*πρωτόγνωρος, πολυετής, μικροπρεπής, ψιλοάγουρος, ημίτρελλος*		2	none
Morphology	Prefixing	Verbal prefix ανα-, απο-, εισ-,επι-,μετα-,παρα-,περι-,προ-,προσ-,συν-	*αναβάλλω, αποβάλλω, εισβάλλω, επιβάλλω, μεταβάλλω, παραβάλλω, περιβάλλω, προβάλλω,*		2	VERBS: Lexical prefixes: ψιλο-,μισο-,κουτσο-,ψευτο-,

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			προσβάλλω, συμβάλλω			
Morphology	Prefixing	VERBS: Lexical prefixes: ψιλο-,μισο-,κουτσο- ,ψευτο-	*ψιλοβλέπω, μισογεμίζω, κουτσοβλέπω, ψευτογελάω*		1	none
Phonology	GPC	initial: σπ - /sp/	*σπίτι*	A	1	none
Phonology	GPC	initial: στ - /st/	*στέκα*		1	none
Phonology	GPC	initial: σκ - /sk/	*σκεπή*		1	none
Phonology	GPC	initial: πρ - /pr/	*πρόκα*	B	1	none
Phonology	GPC	initial: τρ - /tr/	*τρένο*		1	none
Phonology	GPC	initial: κρ - /kr/	*κρόσι*		1	none
Phonology	GPC	initial: πτ - /pt/	*πτώμα*	C	1	none
Phonology	GPC	initial: κτ - /kt/	*κτήμα*		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	initial: φτ - /ft/	*φτύνω*		1	none
Phonology	GPC	initial: φθ - /fθ/	*φθάνω*	D	2	none
Phonology	GPC	initial: σθ - /sθ/	*σθένος*		2	none
Phonology	GPC	initial: χθ - /xθ/	*χθόνιος*		2	none
Phonology	GPC	initial: σφ - /sf/	*σφαίρα*		2	none
Phonology	GPC	initial: σβ - /zv/	*σβήνω*	E	2	none
Phonology	GPC	initial: σγ - /zγ/	*σγουρός*		2	none
Phonology	GPC	initial: βδ - /vð/	*βδέλα*	F	2	none
Phonology	GPC	initial: γδ - /γð/	*γδύνω*		2	none
Phonology	GPC	initial: βγ - /vγ/	*βγάζω*		2	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	initial: σμ - /zm/	*σμήνος*	G	3	none
Phonology	GPC	initial: φρ - /fr/	*φράγμα*	H	3	none
Phonology	GPC	initial: δρ - /ðr/	*δρόμος*		3	none
Phonology	GPC	initial: χρ - /χr/	*χρυσό*		3	none
Phonology	GPC	initial: βρ - /vr/	*βροχή*		3	none
Phonology	GPC	initial: γρ - /γr/	*γράμμα*		3	none
Phonology	GPC	initial: θρ - /θr/	*ϑράσος*		3	none
Phonology	GPC	initial: πλ - /pl/	*πλένω*	I	2	none
Phonology	GPC	initial: κλ - /kl/	*κλείνω*		2	none
Phonology	GPC	initial: φλ - /fl/	*φλόγα,*		2	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	initial: θλ - /θl/	*θλάση*		2	none
Phonology	GPC	initial: σλ - /sl/	*σλάβος*		2	none
Phonology	GPC	initial: χθ - /xl/	*χλωρός*		2	none
Phonology	GPC	initial: βλ - /vl/	*βλέπω*		2	none
Phonology	GPC	initial: γλ - /ɣl/	*γλάρος*		2	none
Phonology	GPC	initial: πν - /pn/	*πνοή*		3	none
Phonology	GPC	initial: κν - /kn/	*κνήμη*		3	none
Phonology	GPC	initial: θν - /θn/	*θνητός*	J	3	none
Phonology	GPC	initial: σν - /sn/	*σνίτσελ*		3	none
Phonology	GPC	initial: χν - /xn/	*χνώτο*		3	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	initial: γν - /ɣn/	*γνέφω*		3	none
Phonology	GPC	initial: τμ - /tm/	*τμήμα*		3	none
Phonology	GPC	initial: σπρ - /spr/	*σπρώχνω*	K	3	initial: /sp/, initial: /pr/
Phonology	GPC	initial: σκρ - /skr/	*σκράπα*		3	initial: /sk/, initial: /kr/
Phonology	GPC	initial: στρ - /str/	*στράτα*		3	initial: /st/, initial: /tr/
Phonology	GPC	initial: σφρ - /sfr/	*σφρίγος*		3	initial:/sf/
Phonology	GPC	internal: σπ - /sp/	*λάσπη*	L	3	initial: /sp/
Phonology	GPC	internal: στ - /st/	*λίστα*		3	initial: /st/
Phonology	GPC	internal: σκ - /sk/	*ασκί*		3	initial: /sk/
Phonology	GPC	internal: πρ - /pr/	*απροετοίμαστος*	M	3	initial: /pr/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	internal: τρ - /tr/	*κίτρο*		3	initial: /tr/
Phonology	GPC	internal: κρ - /kr/	*άκρη*		3	initial: /kr/
Phonology	GPC	internal: πτ - /pt/	*απτός*	N	3	initial: /pt/
Phonology	GPC	internal: κτ - /kt/	*ακτή*		3	initial: /kt/
Phonology	GPC	internal: φτ - /ft/	*κοφτός*		3	initial: /ft/
Phonology	GPC	internal: φθ - /fθ/	*αφθονος*	O	4	initial: /fθ/
Phonology	GPC	internal: σθ - /sθ/	*ασθενής*		4	initial: /sθ/
Phonology	GPC	internal: χθ - /xθ/	*απεχθής*		4	initial: /xθ/
Phonology	GPC	internal σφ - /sf/	*ασφαλής*		4	initial:/sf/
Phonology	GPC	internal: σβ - /zv/	*ασβός*	P	4	initial: /zv/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	internal: βδ - /vð/	*ραβδί*	Q	4	initial: /vð/
Phonology	GPC	internal: γδ - /γð/	*ραγδαίος*		4	initial: /γð/
Phonology	GPC	internal: βγ - /vγ/	*αβγό*		4	initial: /vγ/
Phonology	GPC	internal: φρ - /fr/	*αφρός*	R	4	initial: /fr/
Phonology	GPC	internal: δρ - /ðr/	*αδρός*		4	initial: /ðr/
Phonology	GPC	internal: χρ - /χr/	*ωχρός*		4	initial: /χr/
Phonology	GPC	internal: βρ - /vr/	*ταύρος*		4	initial: /vr/
Phonology	GPC	internal: γρ - /γr/	*αγρός*		4	initial: /γr/
Phonology	GPC	internal: θρ - /θr/	*άϑροισμα*		4	initial: /θr/
Phonology	GPC	Internal, πλ - /pl/	*απλός*	S	4	initial: /pl/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	Internal κλ - /kl/	*πίκλα*		4	initial: /kl/
Phonology	GPC	Internal φλ - /fl/	*έφλεκτο*		4	initial: /fl/
Phonology	GPC	Internal θλ - /θl/	*αθλητής*		4	initial: /θl/
Phonology	GPC	Internal, σλ - /sl/	*Όσλο*		4	initial: /sl/
Phonology	GPC	Internal, χλ - /xl/	*μούχλα*		4	initial: /xl/
Phonology	GPC	Internal βλ - /vl/	*αυλή*		4	initial:/vl/
Phonology	GPC	Internal γλ - /ɣl/	*αίγλη*		4	initial: /ɣl/
Phonology	GPC	Internal, πν - /pn/	*αναπνοή*	T	4	initial: /pn/
Phonology	GPC	Internal κν - /kn/	*οκνηρός*		4	initial: /kn/
Phonology	GPC	Internal χν - /xn/	*αχνός*		4	initial: /θn/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	Internal γν - /ɣn/	*άγνοια*		4	initial: /sn/
Phonology	GPC	Internal, τμ - /tm/	*ατμός*	U	4	initial:/xn/
Phonology	GPC	Internal σμ - /zm/	*οσμή*		4	initial:/ɣn/
Phonology	GPC	internal: σπρ - /spr/	*άσπρο*	V	5	initial: /tm/
Phonology	GPC	internal: στρ - /str/	*άστρο*		5	initial: /zm/
Phonology	GPC	internal: σφρ - /sfr/	*όσφρηση*		5	initial:/sfr/
Phonology	GPC	internal: κστρ - /kstr/	*εκστρατεία*	W	5	internal: /str/), internal: /ks/
Phonology	GPC	internal: ξπρ - /kspr/	*εξπρές*		6	internal: /spr/, internal: /ks/
Phonology	GPC	initial: γκρ - /gr/	*γκρεμός*	X	5	initial: /kr/, initial: /γr/
Phonology	GPC	initial: ντρ - /dr/	*ντροπή*		5	initial: /tr/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	initial: μπρ - /br/	*μπρίκι*		5	initial: /pr/
Phonology	GPC	internal: γκρ - /gr/	*φαγκρί*	Y	5	initial: /gr/
Phonology	GPC	internal: ντρ - /dr/	*άντρας*		5	initial: /dr/
Phonology	GPC	internal: μπρ - /br/	*λαμπρή*		5	initial: /br/
Phonology	GPC	initial: ψ - /ps/	*ψέμα*	Z	5	none
Phonology	GPC	initial: ξ - /ks/	*ξανά*		5	none
Phonology	GPC	internal: ψ - /ps/	*αψίδα*	Zi	5	initial: /ps/
Phonology	GPC	internal: ξ - /ks/	*μετάξι*		5	initial: /ks/
Phonology	GPC	ευ-/ev/	*μαζεύω*	Zii	6	internal: /e/, internal: /v/
Phonology	GPC	ευ-/ef/	*ευχή*		6	initial: /e/, initial: /f/

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Phonology	GPC	αυ-/av/	*αυλή*	Ziii	6	initial: /a/, initial: /v/
Phonology	GPC	αυ-/af/	*αυτός*		6	initial: /a/, initial: /f/
Phonology	GPC	εϋ-/eɪ/	*Σεϋχέλλες*		6	none
Phonology	GPC	οϋ-/oɪ/	*προϋπόθεση*		6	none
Phonology	GPC	αϊ-/aɪ/	*παϊδάκι*		6	none
Phonology	GPC	οΰ-/oi/	*αραχνοΰφαντος*		6	none
Phonology	GPC	αΐ-/ai/	*Μαΐου*		6	none
Phonology	GPC	ια-[ja]	*ψάρια [ˈpsarja]*	Ziv	6	none
Phonology	GPC	ια-[ça]	*παθιασμένος [paθ*çaˈzmenos]		6	none
Syntax	Grammar: function words	*ενός*	*Η τσάντα ενός μαθητή.*	A	1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Grammar: function words	*μιας*	*Η τσάντα μιας μαθήτριας.*		1	none
Syntax	Grammar: function words	*ο*	*ο μαθητής*	B	2	none
Syntax	Grammar: function words	*η*	*η μαθήτρια*		2	none
Syntax	Grammar: function words	*το*	*το παιδί*		2	none
Syntax	Grammar: function words	*του [+N]*	*Το βιβλίο του παιδιού.*	C	3	ο , το
Syntax	Grammar: function words	*της [+N]*	*Το βιβλίο της μαθήτριας.*		3	η
Syntax	Grammar: function words	*οι*	*οι μαθητές*	D	2	ο,η, το
Syntax	Grammar: function words	*τα*	*Τα παιδιά*		2	το
Syntax	Grammar: function words	*των*	*Η τσάντα των δασκάλων.*	E	3	ο, η, οι
Syntax	Grammar: function words	Prepositions: σε	*στο τραπέζι*	F	2	none
Syntax	Grammar:	prepositions: με	*έφυγε με τρένο*		2	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	function words					
Syntax	Grammar: function words	prepositions: για	*έφυγε για Γερμανία*		2	none
Syntax	Grammar: function words	prepositions: από	*από χτες*		2	none
Syntax	Grammar: function words	μέσα σε, πάνω σε, δίπλα σε, κοντά σε		G	3	Prepositions: σε
Syntax	Grammar: function words	πριν από, μετά από, ύστερα από		G	3	prepositions: από
Syntax	Grammar: function words	κάτω από, πάνω από, μέσα από, πίσω από, μπροστά από		G	3	prepositions: από, Prepositions: σε
Syntax	Grammar: function words	μαζί με		G	3	prepositions: με
Syntax	Grammar: function words	αντί για		G	3	prepositions: για
Syntax	Grammar: function words	*[N+] μου*	*το παιδί μου*	H	4	none
Syntax	Grammar: function words	*[N+] σου*	*το παιδί σου*	H	4	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Grammar: function words	*[N+] του*	*το παιδί του*		4	none
Syntax	Grammar: function words	*[N+] της*	*το παιδί της*		4	none
Syntax	Grammar: function words	*[N+] μας*	*το παιδί μας*		4	none
Syntax	Grammar: function words	*[N+] σας*	*το παιδί σας*		4	none
Syntax	Grammar: function words	*[N+] τους*	*το παιδί τους*		4	none
Syntax	Grammar: function words	*μου [+V], σου [+V]*	*μου/σου έστειλαν το γράμμα*	I	5	[N+] μου,[N+] σου
Syntax	Grammar: function words	*του [+V]*	*του έστειλαν το γράμμα*		5	[N+] του
Syntax	Grammar: function words	*της [+V]*	*της έστειλαν το γράμμα*		5	[N+] του
Syntax	Grammar: function words	*μας [+V], σας [+V]*	*μας/σας έστειλαν το γράμμα*		5	[N+] μου,[N+] σου
Syntax	Grammar: function words	*τους [+V]*	*τους έστειλαν το γράμμα*		5	[N+] του
Syntax	Grammar:	*με [+V], σε [+V]*	*Ο Γιάννης με/σε φίλησε*	J	6	μου [+V], σου [+V]

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	function words					
Syntax	Grammar: function words	*τον [+V]*	*Ο Γιάννης τον φίλησε*		6	του [+V]
Syntax	Grammar: function words	*την [+V]*	*Ο Γιάννης την φίλησε*		6	της [+V]
Syntax	Grammar: function words	*το [+V]*	*Ο Γιάννης το φίλησε*		6	του [+V]
Syntax	Grammar: function words	*μας [+V], σας [+V]*	*Ο Γιάννης μας/σας φίλησε*		6	[N+] μας, [N+] σας
Syntax	Grammar: function words	*τους [+V]*	*Ο Γιάννης τους φίλησε*		6	*του [+V]*
Syntax	Grammar: function words	*τις [+V]*	*Ο Γιάννης τις φίλησε*		6	*της [+V]*
Syntax	Grammar: function words	*τα [+V]*	*Ο Γιάννης τα φίλησε*		6	του [+V]
Syntax	Grammar: function words	*μην*	*μην*	K	6	none
Syntax	Grammar: function words	*δεν*	*δεν*		6	none
Orphography	Letter similarity	α		L	1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	(visual)					
Orphography	Letter similarity (visual)	ο			1	none
Orphography	Letter similarity (visual)	φ		M	1	none
Orphography	Letter similarity (visual)	β			1	none
Orphography	Letter similarity (visual)	θ			1	none
Orphography	Letter similarity (visual)	ψ		N	1	none
Orphography	Letter similarity (visual)	ω			1	none
Orphography	Letter similarity	3		O	1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	(visual)					
Orphography	Letter similarity (visual)	ε			1	none
Orphography	Letter similarity (visual)	ξ			1	none
Orphography	Letter similarity (visual)	κ		P	1	none
Orphography	Letter similarity (visual)	χ			1	none
Orphography	Letter similarity (visual)	γ			1	none
Orphography	Letter similarity (visual)	λ			1	none
Orphography	Letter similarity	η		Q	1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	(visual)					
Orphography	Letter similarity (visual)	μ			1	none
Orphography	Letter similarity (visual)	π		R	1	none
Orphography	Letter similarity (visual)	τ			1	none
Orphography	Letter similarity (visual)	δ			2	none
Orphography	Letter similarity (visual)	ρ		S	2	none
Orphography	Letter similarity (visual)	σ			2	none
Orphography	Letter similarity	6			2	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	(visual)					
Orphography	Letter similarity (visual)	ν		T	2	none
Orphography	Letter similarity (visual)	λ			2	none
Orphography	Letter similarity (visual)	β		U	2	none
Orphography	Letter similarity (visual)	δ			2	none
Orphography	Letter similarity (visual)	β		V	2	none
Orphography	Letter similarity (visual)	χ			2	none
Orphography	Letter similarity	ββ	Σάββατο [ˈsavato]	W	3	β

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
	(voiceless)-double consonants					
Orphography	Letter similarity (voiceless)-double consonants	κκ	εκκλησία [ekliˊsia]		3	κ
Orphography	Letter similarity (voiceless)-double consonants	λλ	αλλαγή [ala'ji]		3	λ
Orphography	Letter similarity (voiceless)-double consonants	μμ	γράμμα ['γrama]		3	μ
Orphography	Letter similarity (voiceless)-	νν	γέννηση ['jenisi]		3	ν

Linguistic Level	**Category**	**Feature**	**Details / Examples**	**Group**	**Difficulty level of feature**	**Prerequisites**
	double consonants					
Orphography	Letter similarity (voiceless)- double consonants	ρρ	άρρωστος ['arostos]		3	ρ
Orphography	Letter similarity (voiceless)- double consonants	σσ	θάλασσα ['θalasa]		3	σ
Orphography	Letter similarity (voiceless)- double consonants	ττ	περιττός [peri'tos]		3	τ
Syntax	Embedding	*O/S Relatives - Anaphoric complement Right Branching*	O adras ide to giatro pu htipise ton aθliti xθes. The_NOM man_NOM		1	*none*

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			saw_3SING.PAST the doctor_ACC that hit the_ACC athlete_ACC yesterday. =The man saw the doctor that hit the athlete yesterday.			
Syntax	Embedding	*O/S Relatives - Anaphoric complement Right Branching*	O adras ide to giatro o opjos htipise ton aθliti xθes. The_NOM man_NOM saw_3SING.PAST the doctor_ACC who hit the_ACC athlete_ACC yesterday. =The man saw the doctor who hit the athlete yesterday.		2	oti (subject), pu, O/S Relatives - Anaphoric complement Right Branching
Syntax	Embedding	*O/O Relatives - Anaphoric complement Right Branching*	O adras ide to giatro pu htipise o aθlitis xθes. The_NOM man_NOM		3	oti (subject), pu, O/S Relatives - Anaphoric complement Right Branching

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			saw_3SING.PAST the doctor_ACC that hit the_NOM athlete_NOM yesterday. =The man saw the doctor that the athlete hit yesterday.			
Syntax	Embedding	*S/S Relatives – Anaphoric complement Center Embedding*	*O adras pu htipise ton aθliti ine ilikiomenos.* *The_NOM man_NOM that hit the_ACC athlete_ACC is old.* *=The man that hit the athlete is old.*		4	oti (subject), pu, O/S Relatives - Anaphoric complement Right Branching
Syntax	Embedding	*S/S Relatives – Anaphoric complement Center Embedding*	*O adras o opjos htipise ton aθliti ine ilikiomenos.* *The_NOM man_NOM who hit the_ACC athlete_ACC is old.* *=The man who hit the athlete*		5	oti (subject), pu, O/S Relatives - Anaphoric complement Right Branching

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			is old.			
Syntax	Embedding	*S/O Relatives – Anaphoric complement Center Embedding*	O adras pu htipise o athlitis ine ilikiomenos. The_NOM man_NOM that hit the_NOM athlete_NOM is old. =The man that the athlete kicked is old.		6	oti (subject), pu, O/O Relatives - Anaphoric complement Right Branching
Syntax	Embedding	*afu*	O adras htipise to aγori afu filise ti jineka. The_NOM man_NOM hit_3ST.PAST the_ACC boy_ACC after kissed_3SG.PAST.PERF the_ACC woman_ACC. =The man hit the boy after he kissed the woman.		7	oti (subject), otan, pu

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Embedding	*otan*	O adras htipise to aγori otan filise ti jineka. The_NOM man_NOM hit_3ST.PAST the_ACC boy_ACC when kissed_3SG.PAST.PERF the_ACC woman_ACC. =The man hit the boy when he kissed the woman.		6	*oti (subject), pu*
Syntax	Embedding	*prin*	O adras htipise to aγori prin filisi ti jineka. The_NOM man_NOM hit_3ST.PAST the_ACC boy_ACC before kiss_3SG.NONPAST.PERF the_ACC woman_ACC. =The man hit the boy before kissing the woman.		7	oti (subject), otan

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Embedding	*eno*	O adras htipise to aγori eno filuse ti jineka. The_NOM man_NOM hit_3ST.PAST the_ACC boy_ACC while kiss_3SG.PAST.IMPERF the_ACC woman_ACC. =The man hit the boy while kissing the woman.		6	*oti (subject), pu,*
Syntax	Embedding	*oti (subject)*	O adras pistevi oti o athlitis filise ti jineka. The_NOM man_NOM believes that the_NOM athlete_NOM kissed_3SG.PAST.PERF the_ACC woman_ACC. = The man believes that the athlete kissed the woman.		1	none
Syntax	Embedding	*oti (object)*	O adras pistevi oti ton athliti filise i jineka.		3	oti (subject)

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			The_NOM man_NOM believes that the_ACC athlete_ACC kissed the_NOM woman_NOM. =The man believes that the athlete the woman kissed.			
Syntax	Embedding	*an*	O adras apori an o athlitis filise ti jineka. The_NOM man_NOM wonders if the_NOM athlete_NOM kissed_3SG.PAST.PERF the_ACC woman_ACC. = The man wonders if the athlete kissed the woman.		2	oti (subject), pu
Syntax	Embedding	*pu*	O adras herete pu o athlitis filise ti jineka. The_NOM man_NOM is glad that the_NOM athlete_NOM kissed_3SG.PAST.PERF the_ACC woman_ACC.		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			= The man believes that the athlete kissed the woman.			
Syntax	Embedding	*na*	O adras θeli na filisi ti jineka. The_NOM man_NOM wants to kiss_3SG.NONPAST.PERF the_ACC woman_ACC. = The man believes that the athlete kissed the woman.		2	oti (subject), pu
Syntax	Passives	Passivized verbs with agent *apo*-phrase	O adras me ta jalja sinelifthi apo tin astinomia sti diadilosi. The_Nom man_Nom with the_PL glasses_PL was arrested_PASS by the_ACC police_ACC at the protest. =The man with the glasses was arrested by the police at the protest.		3	Active verbs (control condition)

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Passives	Passivized verbs with non-agent *apo*-phrase	O adras me ta jalja sinelifthi apo laθos sti diadilosi. The_Nom man_Nom with the_PL glasses_PL was arrested_PASS by accident_ACC at the protest. =The man with the glasses was arrested by accident at the protest.		3	Active verbs (control condition)
Syntax	Passives	Active verbs (control condition)	O adras me ta jalja sinelave apo laθos ton astego. The_Nom man_Nom with the_PL glasses_PL arrested_ACT by accident_NOM the homeless man. =The man with the glasses arrested the homeless man by accident.		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Passives	Passive morphology, reflexive meaning with disambiguating *apo*-phrase	O adras me ta jalja ksiristike apo monos tou xθes. The_Nom man_Nom with the_PL glasses_PL was shaved_PAS by himself yesterday. =The man with the glasses shaved yesterday (by himself).		3	Active morphology (control condition)
Syntax	Passives	Passive morphology, reflexive meaning without disambiguating *apo*-phrase	O adras me ta jalja ksiristike ksana meta apo kero. The_Nom man_Nom with the_PL glasses_PL was shaved_PAS again after a long time. =The man with the glasses shaved again after a long time.		4	Passive morphology, reflexive meaning with disambiguating apo-phrase,
Syntax	Passives	Reflexives with passive meaning and agent apo-	O adras me ta jalja ksiristike apo ton barberi		4	Passive morphology, reflexive meaning with disambiguating apo-

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
		phrase	xθes. The_Nom man_Nom with the_PL glasses_PL was shaved_PAS by the barber yesterday. =The man with the glasses was shaved by the barber yesterday.			phrase
Syntax	Passives	Active morphology (control condition)	O adras me ta jalja ksirise to musi tou xθes. The_NOM man_NOM with the_PL glasses_PL shaved_ACT the beard his yesterday. =The man with the glasses shaved his beard yesterday.		1	none
Syntax	Passives	Passive morphology with *apo*-phrase	To pani tu pliu skistike apo ton aera xθes The_NOM sail_NOM the_GEN boat_GEN was torn_3SG.PASS by the air_ACC yesterday		2	Active morphology (control condition)

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			= The sail of the boat was torn by the air yesterday.			
Syntax	Passives	Passive morphology without *apo*-phrase	To pani tu pliu skistike se δjo komatia xθes The_NOM sail_NOM the_GEN boat_GEN was torn_3SG.PASS into two pieces yesterday = The sail of the boat was torn into pieces yesterday.		2	Active morphology (control condition)
Syntax	Passives	Passive morphology, transitive +NP	I fititria eroteftike ton iθopio prin poli cero The_NOM.FEM student_NOM.FEM fell in love with_PAST.PASS the_ACC actor_ACC a long time ago =The student fell in love with the actor a long time ago.		3	Active morphology, transitive +NP (control condition)
Syntax	Passives	Active morphology, transitive +NP (control condition)	I fititria aγapise ton iθopio prin poli cero The_NOM.FEM student_NOM.FEM		1	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			loved_PAST.ACT the_ACC actor_ACC a long time ago =The student loved the actor a long time ago.			
Syntax	Passives	Passive morphology, intransitive +PP	I fititria kimiθike sti veranda to apojevma The_NOM.FEM student_NOM.FEM slept_PAST.PASS at the porch_ACC in the afternoon =The student slept at the porch in the afternoon.		2	Active morphology, transitive +NP (control condition)
Syntax	Passives	Reciprocal Vs with PP	Ta pedja agaliastikan me xara to apojevma The_NOM.NEUT.PL children_NOM.NEUT.PL hugged_3PL.PAST.PASS with joy in the afternoon =The children hugged joyfully in the afternoon.		3	Passives 3rd plural (reciprocals control condition)

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
Syntax	Passives	Reciprocal Vs without PP	Ta pedja agaliastikan metaksi tus to apojevma The_NOM.NEUT.PL children_NOM.NEUT.PL hugged_3PL.PAST.PASS with eachother in the afternoon =The children hugged joyfully in the afternoon.		3	Reciprocal Verbs- Passives 3rd plural (reciprocals control condition)
Syntax	Passives	Passives 3rd plural (reciprocals control condition)	Ta pedja timoriθikan apo ti daskala to apojevma The_NOM.NEUT.PL children_NOM.NEUT.PL were punished_3PL.PAST.PASS by the teacher in the afternoon =The children were punished by the teacher in the afternoon.		2	Active morphology, transitive +NP (control condition)
Syntax	Discourse anaphors	Proper name (and definiteness)	O Jannis efije otan i Maria teleiose to diagonisma TheJannis left _3SG.PAST.PERF when the		3	Subject Position (Nominative): Definite NP

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			Maria finished_3SG.PAST.PERF the test = Giannis left, when Maria finished the test.			
Syntax	Discourse anaphors	Personal Pronoun	O Jannis efije otan afti teleiose to diagonisma The Jannis left _3SG.PAST.PERF when she_FEM.SG. finished_3SG.PAST.PERF the test. = Giannis left, when she finished the test.		5	Accusative: Direct Object: Proper name (and definiteness) - Personal Pronoun
Syntax	Discourse anaphors	Pro (Null subject)	O Jannis efije otan teleiose to diagonisma The Jannis left _3SG.PAST.PERF when finished_3SG.PAST.PERF the test = Giannis left, when (he)		1	*none*

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			finished the test			
Syntax	Discourse anaphors	Definite NP	O Jannis efije otan o maθitis teleiose to diagonisma The Jannis left_3SG.PAST.PERF when the_DEF student_MASC.SG.NOM finished_3SG.PAST.PERF the test = Giannis left when the student finished the test.		2	Subject Position (Nominative)- Pro (Null subject)
Syntax	Discourse anaphors	Indefinite NP	O Jannis efije otan enas maθitis teleiose to diagonisma The Jannis left _3SG.PAST.PERF when a_INDEF student_MASC.SG.NOM finished_3SG.PAST.PERF the		4	Subject Position (Nominative) - Definite NP

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			test = Giannis left, when a student finished the test.			
Syntax	Discourse anaphors	Proper name (and definiteness)	O Jannis efije otan iδe ti Maria The Jannis left_3SG.PAST.PERF when saw_3SG.PAST.PERF ti_FEM.SG.ACC.DEF Maria = Giannis left when he saw Maria.		4	none
Syntax	Discourse anaphors	Personal Pronoun	O Jannis efije otan iδe aftin The Jannis left_3SG.PAST.PERF when saw_3SG.PAST.PERF her_FEM.SG.ACC = Giannis left when he saw her.		3	none
Syntax	Discourse anaphors	Definite NP	O Jannis efije otan iδe ti maθitria The Jannis		3	none

Linguistic Level	Category	Feature	Details / Examples	Group	Difficulty level of feature	Prerequisites
			left_3SG.PAST.PERF when saw_3SG.PAST.PERF the_FEM.SG.ACC.DEF student_FEM.SG.ACC = Giannis left when he saw the student.			
Syntax	Discourse anaphors	Indefinite NP	O Jannis efije otan iδe mia maθitria The Jannis left_3SG.PAST.PERF when saw_3SG.PAST.PERF a_FEM.SG.ACC.INDEF student_FEM.SG.ACC = Giannis left when he saw the student.		5	subject position (Accusative: Direct Object)-Indefinite NP

Appendix III - Pilot results on feature difficulty for Greek domain model

	Reaction times (ms)	
	Dyslexia	**TD children**
Voice		
animate_passive_with by phrase	1137,413793	996,3225806
animate_passive_without by phrase	1217,461538	1143,66129
inanimate_passive_with by phrase	1500,0625	1033,701493
inannimate_passive_without by phrase	1551,973684	1073,609375
animate_active_without by phrase	1362,606061	1171,485294
anim_pass_reflexive_with disambiguate phrase	1046,025641	895,057971
anim_pass_reflexive_with by phrase	1514,428571	1248,016393
anim_pass_reflexive_without disambiguate phrase	1780,21875	1126,308824
inanimate_pass_unaccusative_with by phrase	1426,405405	1145,507937
inanimate_pass_unaccusative_without by phrase	1092,525	1157,439394
anim_pass_reciprocal without by phrase	1108,551724	838,0588235

anim_pass_reciprocal with by phrase	1330,766667	1023,588235
anim_pass_deponent_transit	1368,621622	1130,238806
anim_pass_deponent_intransitive	1074,457143	1135,571429
Embedding		
adverbial_because	729,7352941	757,2898551
adverbial_when	939,84375	728,3382353
adverbial_before	764,8378378	854,1714286
adverbial_while	849,2352941	703,9850746
complement_oti (that) clause	923,54	805,6
complement_if clause	855,68	804,21
complement_pu (that) clause	867,38	778,05
relative_rightbranching_subject_*that*	711,0294118	760,8676471
relative_right branching_object_*that*	825,1428571	741,9285714
relative_right branching_subject_*who*	939,4054054	849,4920635
relative_right branching_object_*who*	978,4722222	808
relative_center embedded_subject_*that*	836,0277778	765,4925373

relative_center embedded_object_*that*	772,9375	858,9152542
relative_center embedded_subject_*who*	997,46875	846,7794118
relative_center embedded_object_*who*	1044,6875	1111,647059
Anaphors		
proper nouns	917,25	922,9545455
pronouns	1281,714286	1319,75
definite NPs	1195,769231	837,0615385
indefinite NPs	1118,642857	951,9206349
null subjects	1060,352941	706,2857143