SAIL 6

CLIL in Higher Education and the Role of Corpora

A Blended Model of Consultation Services and Learning Environments

Giovanna Carloni

Edizioni
Ca'Foscari

CLIL in Higher Education and the Role of Corpora

SAIL
Studi sull'apprendimento e l'insegnamento linguistico

Collana diretta da | A series edited by
Paolo E. Balboni

6

Edizioni
Ca'Foscari

SAIL
Studi sull'apprendimento e l'insegnamento linguistico

http://edizionicafoscari.unive.it/col/exp/38/SAIL

CLIL in Higher Education and the Role of Corpora

A Blended Model of Consultation Services and Learning Environments

Giovanna Carloni

Venezia
Edizioni Ca' Foscari - Digital Publishing
2015

CLIL in Higher Education and the Role of Corpora: a Blended Model of Consultation Services and Learning Environments
Giovanna Carloni

Edizioni Ca' Foscari - Digital Publishing
Università Ca' Foscari Venezia
Dorsoduro 1686
30123 Venezia
http://edizionicafoscari.unive.it/
ecf@unive.it

1a edizione marzo 2015
ISBN 978-88-6969-008-2 (pdf)

Progetto grafico di copertina: Studio Girardi, Venezia

Certificazione scientifica delle Opere pubblicate da Edizioni Ca' Foscari - Digital Publishing: tutti i saggi pubblicati hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima sotto la responsabilità del Comitato scientifico della collana. La valutazione è stata condotta in aderenza ai criteri scientifici ed editoriali di Edizioni Ca' Foscari.
Scientific certification of the works published by Edizioni Ca' Foscari - Digital Publishing: all essays published in this volume have received a favourable opinion by subject-matter experts, through an anonymous peer review process under the responsibility of the Scientific Committee of the series. The evaluations were conducted in adherence to the scientific and editorial criteria established by Edizioni Ca' Foscari.

Abstract

This volume presents an Italian university-based CLIL (Content and Language Integrated Learning) Learning Center. The blended methodological and language support services provided by the CLIL Learning Center to university professors are thoroughly examined. Within a CLIL theoretical framework, the design of corpus-informed content-specific teaching materials in the target language is also analyzed in depth; the use of corpora in CLIL materials design is focused on in the present work in which applied corpus linguistics plays a pivotal role. Furthermore, the online learning environments created to support effective CLIL technology-enhanced teaching/learning are investigated.

Summary

Introduction	9
CLIL in higher education	11
Corpus-informed teaching materials	55
Online methodological support: the CLIL website	135
Conclusion	169
References	171

Introduction

CLIL, Content and Language Integrated Learning, is instrumental in fostering multilingualism and student mobility in higher education. A blended model of consultation services and learning environments targeted to CLIL instructors and students has been developed at the University of Urbino (Italy), where CLIL courses have been implemented across all departments for a few years. The blended model described in the present work can be implemented effectively in similar CLIL contexts in higher education, in Italy or abroad.

A CLIL Learning Center has been established to provide instructors with methodological and linguistic support. The consultation services, imparted by blended model methodology and language experts, are geared to help instructors in lesson planning, materials design, and lesson delivery, with the extensive use of corpora to support instructors in CLIL planning and teaching.

To scaffold learners' content knowledge and foreign language acquisition, a CLIL blended model has also been developed: an online input-rich learning environment devised to foster further communication, end-user generated knowledge, cooperation, critical engagement with discipline-specific content, and learner-directed learning. The use of the target language is consistently enhanced both in face-to-face and online instruction. Academic language development is catered to in a blended mode. Within a metacognitive framework, the use of corpora has been fostered among learners, especially with regard to academic language and content-specific vocabulary. Overall, CLIL technology-enhanced teaching has been promoted to address digital-age learners' needs and create a multilingual, state-of-the-art learning environment, leading to the making of global workers and citizens.

1 CLIL in higher education

Summary 1.1. A CLIL Learning Center. – 1.2. Methodological support services. – 1.3. Corpora and CLIL materials. – 1.4. The Language Triptych and technology-enhanced learning. – 1.5. Lexical priming. – 1.6. Formulaic language in academic prose. – 1.7. Academic Formulas List. – 1.8 Input in CLIL instruction. – 1.9. Classroom observation.

CLIL projects implemented in higher education are likely to feature (a) short-term objectives, such as multilingualism and learning mobility among EU (European Union) and international students, and (b) long-term objectives, such as employability at national and international levels. These objectives are in keeping with those established by the strategic framework for European cooperation in education and training (ET 2020), in line with the Lisbon objectives, agreed upon by the European Council conclusions of May 12, 2009 (European Union 2009). On this occasion:

> Ministers responsible for higher education met in Leuven/Louvain-la-Neuve to establish the priorities for European Higher Education until 2020. The importance of lifelong learning, widening access and mobility were underlined. The goal was set that by 2020 at least 20% of those graduating in the European Higher Education Area should have had a study or training period abroad (European Objectives 2011, p. 59).

The CLIL approach, endorsed by the EU (`http://ec.europa.eu/languages/language-teaching/content-and-language-integrated-learning_en.htm`), entails teaching subject-specific content through a foreign language; «in the teaching and learning process, there is a focus not only on content, and not only on language» (Coyle et al. 2010, p. 1), but on both. CLIL is instrumental in promoting multilingualism, a key objective of EU policy (Commission of the European Communities 2005, p. 4):

> The Commission's long-term objective is to increase individual multilingualism until every citizen has practical skills in at least two languages in addition to his or her mother tongue.

With regard to learning mobility in higher education, supported by EU-funded programs and bilateral agreements, the strategic objective 1 (European Union 2009) reads «Expanding learning mobility: work[ing] together to gradually eliminate barriers and to expand opportunities for learning mobility within Europe and worldwide». Learning mobility entails the acquisition of foreign languages – necessary to interact effectively in everyday life and in academic content-specific settings – and intercultural skills, as well as the development of global workers and citizens. Noticeably, these four elements are the core tenets of the CLIL approach, as envisaged by Coyle et al. (2010, p. 41) in the 4Cs Framework which «integrates four contextualized building blocks: content (subject matter), communication (language learning and using), cognition (learning and thinking processes) and culture (developing intercultural understanding and global citizenship)».

To accomplish the long and short-term objectives mentioned above, face-to-face and online consultation services have been created to support instructors in CLIL lesson planning and delivering at the University of Urbino. Moreover, a blended model of pedagogical practices has been devised to enable students to explore the possibilities of CLIL and to develop autonomy, creativity, and cultural awareness, as well as digital competences. Digital competences can be defined as «[t]he confident, critical and creative use of ICT to achieve goals related to work, employability, learning, leisure, inclusion and/or participation in society» (Ala-Mutka 2012, p. 1). The blended learning model implemented has been designed to help learners engage actively and collaboratively in CLIL learning. Self-directed study has been especially focused on in online learning, where the use of social media has also been fostered to promote intercultural awareness. The integration of Web 2.0 practices in higher education within a CLIL theoretical framework is aligned with the strategic objective 4 of ET 2020 to «promote creativity and innovation by developing specific teaching and learning methods (including the use of new ICT tools)» (European Union 2009) within innovation-friendly institutions. Interlacing CLIL and digital competences is also in keeping with European Objectives (2011, p. 101) in terms of ICT (Information Communication Technology) development.

> The 'Digital Agenda for Europe' is one of the seven flagships of the Europe 2020 strategy for smart, sustainable and inclusive growth. The overall aim of the 'Digital agenda' is to deliver sustainable economic and social benefits from a digital single market based on fast and ultra fast and interoperable applications (Commission of the European Communities [2010] 245 final. p. 3). An adequate level of digital competences across the population is a prerequisite for this goal and this section focuses on the extent to which education systems are delivering this.

The University of Urbino-based CLIL project, with its focus on content subjects, foreign languages, autonomy, creativity, networking, and digital competences, positions itself at the forefront of pedagogically innovative educational practices in higher education. These practices are in line with Europe 2020 Strategy aims (European Objectives 2011, p. 59): «higher education institutions must be able to play their full part in the so-called 'knowledge triangle', in which education, research and innovation interact». The state-of-the-art consultation and instructional framework developed can be used as a model by other higher education institutions to implement CLIL courses effectively.

1.1 A CLIL Learning Center

A CLIL Learning Center has been established at the University of Urbino to scaffold CLIL courses implemented across all departments.[1] The mission of the Center is to promote and support CLIL teaching/learning effectiveness. To accomplish this mission, the Center provides instructors involved in the project with methodological and English language support; noticeably, experts in the field provide consultation services on a one-to-one basis. The CLIL expert is responsible for methodological support geared towards helping instructors develop and carry out the CLIL approach using customized teaching/learning resources and instructional strategies. Methodological support includes needs analysis, syllabus design, lesson planning, activity design, and classroom observation. In this context, it is worth mentioning that over 90% of CLIL courses are taught in English. Thus, native English speakers are available to provide instructors with language support regarding classroom management language, content-specific vocabulary building, and classroom observation. The language support has been organized on the grounds of three macro-interdisciplinary areas, namely the humanities, science as well as law and business.

To effectively address instructors' needs, a link to a web-based preliminary questionnaire is emailed to the professors involved in the project before the beginning of each academic year. The questionnaire, delivered through SurveyMonkey (`http://www.surveymonkey.com`), is tailored to gather data on the CLIL courses in general and to collect information about instructors' needs in terms of methodological and language support services in particular. On submission of each questionnaire, the CLIL expert browses the data submitted and contacts the instructors who request methodological support; a meeting is consequently scheduled.

1 For a brief introduction to the blended model devised see Carloni (2013a).

Furthermore, English native-speaker experts are informed about requests for language support. Meetings are thus arranged by language specialists with instructors. A final questionnaire is also emailed to instructors to gather data on the courses taught.

1.2 Methodological support services

The CLIL expert provides methodological support to instructors in both face-to-face and online modes. During face-to-face meetings, needs analysis is first carried out; then, the expert guides instructors to customize CLIL lessons for the targeted classes.

In CLIL learning environments, subject-specific input, activities, tasks, and classroom oral interaction are all in the target language. Through a variety of classroom setups – in pairs or small groups as well as in lockstep geared to promote communication and negotiation of meaning – learners are consistently required to engage with the input and interact with peers and instructors in the foreign language.

The CLIL lesson plan adopted requires instructors to identify content and language objectives on the grounds of the discipline-specific topics selected. The «triple focus on content, language and learning skills» (Marsh et al. 2011, p. 25) is consistently enhanced in CLIL teaching.

The expert is available to help professors identify content and language objectives as well as possible content learning difficulties the delivery of the topics chosen may entail. With regard to content objectives, the subject topics are usually selected autonomously by instructors while designing their course syllabi. Some topic shifts are, however, negotiated during face-to-face meetings to better address learners' needs and language proficiency levels. Overall, the expert supports professors in planning and designing effective CLIL lessons within a given syllabus, providing them with guidance «to identify key concepts of content subjects and make them accessible to learners by modifying teaching to take into account students' diverse language competences and needs» (Marsh et al. 2011, p. 21). The theoretical model underpinning CLIL lesson design is presented to instructors in detail alongside lesson planning:

> Lesson plans provide a means of formalizing learning activities and a framework for teachers to reflect in a deeper and more creative way about how they design and structure activities for different students and help achieve constructive alignment between theory and practice. (Littlejohn; Conole, Fill) They are particularly useful in helping tutors to plan *blended learning* (i.e. the integration of technology supported methods with face-to-face teaching). (Conole 2007, p. 87)

Within this framework, instructors are guided to identify learners' ZPD, Zone of Proximal Development, «which is […] the layer of skill or knowledge which is just beyond that which the learner is currently capable of coping» (Williams, Burden 1997, p. 40). Professors are also introduced to Anderson and Krathwohl's taxonomy of educational objectives (http://www.celt.iastate.edu/pdfs-docs/teaching/RevisedBloomsHandout.pdf). The taxonomy is a pivotal tool in CLIL planning since, as Coyle et al. (2010, p. 54) suggest, «it explores the relationship between cognitive processing (learning) and knowledge acquisition (of content)». The Knowledge Dimension is mapped according to «four types of knowledge that learners may be expected to acquire or construct – ranging from concrete to abstract» (http://www.celt.iastate.edu/pdfs-docs/teaching/RevisedBloomsHandout.pdf). On the other hand, the Cognitive Process Dimension classifies cognitive processes in terms of «increasing cognitive complexity – from lower order thinking skills to higher order thinking skills» (http://www.celt.iastate.edu/pdfs-docs/teaching/RevisedBloomsHandout.pdf). Teachers are encouraged to devise suitable content learning objectives on the grounds of the different combinations of the knowledge and cognitive dimensions featured in the taxonomy.

Language objectives include, in particular, key content-specific vocabulary items, academic language, language functions, and grammatical structures. As Marsh et al. (2011, p. 42) suggest «Content is related to learning and thinking (cognition). To enable the learner to create their own interpretation of content, it must be analyzed for its linguistic demands». In terms of language objectives, CLIL is informed by the BICS/CALP construct:

> BICS defines basic interpersonal communication skills (survival skills). It refers to a minimum level of competence to handle non-complex interpersonal situations. CALP defines cognitive academic proficiency as a level of linguistic competence needed to operate in more formal academic settings which demand more abstract use of language. (Cummins 1979)

CLIL activity design is carried out using the CLIL Matrix devised by Cummins (1979) to map tasks in terms of low/high cognitive and linguistic demands as well as context-embedded and context-reduced demands:

> BICS is said to occur when there are contextual supports and props for language delivery. Face-to-face 'context embedded' situations provide, for example, non-verbal support to secure understanding. Actions with eyes and hands, instant feedback, cues and clues support verbal language. CALP, on the other hand, is said to occur in 'context reduced' academic situations. Where higher order thinking skills (e.g. analysis, synthesis, evaluation) are required in the curriculum, language is 'dis-

> embedded' from a meaningful, supportive context. Where language is 'disembedded' the situation is often referred to as 'context reduced. (Baker 2006, p. 174)

In university CLIL classes, learners are most likely to tackle CALP context-reduced tasks.

With regard to language objectives, the CLIL expert helps instructors define language functions and identify, in particular, both the key content-specific vocabulary items and the academic language learners need to master in order to process the input and produce output. To identify the correlation between content and language objectives, the Language Triptych devised by Coyle et al. (2010) has been adopted. This framework classifies CLIL language in terms of «language of learning, language for learning and language through learning» (Coyle et al. 2010, p. 36).

1.3 Corpora and CLIL materials

Studies show that corpus linguistics can play a pivotal role in course planning and materials design, especially with regard to content-specific resources:

> Flowerdew (2001) used concordancing techniques in a specialized corpus of biology texts in order to select the most salient linguistic features to include in the ESP syllabus, to extract language examples for instructional materials and to evaluate the authenticity of currently used materials. (Crawford Camiciottoli 2010, p. 97)

Thus, in keeping with corpus linguistics research, the methodology expert creates CLIL course-tailored corpus-informed teaching materials using content-specific corpus-driven findings, namely data which «are extracted from corpora, using the methodology of corpus linguistics, then intellectually processed and turned into results» (Teubert 2004, p. 112). As McCarthy (2004, p. 15) suggests with regard to corpus-informed teaching materials: «corpus data alone [do] not dictate an instructional syllabus. Rather, such data are considered in light of other pedagogical requirements». Thus, subject-specific lexico-grammatical features are employed to design corpus-informed activities in keeping with state-of-the-art Second Language Acquisition theories.

The use of corpus-driven data in CLIL materials is of paramount importance, owing to the pivotal role played by intercollocability in genre-specific texts:

> Sinclair states that an accurate description of a text's content can be achieved by examining 'intercollocability'. Central to this notion are the patterns of co-occurrence of words (collocation) found within texts,

> which are central to meaning-making and contribute, by means of particular phrasings, to genre-specific usage. (Cheng 2009, p. 161)

1.3.1 English academic prose

Corpus-informed activities are specially targeted to help CLIL students to effectively tackle academic prose in English. As research shows, in CLIL settings, foreign language learners are likely to experience:

> difficulty of assimilating often dense and lengthy authentic texts, written to an academic discipline-specific target audience, but characterised by a complete absence of the types of 'caretaking' language that can be exercised in oral discourse where the target audience is [a non-native speaker]. (Ball, Lindsay 2013)

English academic prose is made up of three vocabulary sets: general vocabulary, namely «the higher-frequency vocabulary necessary to achieve a basic functionality in a language» (Schmitt 2010, p. 75); academic vocabulary, that is «vocabulary which is common across academic disciplines […], such as insert, orient, ratio, and technique» (Schmitt 2010, p. 78); and discipline-specific vocabulary, i.e. «[t]echnical words or phrases […] which are recognizably specific to a particular field» (Schmitt 2010, p. 77).

The role of function words must be accounted for while analyzing scientific writing and devising corpus-informed, content-specific materials. As Schmitt (2010, p. 55) suggests with regard to the high number of function words in texts, «Corpus research […] found that approximately 270 function word types in English (176 word families) accounted for 43-44% of the running words in most texts». This is the reason why function words usually rank high in corpus-retrieved frequency lists.

Nominalisation is a key feature of written academic English: «Written registers use nouns to a much greater extent than any other word class» (Biber 2006, p. 47). In this respect, research shows that content words account for more than half of scientific texts: «Approximately 60% of all content words in academic prose are nouns» (Biber 2006, p. 15). Furthermore, it is important to note that content-specific vocabulary is pivotal in academic writing since it conveys key discipline-specific concepts: «Technical vocabulary is essential to understanding discourse in a field, and can cover 10% or more of the running words in a text from the field (Sutarsyah, Nation, Kennedy)» (Schmitt 2010, p. 77).

While devising teaching materials, one also has to bear in mind that technical lexical items can often be difficult for students to process and acquire due to their low degree of imageability and concreteness:

> *Imageability* refers to how easy it is to imagine a concept. *Concreteness* is «a variable that expresses the degree to which a word (or rather, the entity the word refers to) can be experienced by the senses» (de Groot). […] The degree of imageability/concreteness is important because it has been shown that more concrete/imageable words are learned far better than less concrete/imageable words, with the effect of being both large and robust (de Groot; Ellis, Beaton). (Schmitt 2010, p. 53)

1.4 The Language Triptych and technology-enhanced learning

The Language Triptych previously mentioned is the result of a longstanding immersion-based research; within this theoretical framework, Lyster (cited in Coyle et al. 2010, p. 35) advocated a new linguistic system for immersion to «combine the program's communicative agents with a more systematic and graded language component aimed at second language learner[s]». Other scholars researching immersion language programs argued for similar educational frameworks:

> Snow, Met, Genesee […] suggested identifying content-obligatory language (essential for learning the content) and content-compatible language (which «supports the content of a lesson, as well as the linguistic cultural objectives of the curriculum»). (Coyle et al. 2010, p. 36)

The Language Triptych was thus developed by Coyle (Coyle et al. 2010) to meet the needs identified by researchers in this field.

1.4.1 Language of learning

One of the main objectives of CLIL instruction consists in promoting the acquisition of «content-obligatory language [which] includes technical vocabulary and other domain specific expressions» (Lyster 2007, p. 28). In the Language Triptych mentioned earlier, discipline-specific language is defined as language of learning – that is «language needed for learners to access basic concepts and skills relating to the subject theme or topic» (Coyle et al. 2010, p. 35). To help instructors devise content-specific vocabulary language objectives, the CLIL expert uses Sketch Engine (`http://www.sketchengine.co.uk`), a web-based software enabling users to both search existing corpora and compile new corpora, and to identify keywords –that is, «words in a corpus whose frequency is unusually high (positive keywords) or low (negative keywords) in comparison [to] a reference corpus» (McEnery et al. 2006, p. 347). Positive keywords are usually the most frequently used domain-specific vocabulary items in the

texts analyzed. Sketch Engine is used to generate keyword lists based on both the reading materials adopted in each course and the slides created by instructors for the CLIL lessons. Focusing only on keywords is not enough to generate a thorough lexical profile of genre-specific texts: «The over-reliance on keywords means that most of the important information regarding the content of individual texts has not been utilized» (Cheng 2009, p. 160). For lexical profiling to be complete, lexical bundles also need to be investigated in academic texts: «three-word […], four-word […], five-word […], and six-word […] combinations […] were very frequent, making up about […] 21% of academic prose» (Schmitt 2010, p. 123). Lists of n-grams, that is, of subject-specific multi-word sequences, are thus also retrieved with Sketch Engine by the methodology expert and made available to instructors. As Biber (2006, pp. 133-134) suggests, multi-word units are «'lexical bundles', defined simply as the most frequent recurring sequences of words. […] [T]hey are usually not idiomatic in meaning, and they are usually not complete grammatical structures». The crucial role played by multi-word units in fluency acquisition in a foreign language is strongly supported by McCarthy (2006, p. 9) on the grounds of corpus-based studies:

> Corpora reveal that much of our lexical output consists of multi-word units; language occurs in ready-made chunks to a far greater extent than could ever be accommodated by a theory of language insistent upon the primacy of syntax.

In a CLIL setting, content-specific multi-word unit acquisition is therefore especially instrumental in fostering foreign language learners' output. Noticeably, fluency relies largely on the activation of top-down automatic production of clusters: «learners create language bottom-up. Native speakers and expert users, because they have stored thousands of ready-made collocations,[2] are working top-down, and [are] simply assessing collocations from memory» (McCarthy et al. 2010, p. 34).

The keyword and lexical bundle lists generated with Sketch Engine are emailed by the CLIL expert to instructors, who can use them to devise language objectives and to create paper-based or online glossaries to be provided to students before class.

Furthermore, Sketch Engine-retrieved keyword and n-gram lists are used by the CLIL expert to generate course-tailored word clouds (Fig. 1). To this purpose, Wordle (`http://www.wordle.net`), a free word cloud generating software, is used. Word clouds graphically feature vocabulary items

2 A collocation is «the habitual co-occurrence of words/linguistic items in close proximity to one another» (Hoffmann et al. 2008, p. 264).

on the grounds of their frequency; thus, words that occur more often in a text are featured with a bigger font size. As a result, the distribution of content-specific lexical items in a text affects the dimension of the words in the word clouds, which makes statistically based data visually accessible to users at a glance. Texts can be pasted directly in the Wordle generating box; the web-based tool itself counts word frequency. Colours, fonts, and layouts can be tweaked, which makes word clouds more motivating for learners. On the other hand, customized course-tailored word clouds can also be created with Wordle Advanced (`http://www.wordle.net/advanced`), which requires users to provide words and their frequency. Thanks to the statistically based data retrieved through the Sketch Engine-generated keyword lists previously mentioned, the CLIL expert and instructors usually create word clouds using Wordle Advanced. Course-tailored word clouds are made available for instructors and students online before class on the CLIL website (`http://clilteaching.weebly.com`), which will be described later in this work.

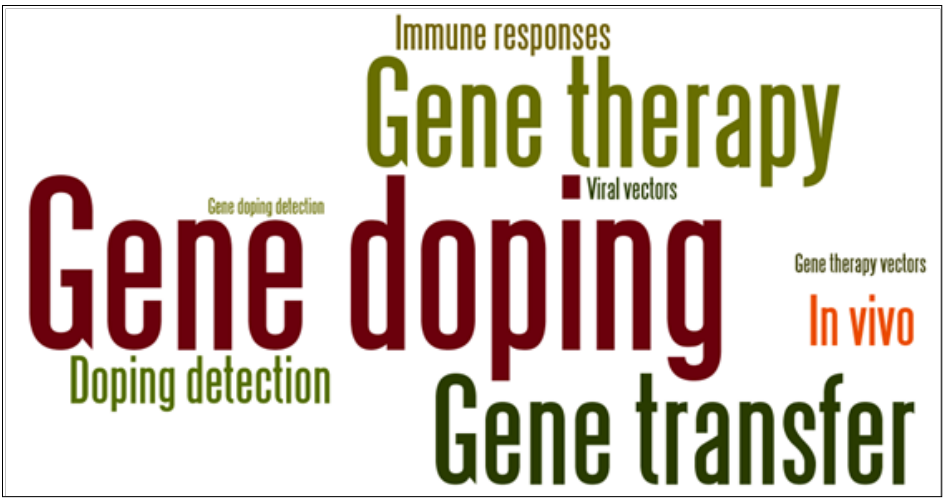


Figure 1. A word cloud

1.4.2 Language for learning

Language for learning, the second parameter of the Language Triptych, refers to academic language:

> [I]n CLIL settings […] the learners will need to be supported in developing skills such as those required for pair work, cooperative group work, asking questions, debating, chatting, enquiring, thinking, memorizing and so on. […] Developing a repertoire of speech acts which relate to the content, such as describing, evaluating and drawing conclusions, is essential for tasks to be carried out effectively. (Coyle et al. 2010, p. 37)

Various tools have been identified and used by the CLIL expert to help instructors to detect the academic language learners need to master to engage effectively in CLIL learning environments. In particular, to firstly familiarize and afterwards enable instructors to identify the key academic vocabulary targeted in their CLIL lessons autonomously, the AWL (Academic Word List), devised by Averil Coxhead (`http://www.victoria.ac.nz/lals/resources/academicwordlist`), is introduced and uploaded on the CLIL website:

> Coxhead used a 3.5 million word corpus consisting of written academic texts from journals, textbooks and coursebooks originating in different parts of the native English-speaking world, covering 28 subject areas subsumed under four major disciplinary areas (arts, science, commerce and law). She examined the distribution of words not included in the most frequent 2,000 English words form [sic] West's General Service List (West). Based on criteria of frequency (at least 100 occurrences in the corpus for members of each word family) and range (i.e. a minimum number of occurrences across the different disciplines and subject areas), Coxhead produced a list of 570 word families (base forms and their related inflected and derived forms) which accounted for around 10% of the total tokens in the corpus. The same word-families found to cover less than 1.5% of the total words in an equally-sized written corpus consisting of fiction texts. The AWL therefore offers a 'fingerprint' of written academic vocabulary, the common core items which make it different from other types of writing. Most fruitfully, focusing on AWL in vocabulary teaching and learning offers the possibility of increasing comprehension of academic text far more rapidly and efficiently than through just enlarging one's general vocabulary […]. Coxhead proposes dividing the AWL into sublists of 60 items for practical learning purposes to provide a systematic framework for vocabulary teaching, and even though the AWL is simply a list, advocates teaching its members in context. (O'Keeffe et al. 2007, pp. 198-199)

AWL sublists feature academic vocabulary according to their frequency of occurrence. So, for example, «Sublist 1 contains the most common words in the AWL. Sublist 2 contains the next most common words, and so on» (http://www.victoria.ac.nz/lals/resources/academicwordlist/sublists).

Instructors need to familiarize themselves with word lists to better cater to their learners' needs and foreign language proficiency while planning activities:

> In presenting the case for the use of frequency lists in language learning, Cobb states that the most frequent 1,000 word families make up 72 per cent of all text, and the most frequent 2,000 words make up 79.7 per cent. Knowledge of these word families is therefore a powerful tool for increasing text comprehension at relative speed. However, this is not necessarily sufficient for a reader to cope with authentic text. Cobb quotes research indicating that for reading to be reliably successful, and to support further vocabulary acquisition, at least 95 per cent of text tokens should already be 'known'.
> The gap between the 95 per cent target, and the 2000 word plateau can be addressed, Cobb suggests, either by continuing down frequency lists, or through the use of more specialized lists such as the Academic Word List (Coxhead) which consists of 570 additional word families most commonly found in academic text. The new combined list of 2,570 words, according to Cobb, would then enable a reader of academic text to recognize about 90 per cent of the tokens in an average academic text. Cobb's Vocabulary Profiler, which he has generously and freely made available on the world-wide web, analyzes texts according to these lists and can tell us exactly what the proportions of GSL (General Service List), AWL (Academic Word List) and 'off-list' (p. 330) tokens in a text are. Theoretically, we could then predict how an idealized learner taught according to these lists would be able to cope with a particular text. (Hancioglu, Eldridge 2007, p. 331)

The CLIL expert provides instructors with frequency lists of the academic vocabulary featured in the English reading materials adopted. These lists enable professors to plan language objectives in terms of academic language and thus gradually build learners' academic vocabulary competences. Course-tailored academic word lists are generated by the expert through the freely downloadable Range (http://www.victoria.ac.nz/lals/resources/range.aspx) software, which runs three base word lists:

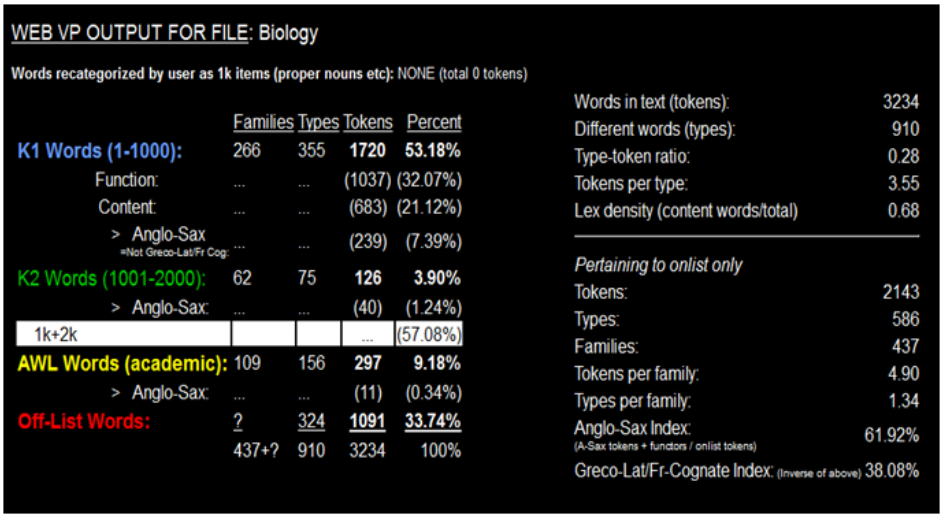
> Three ready made base lists are available. The first (BASEWRD1.txt) includes the most frequent 1000 words of English. The second (BASEWRD2.txt) includes the 2nd 1000 most frequent words, and the third

> (BASEWRD3.txt) includes words not in the first 2000 words of English but which are frequent in upper secondary school and university texts from a wide range of subjects. All of these base lists include the base forms of words and derived forms. The first 1000 words thus consist of around 4000 forms or types. The sources of these lists are A General Service List of English Words by Michael West (Longman, London) for the first 2000 words, and The Academic Word List by Coxhead containing 570 word families [for the rest]. (http://www.victoria.ac.nz/lals/resources/range.aspx)

The list of words retrieved by Range not included in the three base lists is also made available to users. This list is most likely to feature content-specific words. Through the Range program, the percentage of common English words, academic language, and content-specific language of English reading materials can be estimated. Instructors can use the online user-friendly version of Range (http://www.er.uqam.ca/nobel/r21270/cgi-bin/webfreqs/web_vp.html) to analyze academic texts on their own and professors can thus retrieve vocabulary profiles autonomously.

Within the pedagogical framework described so far, academic texts are also analyzed by the CLIL expert with the VP (Vocabulary Profiler), free online software adapted for the web by Tom Cobb (http://conc.lextutor.ca/vp). The VP, which is one of the tools provided by Compleat Lexical Tutor (http://www.lextutor.ca), runs the GSL (General Service List)[3] and the AWL (Academic Word List). An off-list including the words not featured in the GSL and AWL lists is also made available to users. To obtain a vocabulary profile of the texts adopted by instructors in the CLIL classes, the expert pastes parts of the targeted English reading materials into the box provided by the VP. The data thereby retrieved show the percentage of words belonging to the three lists as featured in the screenshot below (Fig. 2).

3 «The General Service List was developed by Michael West in the 1950's to represent the 2,000 most common words appearing in the English language. The list has gone through some refinements over the years, but it is generally quite stable, and useful for teachers of English as a second language». (http://www.wordsift.com/wordlists)



WEB VP OUTPUT FOR FILE: Biology

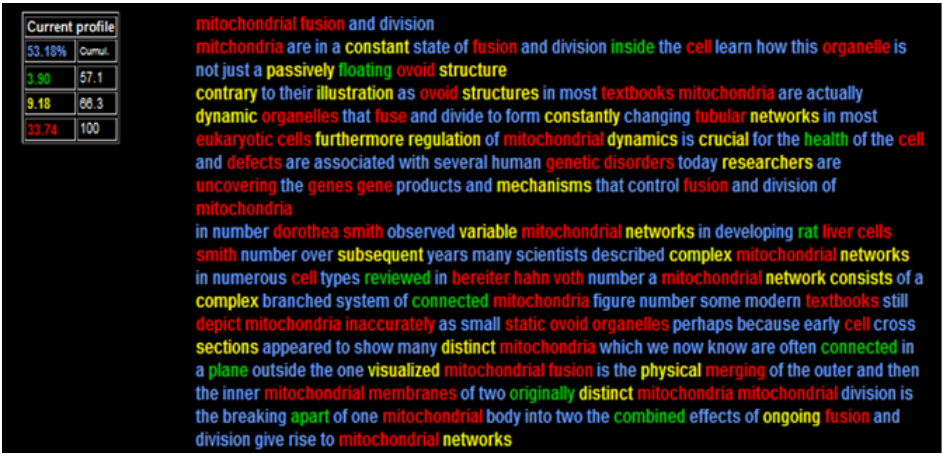
Words recategorized by user as 1k items (proper nouns etc): NONE (total 0 tokens)

	Families	Types	Tokens	Percent
K1 Words (1-1000):	266	355	1720	53.18%
Function:	...	...	(1037)	(32.07%)
Content:	...	...	(683)	(21.12%)
> Anglo-Sax =Not Greco-Lat/Fr Cog:	...	...	(239)	(7.39%)
K2 Words (1001-2000):	62	75	126	3.90%
> Anglo-Sax:	...	...	(40)	(1.24%)
1k+2k			...	(57.08%)
AWL Words (academic):	109	156	297	9.18%
> Anglo-Sax:	...	...	(11)	(0.34%)
Off-List Words:	?	324	1091	33.74%
	437+?	910	3234	100%

Words in text (tokens):	3234
Different words (types):	910
Type-token ratio:	0.28
Tokens per type:	3.55
Lex density (content words/total)	0.68
Pertaining to onlist only	
Tokens:	2143
Types:	586
Families:	437
Tokens per family:	4.90
Types per family:	1.34
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	61.92%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	38.08%

Figure 2. Text profile

Scrolling down the same page of data presented above, users can obtain visual access – through a colour-coded system – to how vocabulary belonging to the three lists is distributed across the texts processed (Fig. 3).



Current profile	
53.18%	Cumul.
3.90	57.1
9.18	66.3
33.74	100

mitochondrial fusion and division
mitchondria are in a constant state of fusion and division inside the cell learn how this organelle is not just a passively floating ovoid structure
contrary to their illustration as ovoid structures in most textbooks mitochondria are actually dynamic organelles that fuse and divide to form constantly changing tubular networks in most eukaryotic cells furthermore regulation of mitochondrial dynamics is crucial for the health of the cell and defects are associated with several human genetic disorders today researchers are uncovering the genes gene products and mechanisms that control fusion and division of mitochondria
in number dorothea smith observed variable mitochondrial networks in developing rat liver cells smith number over subsequent years many scientists described complex mitochondrial networks in numerous cell types reviewed in bereiter hahn voth number a mitochondrial network consists of a complex branched system of connected mitochondria figure number some modern textbooks still depict mitochondria inaccurately as small static ovoid organelles perhaps because early cell cross sections appeared to show many distinct mitochondria which we now know are often connected in a plane outside the one visualized mitochondrial fusion is the physical merging of the outer and then the inner mitochondrial membranes of two originally distinct mitochondria mitochondrial division is the breaking apart of one mitochondrial body into two the combined effects of ongoing fusion and division give rise to mitochondrial networks

Figure 3. Text profile

Instructors and learners are invited to use the VP autonomously; instructors can use it to plan language objectives and activities, while on the other hand, for the learners, the consistent use of the VP is likely to promote their metacognition and enhance autonomous management of academic language building.

The online VP and off-line Range enable the CLIL expert and instructors to answer the following questions which are pivotal to devising suitable learning objectives:

- How large a vocabulary is needed for reading this text?
- If a learner has a vocabulary of 2,000 words, how much of the vocabulary in the text will be familiar to the learner?
- What are the words in the text which the learner is not likely to know? (`http://www.victoria.ac.nz/lals/resources/range.aspx`).

Students are also encouraged to use WordSift (`http://www.wordsift.com`) to analyze the academic texts they have to engage with autonomously. Learners can paste parts of their English reading materials into WordSift to obtain a word cloud visualizing the fifty most frequent vocabulary items featured in the targeted text (Fig. 4). The more frequent words are, the larger their font size is. Learners can choose to have the words displayed alphabetically or from the most to the least common, amongst other ways.

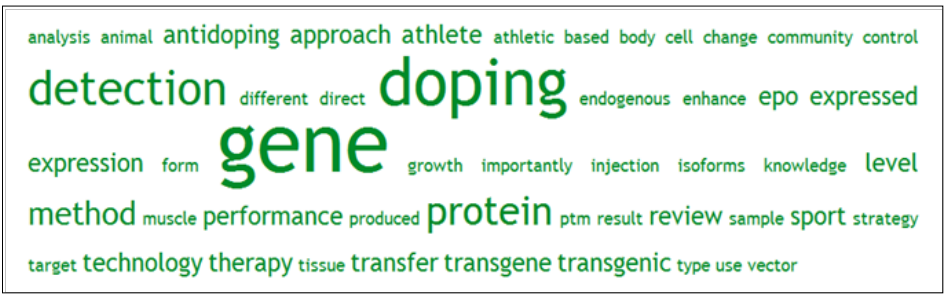


Figure 4. A WordSift-generated word cloud

Students can have the academic words highlighted in the word clouds. The process is supported by Coxhead's (2000) Academic Word List (Fig. 5).

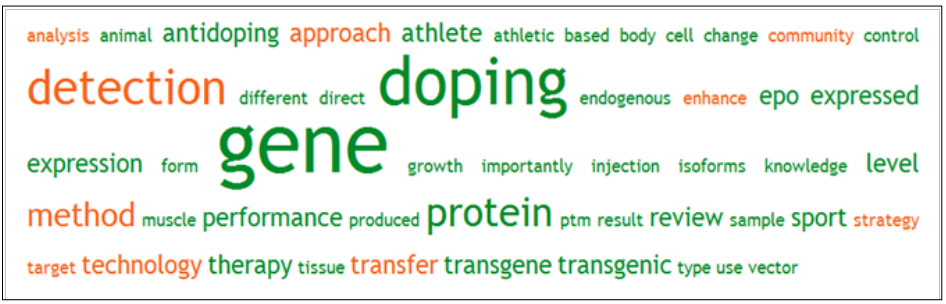


Figure 5. A WordSift-generated word cloud with academic words highlighted

Students can also decide to have the General Service List, Social Studies, Language Arts, Science and Math words[4] highlighted (Fig. 6). Users can create customized lists to have their texts analyzed.

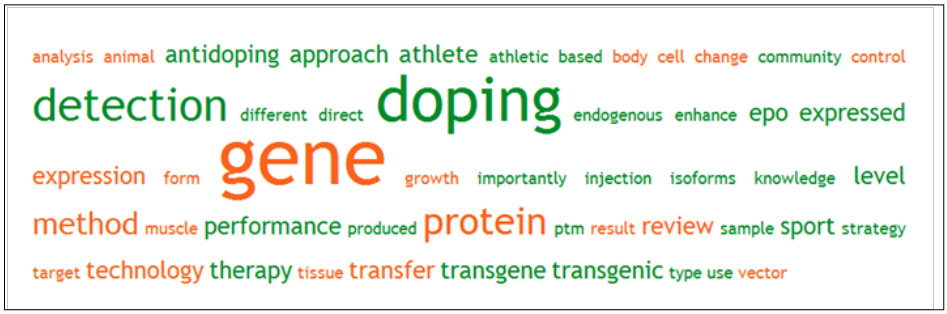


Figure 6. A WordSift-generated word cloud with science-specific words highlighted

The most common word of the text pasted in WordSift is featured through the interactive Visual Thesaurus (`http://www.visualthesaurus.com`), together with its synonyms, antonyms, most common collocates, and related words (Fig. 7).

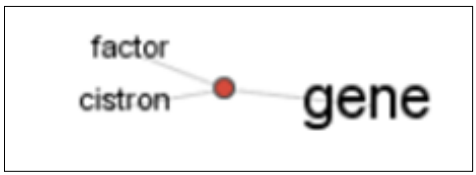


Figure 7. The most common word in a WordSift-generated word cloud featured through the interactive Visual Thesaurus

By rolling the mouse over the elements provided by the web-based thesaurus, further meaning-related information is provided. If students are interested in investigating one of the synonyms, antonyms, or collocates provided, it is enough for them to click on the item to have it displayed at the center, together with its related features (Fig. 8).

4 «Language Arts, Math, Science and Social Studies Words. These words include those identified in Robert Marzano's important work (Marzano, 2004), but his list contains many compound words (e.g. nuclear reaction) which do not work well in the WordSift context, since WordSift operates at the individual word level. So, with apologies, we chopped up those compounds into individual words. We then also had several teachers with content specialization go through the list to delete and add items, consulting their own glossaries and wealth of teaching experience. In spite of these transformations, teachers using the Marzano materials should be able to make use of WordSift to aid in their work».(`http://www.wordsift.com/wordlists`)

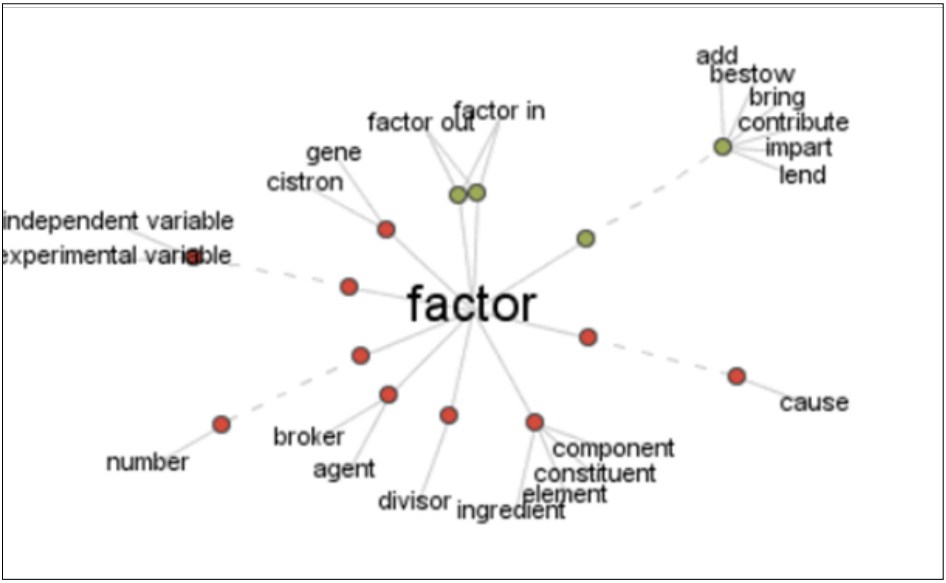


Figure 8. One of the synonyms related to the most common word in a WordSift-generated word cloud featured through the interactive Visual Thesaurus

Learners can also get information on each lexical item featured in the word cloud by simply clicking on the targeted element.

The service Word and Phrase – Academic (`http://www.wordandphrase.info/academic`), provided by the COCA, Corpus of Contemporary American English (`http://corpus.byu.edu/coca`), which targets academic texts in particular, is also used by the CLIL expert to promote instructors' and learners' awareness of academic and content-specific language. Students are encouraged to employ the user-friendly tool on their own in a self-directed learning perspective. They can paste a text into the box provided and retrieve data about the academic (AWL) and content-specific words featured in the targeted text (Fig. 9). The AWL is divided into two ranges: range one, featuring AWL list 0–500, and range two, featuring AWL list 501–3,000. The academic words featured in range one are more frequent in academic English than those featured in range two. Through a drop-down menu, users can further customize their search by selecting the subject-specific content of the text to be analyzed. The results thereby retrieved are colour-coded as the image below shows:

Figure 9. Colour-coded academic and content-specific words retrieved with Word and Phrase – Academic

The lists of words retrieved are also available for offline use (Fig. 10).

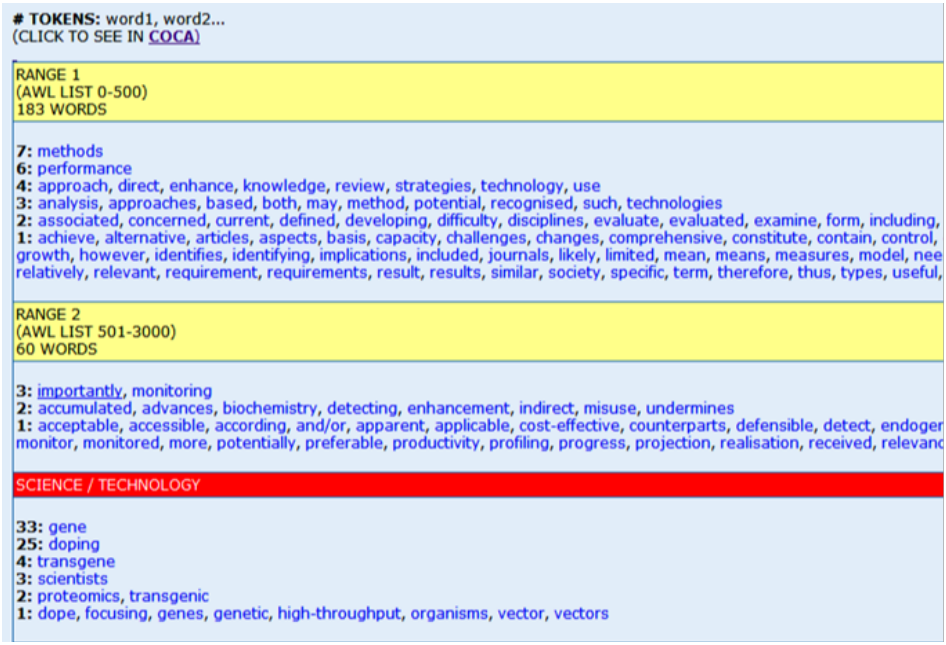


Figure 10. Wordlists available offline (Word and Phrase – Academic)

Students can acquire more information about single words featured in the colour-coded text by clicking on the targeted vocabulary element, thereby generating concordances of the query item (Fig. 11) – that is «list[s] of all the occurrences in the corpus of the query item together with some sur-

rounding context in the form of words to the left and the right» (Hoffman et al. 2008, p. 264). Thus, collocations can be retrieved. As the instructions read:

> you can click through the words in the text to see a detailed 'word sketch' of any of the words –showing their definition, and detailed information for the word from COCA– collocates (which provide meaning into the meaning and usage of the word), re-sortable concordance lines, and the frequency of the word (overall, and by genre). (`http://www.wordandphrase.info/analyzeText.asp`)

Figure 11. Concordances retrieved (Word and Phrase – Academic)

Learners can also search for phrases featured in the texts, and «[i]n this way, this resource is like a 'collocational thesaurus' to see what related phrases are most likely in different styles of English» (`http://www.wordandphrase.info/analyzeText.asp`).

With the Word and Phrase – Academic tool, learners can monitor and manage their academic language development autonomously. Instructors can use the tool to plan language objectives and activities.

To help learners carry out academic written tasks in English, learners are introduced to Concord Writer (`http://conc.lextutor.ca/concord_writer/index.pl?lingo=English`), free software using the Brown Corpus, which enables users to generate concordances (Fig. 12). Students can pick the search item by clicking on the targeted word while writing a text in the box provided.

Figure 12. Concord Writer

Learners are also introduced to Just the Word (`http://www.just-the-word.com`). Students can access this free web-based tool, which uses a subset of the British National Corpus, to search collocations without having to go through concordances, thus simplifying the process. Bar graphs indicate the degree of frequency of the collocations retrieved (Fig. 13), which can be accessed by clicking on the collocations listed. In addition, by clicking on the button provided, learners can visualize collocations through Wordle-generated word clouds, which makes it easier for users to detect the most commonly used collocations of the entry word queried.

Figure 13. Collocations retrieved with Just the Word

To foster students' pushed output in class, the CLIL expert provides instructors with a handout featuring academic language sentence stems, that is, phrases suitable to express the most common academic language functions. For example, the 'Social Studies Questioning Stem Guide', devised by John Seidlitz (cited in Short et al. 2011, pp. 13-14), is made available for this purpose. Professors are invited to give the handout to learners before CLIL classes and explain to them how to use sentence stems effectively during class discussions. Learners' output is therefore scaffolded in keeping with the CLIL Framework requiring CLIL teachers «to scaffold language learning during content classes» (Marsh et al. 2011, p. 19).

Academic language sentence stems are, however, not enough to completely scaffold language output and foster academic language proficiency in CLIL settings. Therefore, structured academic-language activities to be carried out autonomously by learners are provided online on the CLIL website.

Overall, output language scaffolding in terms of academic and subject-specific vocabulary as well as lexical bundles is consistently provided to CLIL learners. The pivotal role played by output language scaffolding in CLIL learning environments is highlighted by Meyer (2010, p. 15):

> Scaffolding also supports language production (= pushed output) by providing phrases, subject-specific vocabulary and collocations needed to complete assignments. It helps students to verbalize their thoughts appropriate to the subject manner. In other words, scaffolding done right will boost students' cognitive academic language proficiency (CALP).

Students can get acquainted with the tools mentioned thus far on the CLIL website.

1.4.3 Language through learning

Language through learning, which conceives learning «as conceptual change and […] key to building knowledge» (Harasim 2012, p. 90), is the language learners need to voice newly elaborated ideas emerging on the spot in a CLIL setting. In this perspective, language through learning «is to do with capturing language as it is needed by individual learners during their learning process –and this by definition cannot be predicted in advance» (Coyle et al. 2010, p. 38). Besides instructors, English native-speaker experts may be made available in class to provide language through learning on the spot; it is necessary to consistently promote the use and review of language through learning in class. As Coyle et al. (2010, p. 37) suggest, «this emerging language needs to be captured, recycled and developed strategically by teachers and learners».

1.5 Lexical priming

While devising corpus-informed activities, lexical priming is consistently enhanced; it is the process through which readers or speakers can recognize or produce a word faster if they have been previously exposed to a semantically related lexical item or co-text:

> priming refers to the phenomenon in which prior exposure to language somehow influences subsequent language processing, which may occur in the form of recognition or production. […] [T]he initial language form or aspects of its meaning, referred to as the prime, facilitates the recognition or production of a subsequent form or aspects of its meaning, which is referred to as the response or target. (McDonough, Trofimovich 2008, pp. 1-2)

Lexical priming is based on the activation of semantic associations, which «exist[…] when a word or word sequence is associated in the mind of a language user with a semantic set or class, some members of which are also collocates for that user» (Hoey 2005, p. 24). Semantic associations are primed by specific language items or co-texts:

> the notion of semantic priming is used to describe the way a 'priming' word may provoke a particular 'target' word. For example, a listener, previously given the word *body*, will recognize the word *heart* more quickly than if they had previously been given an unrelated word such as *trick*; in this sense, *body* primes the listener for *heart*. This has an obvious connection with word association games. The word *body* sets up a word association with *heart*, which the word *trick* does not. (Hoey 2005, p. 8)

The associational system underpinning the language acquisition process informs lexical priming:

> Word association research has illustrated the connections between lexical items in the mental lexicon. Research has also shown that lexical items are not processed in isolation, but are affected by their surrounding context. (Schmitt 2010, p. 107)

Within this theoretical framework, priming is considered as the basic paradigm of language: «priming is the driving force behind language use, language structure and language change» (Hoey 2005, p. 12). In particular, lexis is conceived as the structuring element of language: «The theory reverses the roles of lexis and grammar, arguing that lexis is complexly and systematically structured and the grammar is an outcome of this lexical structure» (Hoey 2005, p. 1).

Lexical priming expects interactants' minds to work like a concordancer while processing incoming input or producing output:

> The notion of priming […] assumes that the mind has a mental concordance of every word it has encountered, a concordance that has been glossed for social, physical, discoursal, generic and interpersonal context. (Hoey 2005, p. 11)

Cognitive processes in terms of language processing and production are therefore envisioned as working on the basis of collocational processing:

> We can only account for collocation if we assume that every word is mentally primed for collocational use. As a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. (Hoey 2005, p. 8)

There are two kinds of priming; that is, productive and receptive priming:

> Productive primings occur when a word or word sequence is repeatedly encountered in discourse and genres in which we are ourselves expected (or aspire) to participate and when the speakers or writers are those whom we like or wish to emulate. Receptive primings occur when a word or word sequence is encountered in contexts in which there is no probability, or even possibility, of our ever being an active participant – party political broadcasting, interviews with film stars, eighteenth-century novels – or where the speaker or writer is someone we dislike or have no empathy with – drunken football supporters, racists, but also sometimes stern teachers and people of a different age group. (Hoey 2005, pp. 11-12)

It is productive priming which needs to be fostered in CLIL.

Lexical priming seems to be highly effective in content-specific learning environments since, as Hoey (2005, p. 10) suggests, «domain-specific primings […] are the norm». Lexical priming needs to be implemented in CLIL activities to promote content and foreign language processing as well as effective acquisition. To this purpose, prior to reading a scientific text or listening to a subject-specific lecture, students need to be provided with key content-specific vocabulary items –working as primes– in order for priming to occur effectively during the reading/listening process. Therefore, it is of paramount importance in CLIL to design activities targeted to introduce discipline-specific lexical items and collocations, previously identified through keyword lists and concordance outputs, in order for

lexical priming to occur while learners are engaged in discipline-specific reading and listening comprehension. Priming also occurs in conversations: «Priming leads to a speaker unintentionally reproducing some aspect of the language, and that aspect, thereby reproduced, in turn primes the hearer» (Hoey 2005, p. 9). Thus, CLIL activities targeted to foster output production also need to be created in order for priming to occur during interactional exchanges.

While designing corpus-informed, content-specific activities, one has to bear in mind that priming is a transitory system which is affected on an on-going basis by speakers' contact with new communicative contexts:

> Priming need not be a permanent feature of the word or word sequence; in principle, indeed, it never is. Every time we use a word, and every time we encounter it anew, the experience either reinforces the priming by confirming an existing association between the word and its co-texts, or it weakens the priming, if the encounter introduces the word in an unfamiliar context or co-text or if we have chosen in our own use of it to override its current priming. It follows that the priming of a word or word sequence is liable to shift in the course of an individual's lifetime, and if it does so, and to the extent that it does so, the word or word sequence shifts slightly in meaning and/or function for that individual. This may be referred to as a drift in the priming. (Hoey 2005, p. 9)

In CLIL settings, the transitory state of lexical priming can thus be consistently affected through course-tailored activities.

In keeping with the theoretical principles presented so far, the data retrieved through corpus investigation are used to create CLIL course-tailored activities aimed to foster lexical priming. Furthermore, content lecturers are sensitized to how lexical priming occurs in order for them to plan lectures, namely their speech, and supporting materials such as Power Point presentations, which are liable to foster lexical priming and thereby scaffold CLIL students' cognitive processes.

In CLIL, corpus-informed, content-specific activities, mainly associative and category primings, are enhanced as they are cognitively more direct and therefore more effective:

> Associative priming is a form of semantic priming for prime and target words that are close semantic associates of each other but are not members of the same semantic category. For example, associatively related pairs of words are sugar-sweet and grass-green.
> Category priming is a form of semantic priming for prime and target words that are members of the same semantic category. For example, bird-robin and furniture-table are categorically related pairs of words.
> Mediated priming, as opposed to direct semantic priming, refers to prim-

> ing between words that are not related directly. Examples of such words are stripes-lion. The semantic relationship between these two words is mediated by the word tiger. (McDonough, Trofimovich 2008, p. 62)

1.6 Formulaic language in academic prose

The role of formulaic language in reading/listening comprehension, output production and acquisition processes is pivotal. As Schmitt (2010, p. 109) suggests in terms of reading comprehension: «The Conklin and Schmitt self-timed reading methodology showed that formulaic sequences were read more quickly than non-formulaic sequences». With regard to output production, research provides interesting data:

> Erman and Warren calculated that 52-58% of the language they analyzed was formulaic, and Foster came up with a figure of 32% […]. Biber *et al.* found that around 30% of the words in their conversation corpus consisted of lexical bundles, and about 21% of their academic prose corpus. (Schmitt 2010, pp. 117-118)

Lexical bundles are likely to enhance both language intake and output production:

> The acquisition approach is related to what might be called the 'psycholinguistic' approach, where formulaic language is assumed to be holistically stored in the mind. There is evidence for this on the phonological front: formulaic sequences are typically spoken more fluently, with a coherent intonation contour, to the extent that this has been accepted as one criterion of formulaicity (e.g. van Lancker, Canter, Terbeek; Peters). (Schmitt 2010, p. 121)

Studies highlight the assets of lexical bundles in fostering oral output production: «There is now considerable converging support for the notion that formulaic language provides processing advantages over creatively generated (i.e. non-formulaic) language» (Schmitt 2010, p. 137). The interrelationship between lexical bundles and foreign language acquisition is at the cutting edge of today's research in SLA:

> The nature of formulaic language and its acquisition is likely to become of ever-greater interest as the field turns to more pattern-based models of language acquisition (e.g. pattern grammar – Hunston and Francis) and construction grammar (Tomasello), which posit that the human facility for language learning is based on the ability to extract patterns from input, rather than being under the guidance of innate principles and

> parameters which determine what aspects of grammar can and cannot be acquired (Ellis). (Schmitt 2010, pp. 140-141)

Working on formulaic language,[5] and therefore promoting the processing and acquisition of multi-word units, is especially important in CLIL where the integration of language and content is essential: «In CLIL, successful content learning is particularly dependent on language: enhanced language learning is dependent on content learning» (Marsh et al. 2011, p. 18). Working on phraseology, with a special focus on content-specific and academic vocabulary, is pivotal in a CLIL learning environment: «There is a widespread feeling that formulaic language is especially problematic for L2 learners, and its lack/misuse is a major reason why L2 output can feel unnatural and non native-like» (Schmitt 2010, p. 142).

In this context, it is worth mentioning that there are different kinds of formulaic language. On the one hand, there are n-grams featuring contiguous lexical items – «'N-grams' […], e.g. *a lot of, one of the, you know, and I don't think*, are contiguous words that constitute a phrase, or a pattern of use» (Cheng 2009, p. 161) – which are likely to be genre-specific. On the other hand, there are non-contiguous multi-word units defined as skipgrams or phrase-frames exemplifying the 'open slot' variety: «Skipgrams, or 'phrase-frames', (Fletcher) such as *the past three years*, *the past few years* uncover patterns of phraseology with non-contiguous sequences, but not those with positional variation» (Cheng 2009, p. 161).

Research has tried to shed new light on non-contiguous word units featuring positional variation in the last few years. To this purpose, ConcGram, a software tool, has been specially created. As Greaves (2009, p. 8) suggests with regard to keyness and phraseology: «Given that phraseology is all pervasive in language, ConcGram can be used to extend the notion of keyness beyond keywords to include the full range of phraseology». ConcGram can be used to detect non-contiguous multi-word units, concgrams, which are interrelated with constituency and positional variation:

> ConcGram […] on top of its capability to handle both constituency variation (i.e. AB, A*B) and positional variation (i.e. AB, BA), conducts fully automated searches [to find] word-combinations comprising [of] up to five words. The products of the searches are termed 'concgrams'. A concgram is all of the permutations of constituency variation and positional variation generated by the co-occurrence of two or more words. (Cheng 2009, p. 162)

5 For an analysis of lexical bundles in academic spoken English see Carloni (2013b).

Concgrams are additionally useful to determine the aboutness of a text:

> According to Philips (1989), aboutness is a product of global patternings of a text. […] The identification of the phraseological profile of a text is linked to what Philips refers to as the aboutness of a text. The phraseological profile is all of the word associations in a text or corpus, and the aboutness of the text or corpus can be determined from the word associations that are specific to that particular text or corpus. Word associations which are specific to a text or corpus are termed 'aboutgrams'. (Greaves 2009, p. 8)

Aboutgrams are therefore the result of the patterning of all text-specific word associations retrieved through the phraseological profiling of the targeted text or corpus.

Discipline-specific aboutgrams seem to be particularly useful to promote lexical priming and enhance language awareness in CLIL content-specific materials. As Cheng (2009, p. 173) suggests, aboutgrams can be used to help learners identify the aboutness of the content-specific materials they have to study: «[ConcGrams have] implications for ESP teaching, learning and research in raising language awareness and increasing knowledge about the aboutness associated with the topics in the discipline- and professional-specific discourses».

ConcGrams are retrieved by the methodology expert to help instructors devise corpus-informed teaching materials.

1.7 Academic Formulas List

Corpus-based genre-specific studies have recently dealt quite extensively with academic language. Coxhead's (2000) AWL (Academic Word List) is the product of this research supporting the common-core hypothesis:

> Corpus-based research on different academic genres has also reignited the debate on the 'common-core hypothesis' (Bloor). Proponents of this approach maintain that there exists a common set of linguistic structures and vocabulary that will be found across a range of academic texts, regardless of discipline and genre. Coxhead's (2000) corpus-based research of vocabulary items in a 3.5-million-word corpus composed of written academic texts drawn from the disciplines of arts, commerce, law and sciences, seems to support the hypothesis with its extraction of common core vocabulary items forming an academic word list (AWL). (Flowerdew 2012, p. 192)

In this light, a list of multi-word units more frequently occurring in academic language was devised by Simpson-Vlach and Ellis (2010). Within a pedagogical framework, this newly devised list carried Coxhead's (2000) analysis of academic language a step further:

> Simpson-Vlach and Ellis [...] have taken Coxhead's idea from word to phrase level and devised an Academic Formulas List (AFL) which consists of word combinations that occur significantly more often in academic than in non-academic speech and writing. Both studies (AWL and AFL) take [an] indirect approach to using specialized corpora in language teaching and contribute to improving the teaching of English for academic purposes through informing syllabus design. (Römer 2010, p. 27)

Simpson-Vlach and Ellis's (2010, p. 487) AFL (Academic Formulas List) is divided into various subcategories:

> The AFL includes formulaic sequences identified as (i) frequent recurrent patterns in corpora of written and spoken language, which (ii) occur significantly more often in academic than in non-academic discourse, and (iii) inhabit a wide range of academic genres. It separately lists formulas that are common in academic spoken *and* academic written language, as well as those that are special to academic written language alone and academic spoken language alone. The AFL further prioritizes these formulas using an empirically derived measure of utility that is educationally and psychologically valid and operationalizable with corpus linguistic metrics. The formulas are classified according to their predominant pragmatic function for descriptive analysis and in order to marshal the AFL for inclusion in English for Academic Purposes instruction.

The Academic Formulas List is used by the methodology expert to help instructors with CLIL materials design.

1.8 Input in CLIL instruction

In CLIL learning environments, learners are provided with comprehensible input that needs to be slightly above learners' competence as Krashen's Input Hypothesis suggests (cited in Richards, Rodgers 2001, p. 182):

> people acquire language best by understanding input that is slightly beyond their current level of competence:
> An acquirer can 'move' from a stage I (where I is the acquirer's level

> of competence) to a stage I + 1 (where I + 1 is the stage immediately following I along some natural order) by understanding language containing I + 1. (Krashen and Terrell).
> Clues based on the situation and the context, extralinguistic information, and knowledge of the world make comprehension possible.

Providing learners with comprehensible input is necessary to start fostering intake,[6] as Krashen holds (cited in Larsen-Freeman, Long 1991, p. 140):

> Comprehension is necessary, Krashen believes, in order for the input to become *intake*, i.e. data taken in or assimilated by the learner and used by the learner to promote IL [Interlanguage] development. The ability to understand items not yet in the IL grammar derives, Krashen maintains, from the speech adjustments made [or any kind of scaffolding provided] to learners, plus the learners' use of shared knowledge and (linguistic and extralinguistic) context.

However, for intake to occur, learners also need to produce pushed output as advocated by Swain's Output Hypothesis (cited in Meyer 2010, p. 19):

> Research into the complex relationship between language and thinking and its effect on language learning/acquisition has led Swain to formulate the idea of 'languaging' which she defines as «the process of making meaning and shaping knowledge and experience through language» (Swain). Languaging completes our thoughts/cognition/ideas and transforms them into artifacts that allow for further contemplation, which, in turn, transforms thought. While speaking (or writing), a new or deeper understanding may be achieved (O'Connell).

The social constructivist interactionist framework relies on the interrelationship between the outer socio-linguistic context, where language learning occurs, and learners' inner individual differences. In this acquisitional perspective, oral interaction and negotiation of meaning are instrumental in fostering content learning and foreign language development. Notably, in an interactionist perspective, the comprehensibility of input is fostered through negotiation of meaning, which provides learners with «opportunities […] to use the second language to mediate content learning during academic tasks» (Lyster 2007, p. 134). Negotiation of meaning and therefore modified input are likely to promote content and foreign language acquisition:

6 «Intake is that part of the input that learners notice and therefore take into temporary memory. Intake may subsequently be accommodated in the learner's interlanguage system (i.e. become part of long-term memory). However, not all intakes are accommodated» (Ellis 1999, p. 708).

> negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or the more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways. (Long 1996, pp. 451-452)

Negotiation of meaning in the target language is thus pivotal to promote intake in a CLIL learning environment: «The principle that informs these theories is that verbal interaction is of crucial importance for language learning as it helps to make the 'facts' of the L2 salient to the learner» (Ellis 1999, p. 244). Learners' consistent interaction with the input through negotiation of meaning therefore needs to be implemented. As Kost (2008, pp. 154-155) claims:

> During the negotiation process, learners receive input, attend to feedback, and produce output […] (Pellettieri). By modifying language interactions and increasing input comprehensibility, they negotiate for both meaning and form. These features are argued to be facilitative for language acquisition within the interactionist framework (Long; Gass).

Lightbown and Spada (1997, pp. 29-30) summarize the tenets of interactionism as follows:

> Interactionists claim that a crucial element in the language acquisition process is the modified input that learners are exposed to […]. Proponents of the interactionist view, such as Michael Long, agree with Krashen that comprehensible input is necessary for language acquisition. However, they are more concerned with the question of how input is made comprehensible. They see interactional modifications which take place in conversations between speakers and non-native speakers as the necessary mechanism for this to take place (Long).
> For Long and others, modified interaction must be necessary for language acquisition. This relationship has been summarized as follows:
> 1. Interactional modification makes input comprehensible;
> 2. comprehensible input promotes acquisition.
> Therefore,
> 3. interactional modification promotes acquisition.

To make content-specific input comprehensible in a CLIL classroom, various instructional strategies need to be applied. To this purpose, the guidelines, devised by Coonan (2008, p. 18) to tackle the main issues taken into account while creating CLIL materials, have been adopted to carry out the consultation services; they are «a synthesis based on Ellis, [which] places features of input (code used, code complexity, cognitive complexity),

'working' conditions (interaction, dialogue), cognitive processes called into play, and outcomes (code used, dimension and mode) on a continuum from easier to more difficult».

To conclude, for effective content and foreign language learning to take place in CLIL settings, comprehensible input, intake, output, and negotiation of meaning need to be consistently enhanced and carefully orchestrated.

1.8.1 Preparation process

The authentic content-specific input that learners deal with in CLIL learning is an added value, even if, however, students need scaffolding to process and interact with it effectively. Thus, in order for learners to cope successfully with CLIL classes, instructors need to draft every step carefully, starting with the preparation process, scheduled to be carried out autonomously by students before class instruction.

Learners' prior knowledge construction, which plays a pivotal role in learning processes in general and in CLIL learning environments in particular, needs to be consistently enhanced through various strategies. Noticeably, to effectively process incoming input, learners need to be able «to build on prior language and content knowledge, skills, attitudes and experiences» (Marsh et al. 2011, p. 22). In this light, the self-directed preparation phase scheduled to take place prior to class is targeted to help the students involved in the project to build content-specific background knowledge autonomously. The resources and activities provided to learners before class aim to promote students' overall comprehension of the main discipline-specific topics scheduled to be introduced and discussed in class. Within a metacognitive framework, the content prerequisite knowledge needed is made explicit to learners.

Students' acquisition of new content knowledge structures prior to CLIL classes also aims to reduce learners' cognitive load during face-to-face instruction. As Kalyuga (2011, p. 205) suggests, «if task-relevant knowledge structures are not available in long-term memory, learners may need to simultaneously process many new elements of information in working memory resulting in a cognitive overload». The preparation process is thus also targeted to decrease the amount of information learners' working memory is required to process while attending to new content-specific input taught through a foreign language in class.

During the self-directed preparation phase, learners can construct knowledge base in an online, non threatening learning environment customized for each course by the CLIL expert: «Appropriate external guidance may be required to assist […] [CLIL] learners in acquiring new knowledge structures in a cognitively efficient and non stressful manner» (Kalyuga 2011, p. 205). If the preparation phase is effectively accom-

plished, during lectures, students will experience only intrinsic cognitive load, that is, the «essential, necessary, productive, and useful load required for achieving specific learning goals» (Kalyuga 2011, p. 206). As Kalyuga (2011, p. 206) holds:

> Intrinsic cognitive load is caused by specific cognitive activities resulting in new or modified knowledge structures in long-term memory. Such activities involve concurrent processing of interacting elements of information in working memory and integrating them with available knowledge structures in accordance with specific learning goals. Intrinsic load is associated with internal complexity of the learning task and, therefore, is always relative to the level of learner expertise since what is complex for novices may be simple for experts. In order to achieve meaningful learning outcomes, it is necessary to accommodate this load without exceeding available working memory resources.

Prior to CLIL lessons, instructors are invited to provide learners with the Power Point slides that will be used in class during lectures: key content-specific concepts and their definitions, key content-specific glossaries, word clouds featuring keyword lists, reading assignments for each class, graphic organizers targeted to scaffold the comprehension of daily assigned reading materials, graphic organizers geared to detect, understand, and analyze key content-specific vocabulary items featured in the reading materials, preparatory podcasts providing a brief overview of the main concepts that will be introduced in class and/or research questions students are expected to answer during class instruction, academic language sentence stems, and 'Recommendations for students', that is, advice for students aimed at helping them to cope with cognitively challenging CLIL classes. The students' tips have been devised on the grounds of the results which emerged from a CLIL university-based research project (Bartik et al. 2009) carried out at the Faculty of Engineering in Brussels.

The resources created for the preparation phase are uploaded on the CLIL website and on the instructor's homepage. In particular, instructors are invited to create activities focusing on key content-specific vocabulary items using Word Dynamo (`http://dynamo.dictionary.com`), a free web-based, user-friendly tool. With Word Dynamo, professors can create matching activities and flash cards providing key vocabulary items and their definitions both in the written and audio mode. Students can do the activities online or they can print them out.

If learners feel overwhelmed by the length of the assigned English reading articles, they can use Text Compactor (`http://textcompactor.com/about`), a free online tool especially suitable for expository texts, to summarize the articles:

> After text is placed on the page, the web app calculates the frequency of each word in the passage. Then, a score is calculated for each sentence based on the frequency count associated with the words it contains. The most important sentence is deemed to be the sentence with the highest frequency count. (http://textcompactor.com/about)

Through a slider, users can decide the percentage of the text they want to be kept in the summary. When students are happy with the summary, they can listen to the text through vozMe (http://vozme.com/index.php?lang=en), free text-to-speech software. With Text Compactor, learners can come into contact with content materials in a less threatening way; after accessing the text globally, students can tackle the full English articles more easily.

By means of the scaffolding implemented through the CLIL website, the use of metacognitive strategies is enhanced. As a result, students can take responsibility for effectively managing their learning process prior to class. Learners can select the resources and activities suitable to accomplish this goal; personalized learning is enhanced. In this metacognitive framework, the preparation phase also works as an awareness-raising strategy. Students are given the opportunity to realize that «more knowledgeable learners use their available knowledge structures for managing cognitive load» (Kalyuga 2011, p. 205). Students can thus start being active learners and prepare to engage effectively in face-to-face CLIL classroom instruction.

1.8.2 Face-to-face CLIL classroom instruction

In face-to-face CLIL lessons, instructors and students consistently use the target language throughout the whole class period, and this can be quite challenging for students. Thus, to help students on the first day of CLIL lessons, an English language expert can be invited to the class to present the methodological and language support available to students at the CLIL Learning Center.

Teachers are provided with guidelines to effectively manage face-to-face CLIL instruction. At the beginning of each class, instructors can briefly summarize, either in Italian or in the target language, the main issues introduced in the previous lesson. This strategy aims to help students, who may have missed some information, to catch up. This will help students not to panic if they do not understand everything in class, since they know the main ideas will be outlined again.

Instructors then present the learning objectives of the lesson orally and in the written mode to the class through a Power Point presentation. Students' awareness of daily outcomes is thus promoted. The interrelationship between metacognition and motivation plays a key role in CLIL learning.

This connection is especially important when dealing with less competent students. As Graesser and D'Mello (2011, p. 17) suggest, «Learners […] persist on content that is relevant to their goals even when their prior knowledge about the material is modest (McCrudden, Schraw)». CLIL learning environments benefit from metacognition:

> When motivated by consciously chosen goals, learners experience a sense of control and meaningfulness of their experience. Learning goals represent an important part of a learner prior knowledge that performs an important guiding role in cognitive processing. Balancing external guidance with learner internal goal structures is important for creating positive affective states and higher levels of motivation. (Kalyuga 2011, p. 212)

The introduction of daily learning objectives is pivotal to foster learners' self-evaluation at the end of each class. Students are then enabled to evaluate to what degree they accomplished the learning outcomes stated at the beginning of the lesson. Self-evaluation is likely to enhance students' self-efficacy, and as Barnhardt suggests (1997, p. 26):

> Self-efficacy forms the basis for self-esteem and learning motivation. Self-efficacious learners feel confident about solving a problem because they have developed an approach to problem solving that has worked in the past. They attribute their success mainly to their own efforts and strategies, believe that their own abilities will improve as they learn more, and recognize that errors are part of learning.

Self-efficacy is also instrumental in fostering active learning in a CLIL setting:

> Social-cognitive theory places special emphasis on the role of personal motivation. When a learner experiences success at a valued task, he or she develops a sense of self-efficacy – a belief that one has the capability to succeed at that kind of task. Self-efficacy can, in turn, affect whether the student is willing to try a task, as well as the students' persistence at the task, thoughts during the task, and eventual performance […]. Using appropriate strategies can help build self-efficacy by creating success experiences and by giving students the tools for future successes. Social-affective strategies, such as self-talk, can also help students work through tasks by providing direct self-motivation. In our own research, we found that foreign language students with high self-efficacy also report using more learning strategies than do foreign language students with low self-efficacy. (Chamot et al. 1999, p. 159)

1.8.3 Motivation

The motivation phase aims at building learners' background knowledge in class. Firstly, instructors activate learners' prior knowledge. Brainstorming and graphic organizers can be used to recall information and make it visually accessible to the entire class. Eliciting students' past learning and experience is pivotal to fostering comprehension of new information. In effective comprehension tasks, students draw extensively on pre-existing knowledge and beliefs linking new input to their background knowledge and experiences. Therefore, instructors need to apply instructional strategies «for fostering in students the habit of linking new learning with their personal experience (e.g. language, content subjects, personal experience and the out-of-school world)» (Marsh et al. 2011, p. 19).

After activating students' available knowledge base, instructors need to introduce key content-specific vocabulary items and concepts through various activities. Promoting learners' hypothesis-making skills before delivering lectures is also of paramount importance. On the grounds of the discipline-specific words, concepts, and images provided to students thus far, instructors can invite learners to anticipate information and make predictions about the content of the lecture they are about to listen to. These strategies are geared to activate top-down processing that is instrumental in:

> [a]chieving interpretation and production of language meaning through prior knowledge of content, context and culture. This process may allow learners to guess the meanings of words they have not met before, and to make some sense out of larger chunks of written and oral texts. (Saville-Troike 2006, p. 195)

On the other hand, bottom-up processing enables learners to «[a]chiev[e] interpretation and production of language meaning through prior knowledge of the language system and of physical (graphic and auditory) cues» (Saville-Troike 2006, p. 186). Successful learning requires effective interaction between top-down and bottom-up processing.

1.8.4 Presentation

In the presentation phase, instructors deliver their lectures. The authenticity of the input is a key component of CLIL; however, as previously mentioned, it is also a challenge for students. Therefore, the subject-specific content provided by instructors at this stage needs to be adapted and made more comprehensible to learners, while also preventing them from experiencing cognitive overload. To this purpose, instructors need to provide students with while-listening activities.

To scaffold students' listening comprehension processes, professors need to lecture for about 20/25 minutes at the most during a 60-minute lesson, control their speech rate, use Power Point presentations providing learners with multimodal input catering for diverse needs and cognitive styles, and give students while-listening tasks, such as completing diagrams, filling in blank spaces in the slides provided before class, taking notes using customized graphic organizers, answering multiple-choice questions, etc.

The CLIL expert is available to support instructors in devising Power Point presentations suitable to students' language proficiencies. Input enhancement, targeted to make input salient, is implemented in Power Point instructional material design with boldfacing, colour coding, and key content-specific vocabulary labeling. Input enhancement aims to foster noticing:

> Sharwood Smith has coined the term 'input enhancement' to discuss focus on form. In his framework, input enhancement is any external attempt (by instructors or materials) to make features of the input more salient to learners and could come in any forms. […] Positive enhancement could be as simple as colour coding or boldfacing forms in reading texts. (VanPatten 1996, p. 84)

Images, videos, and flash animations can be incorporated into lectures. However, while selecting multimedia materials, instructors need to bear in mind the cognitive load each kind of material entails for learners with different language proficiencies:

> Animated visualizations that represent movements and processes are commonly believed to be appropriate for presenting dynamic information. However, no convincing evidence has been obtained indicating higher learning effects of animations as compared to static diagrams (Hegarty, Kriz, Cate; Park, Hopkins; Tversky, Morrison, Betrancourt). According to cognitive load theory, continuous animations could be too cognitively demanding for novice learners because of high levels of transitivity. These learners could benefit more from studying a set of static diagrams. On the other hand, knowledge structures of more experienced learners may help them in handling the transitivity of animations, as they may have sufficient working memory resources for constructing and running dynamic mental representations. Static graphics could be less beneficial for these learners because their knowledge structures would need to be integrated and reconciled with redundant (for them) details displayed in graphics, thus unnecessarily consuming additional cognitive resources. (Kalyuga 2011, p. 209)

The CLIL expert is also available to help instructors create while-listening/viewing activities geared towards promoting input comprehension. Scaffolding students' comprehension processes with structured, cognitively and linguistically appropriate while-listening or while-viewing activities is pivotal in CLIL settings: «Scaffolding […] reduces the cognitive and linguistic load of the content/input (= input-scaffolding) which means that scaffolding helps students understand the content and language of any given material» (Meyer 2010, p. 15).

To be effective, while-listening/viewing activities need to foster intrinsic cognitive load while preventing learners from experiencing extraneous cognitive load, the latter being one of the most likely drawbacks of CLIL classes:

> Working memory resources that are actually devoted to dealing with intrinsic cognitive load and lead to learning are germane resources in contrast to extraneous working memory resources that are devoted to dealing with extraneous cognitive load (Sweller). This separate dimension of actually allocated working memory resources stresses the role of germane resources in learner engagement in processing relevant aspects of a task and importance of instructional methods that motivate and engage students in learning-effective cognitive activities. More engaged and motivated learners invest more of their working memory resources into dealing with intrinsic load thus leading to better learning (Schnotz). (Kalyuga 2011, p. 206)

While-listening/viewing activities are especially important for lower language proficiency learners, who may otherwise feel overwhelmed:

> applying to novice learners […] insufficient external instructional guidance may not compensate for [the] limited knowledge of these learners, thus forcing them into applying search-based processes resulting in extraneous cognitive load. (Kalyuga 2011, p. 207)

Instructors can also use online tools to devise real-time while-listening/viewing targeted to enhance and monitor students' comprehension of the domain-specific content delivered in lectures. Free tools, such as Mentimeter (`http://mentimeter.com`), can be used to provide learners with while-listening multiple-choice/true-false questions in real time. Mentimeter allows instructors to create and post customized questions students can answer using any device connected to the Internet. To enter the virtual room to answer the questions, learners need to connect to the URL address provided by the instructor; professors can also embed the questions in the CLIL website, which makes answering faster. Professors can visualize students' results immediately on their computer screen; to promote metacognitive awareness, learners can be shown the results. Online real-time

monitoring enables teachers to realize, for example, what concepts have not been fully understood by students during lectures and therefore review them right away. Learners can then be asked to answer the same question twice to check whether the instructional practice deployed to explain the concept again was effective enough.

Furthermore, during lectures, learners can be invited to use Twitter (`http://www.twitter.com`), a free real-time micro-blogging tool, to tweet, in the target language, their questions and doubts about the subject-specific content being delivered. After lecturing, instructors can check their students' tweets and answer the questions tweeted. To display the tweets, instructors can use Tweetchat (`http://tweetchat.com`), which allows users to pause and refresh the tweeting stream as well as setting the refreshing speed; this tool is useful especially if one needs to focus on certain issues. TweetShow (`http://tweet-show.com/index.html`), a web application that allows users to display tweets retrieved through hashtags in full-screen, can also be used. However, the stream cannot be paused. To use Twitter effectively in class, learners need to be told in advance how and why to use this tool. At the beginning of each lesson, on a Power Point slide, instructors can display their Twitter username and the daily class hashtag instrumental to identify the targeted Twitter stream. Thus, the backchannel use of Twitter in class can be fostered, and through real-time feedback, instructors can further monitor students' comprehension. Perry (cited in Conole, Alevizou 2010, p. 34) identifies the following as an asset of the synchronous use of Twitter in class:

> Using backchannels to generate instant feedback within lectures is another factor for potential success. This is consistent with Yardi who notes that:
> Online backchannel chat rooms offer the potential to transform classroom learning in unexpected and powerful ways. However, the specific ways in which they can influence teaching pedagogy and learning opportunities are less well understood. Activities in a backchannel may include the dissemination of ideas, knowledge building, asking and answering questions, engaging in critical discourse, and sharing information and resources.

Moreover, after lecturing, instructors are encouraged to invite students to ask questions about issues in the foreign language they did not understand or found interesting. Interacting in the foreign language can be face threatening for learners, especially at the beginning, and for lower-proficiency students in particular. Instructors can therefore encourage students to ask questions in the target language using TodaysMeet (`http://todaysmeet.com`), a free online live stream backchannel tool learners can use to express ideas in a written mode.

To sum up, the practices and strategies presented thus far enable instructors to help learners process and understand novel disciplinary information delivered in a foreign language while effectively managing intrinsic and extraneous cognitive load, which is of paramount importance in CLIL classes:

> Together, the added intrinsic and extraneous cognitive load determines the total cognitive load imposed on the learner by the learning task. This load determines working memory resources required for processing all the involved elements of information and achieving learning goals by a fully engaged learner. (Kalyuga 2011, p. 207)

1.8.5 Application

To enhance content knowledge, after answering students' questions about the lecture delivered, instructors are encouraged to provide learners with collaborative tasks[7] triggering output, negotiation of meaning, and cognitive engagement liable to activate learners' higher-order thinking skills necessary for students to experience deep learning. As Coyle et al. (2010, p. 39) suggest, deep learning:

> involves the critical analysis of new ideas, connecting them to already-known concepts, and leads to understanding and long-term retention of those concepts so that they can be used for problem solving in unfamiliar contexts.

This practice is in keeping with the European Framework for CLIL Teacher Education that requires «CLIL teachers […] to apply strategies for fostering critical thinking by students about content and language» (Marsh et al. 2011, p. 19). To enhance intake in CLIL settings, students need to be intellectually challenged with activities heightening reasoning skills, hypothesis-making, inquiring, evaluating, problem solving, and creative thinking. As Coyle et al. (2010, p. 54) hold about cognition:

> For CLIL to be effective, it must challenge learners to create new knowledge and develop new skills through reflection and engagement in higher-order as well as lower-order thinking. CLIL is not about the transfer of knowledge from an expert to a novice. CLIL is about allowing individuals to construct their own understandings and be challenged –whatever their age and ability.

7 For an analysis of tasks see Carloni (2014).

In CLIL teaching, practitioners need «to adopt an inquiry-based approach» (Coyle et al. 2010, p. 44) implemented mainly through tasks. As Lyster (2007, pp. 73-74) suggests, in tasks:

> (a) meaning is primary, (b) there is a goal to work towards or a communication problem to solve, (c) there is a relationship with real-world activities, and (d) assessment is in terms of outcomes (e.g. Ellis; Skehan; Nunan).

Problem solving and other kinds of tasks –carried out through various classroom setups, such as pair and group work– can thus be used to foster content and language acquisition through meaning-focused learning:

> learners are occupied with understanding, extending (e.g. through reasoning), or conveying meaning, and cope with language forms as demanded by that process. Attention to language forms is thus not intentional but incidental to perceiving, expressing, and organizing meaning. (Prabhu 1987, p. 27)

CLIL context-reduced tasks, operationalized at CALP level, need to be carefully drafted in order for them to be «linguistically accessible whilst being cognitively demanding» (Coyle et al. 2010, p. 67).

Tasks can promote either divergent or convergent thinking, respectively allowing learners to come up with different solutions or requiring them to find the only solution available. Tasks can also entail both divergent and convergent thinking as long as they are correctly sequenced:

> Divergent thinking refers to a process that generates many questions, ideas, responses or solutions. It is associated with brainstorming and creative thought, that is generating questions and drawing on ideas from different perspectives and many sources (including personal observations and experiences). While divergent thinking involves generating many ideas, the process associated with identifying the best ideas and discarding the weak ones is called convergent thinking. Convergent thinking refers to narrowing down the options based on existing information and analysis, and selecting the best. (Harasim 2012, p. 92)

In brief, content and language intake is concurrently fostered through tasks.

As a follow up in a plenary session, instructors and students can cooperatively carry out a review of the key content concepts dealt with in class. To heighten the effectiveness of the activity, scaffolding can be provided with graphic organizers. Key concepts and their interplay can thus be promoted as well as assessed. Instructors can use Popplet (`www.popplet.com`), a free

web-based tool, to create Web 2.0 visual representations of the concepts analyzed. Popplets –which users can resize, move around on the canvas, connect, and organize as they prefer– can feature text, which simply needs to be typed in, and notes, added with a drawing tool. Files and images saved on the computer can also be uploaded. Contents from Flickr, YouTube, and Google maps can be imported. To make Popplet an effective teaching and learning tool, users can move the whole popplet board to make its elements more easily visible and focused on, and arrange popplets in different ways, such as in columns, rows, vertically, or horizontally. At the end of class, instructors can upload the mind maps generated on the CLIL website to share them with students. Popplets can be exported as PDF or JPEG files and printed. Instructors can also let students further collaborate in popplet creation after class to «engage people through communication, co-production, and sharing» (McLoughlin, Lee 2008, p. 12). Learners can add ideas to each popplet through the comment function to become knowledge builders. Co-construction of knowledge can thus be promoted as the result of the aggregation of student-generated knowledge; learners are thereby given the opportunity to act proactively within an agency framework. Through Popplet, written pushed output can be promoted within a shared editable networked space where «people meet, as nodes on networks, while communicating with others» (Kop, Carroll 2011, p. 4). As McLoughlin and Lee (2008, p. 17) suggest:

> These creations, while enabling personal expression and publication, also allow for social constructivist forms of participation by allowing comments and annotations by others, and, furthermore, by contributing to extant communities of interest by sharing resources. Therefore, not only is this element of Pedagogy 2.0 reflective of the «participation model of learning» (Sfard), as opposed to the 'acquisition' model, but it also adds a further dimension to participative learning by increasing the level of socialization and collaboration with experts, community, and peer groups, and by fostering connections that are often global in reach.

Instructors can turn Popplet mind maps into presentations: to set the desired presentation sequence, they simply need to click on popplets. Instructors can use the co-constructed mind maps to brainstorm ideas at the beginning of the following lesson. Furthermore, the popplets generated at the end of each instructional session can be linked together. Learners can thus access the collaboratively constructed mind map of an entire CLIL module and use it to review for exams.

At the end of the lesson, instructors are encouraged to provide students with brief self-evaluation questionnaires focusing on the daily content objectives. Questionnaires can be administered either in a paper-based format or online through web-based tools, such as Mentimeter.

To sum up, learning units designed with the methodological support of the expert aim to be in keeping with CLIL core features as envisaged and endorsed by Marsh et al. (2011, p. 33):

- Multiple focus
- Safe and enriching learning environments
- Authenticity
- Active learning
- Scaffolding
- Cooperation.

1.9 Classroom observation

Instructors can ask for classroom observation, which, after an initial consultation with the professors, the CLIL expert carries out using a customized grid. A follow-up meeting with the instructors is then scheduled to discuss the findings and provide suggestions to improve CLIL practices, if necessary.

Classroom observation is also carried out by English native speakers, who mainly focus on instructors' classroom management language. Language specialists also observe how students use the foreign language and are available to provide them with language tips.

Form-focused instruction is necessary in CLIL teaching, as it:

> refers to «any pedagogical effort which is used to draw the learners' attention to language form either implicitly or explicitly. This can include the direct teaching of language (e.g. through grammatical rules) and/or reactions to learners' errors (e.g. corrective feedback)» (Lyster 2007, p. 43). […] Form-focused instruction has been portrayed as either 'proactive' or 'reactive'. (Lyster 2007, pp. 44-45)

During face-to-face CLIL lessons, professors are unlikely to provide learners with reactive form-focused instruction, that is, with «corrective feedback as well as other attempts to draw learners' attention to language features in relatively unplanned and spontaneous ways […] during teacher-student interaction» (Lyster 2007, p. 47). Therefore, during classroom observation, language specialists are available to provide learners with corrective feedback through Twitter, where the experts can tweet their corrective feedback. The hashtag for each lesson is communicated to students at the beginning of class. TodaysMeet can also be used for this purpose. This practice is likely to further improve students' acquisition of subject-specific lexical items. Studies carried out by Lapkin, Swain, and Lyster (cited in Lyster 2007, p. 47) showed how «a reactive approach is ideal for pushing students in their lexical choices». Moreover, reactive-focused

approaches, targeted to lead students to notice their mistakes to correct them through self-repair, are necessary to promote learners' output in terms of accuracy:

> Swain's Output Hypothesis purports that learners have to be pushed to make their output more precise, coherent, and appropriate. One prerequisite for this push is that the learners notice the gap between their own output and the target language in terms of structures and pragmatics. (Kost 2008, p. 168)

During classroom observation, learners can also use Twitter or TodaysMeet to ask language specialists for language tips.

2 Corpus-informed teaching materials

Summary 2.1. Corpus-informed teaching materials: an example. – 2.2. Readability. – 2.3. Vocabulary coverage in content-specific texts. – 2.4. Specialized corpora and the investigation of content-specific texts. – 2.5. Content-specific vocabulary distribution. – 2.6. Online corpus-informed activities. – 2.7. Lexical priming: technology-enhanced activities. – 2.8. ConcGrams in a biology text. – 2.9. The Academic Formulas List in a biology text. – 2.10. Data-Driven Learning. – 2.11. Reading academic prose: online tools. – 2.12. Corpus-informed materials and academic skills.

2.1 Corpus-informed teaching materials: an example

The procedures the CLIL expert used to both investigate the English article featured in a CLIL first-year biology course reading list and design CLIL course-tailored corpus-informed materials are presented in this section.[1]

The CLIL expert decided to work on just one article because in most CLIL courses, learners are required to study only one scientific essay in English. Furthermore, it is necessary to work on the lexico-grammatical features of each content-specific study material to retrieve collocational findings useful for customizing course-tailored, corpus-informed, discipline-specific activities, such as pre- reading/listening activities, while- reading/listening activities and language awareness.

2.2 Readability

Foreign students often find it difficult to read academic prose in English: «One of the most common complaints about scientific documents is that they [are] difficult to read because of the ponderous complexity and length of their words, sentences, and paragraphs» (Matthews et al. 2000, p. 107). As such, to make sure that the English scientific article featured in the biology reading list mentioned above was suitable for the CLIL biology students, its readability degree was calculated:

1 For further examples in other subject-specific fields see Carloni (2013c, 2013d).

> the term 'readability' refers to aspects of writing that can be measured and subjected to a formula. Readability formulas such as Fry, Flesch, Fog, and Kincaid are based on the relationships between average word length (or number of syllables) and average sentence length. Various ratios between the two allow a document's readability to be ranked by difficulty or by grade. (Matthews et al. 2000, p. 107)

A tool featured in Microsoft Word was used to calculate the readability scores of the targeted biology text. The Word-retrieved data are shown below (Fig. 14).

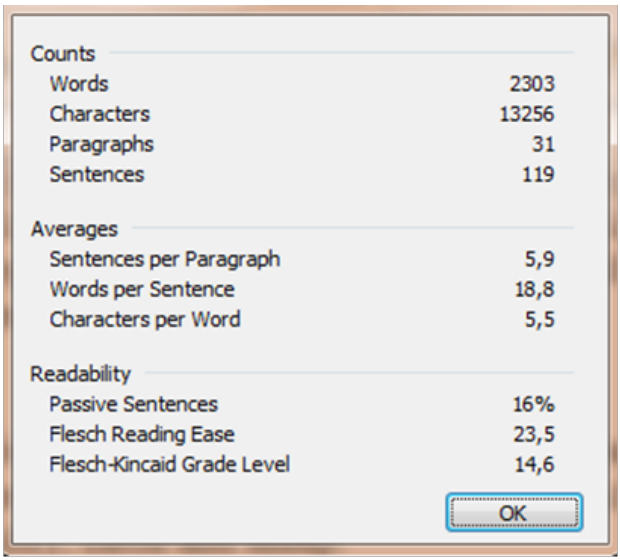


Figure 14. Word-retrieved data

The score of the Flesch-Kincaid Grade Level was dealt with first:

> The Flesch-Kincaid Grade Level index is one way to measure and report the readability of English text. The Flesch-Kincaid formula considers the average number of words per sentence (average sentence length, or ASL) and the average number of syllables per word (ASW) within a given passage in order to estimate the complexity of the text. The formula then converts that complexity level into a score that roughly equates with a grade level (K-12) in the United States. The formula is:
> Flesch-Kincaid Grade Level = (.39 × ASL) + (11.8 × ASW) - 15.59.
> (`http://etc.usf.edu/lit2go/welcome/faq`)

The Flesch-Kincaid Grade Level of the biology article was 14.6, which entailed that English native-speaker undergraduates could understand the text. Owing to the high readability score however, the article was likely to be difficult for non-native, first-year English university speakers.

The data retrieved showed that 16% of the text was composed of passive sentences. The distribution of passive and active verbs (16% vs. 84%) in the biology article was in keeping with Biber's (2006, pp. 64-65) studies: «the written university registers show a great reliance on passive voice: c. 20% passive voice vs. c. 80% active voice». The high percentage of active voice verbs, characterized by the standard SVO (subject, verb, object), made it easier for CLIL students to understand the text. Passive sentences, on the other hand, were likely to make the reading process more difficult by increasing learners' cognitive load. As Matthews et al. (2000, p. 109) suggest in this respect:

> If […] one writes, «A magnet was swallowed by the cow», then the reader unconsciously unscrambles the backward construction, converting it back into SVO order before grasping it fully. This takes extra mental energy and always interferes to some degree with effective communication.

The article featured about 18.8 words per sentence. Sentence length was thus in line with the readability benchmarks specific of English academic writing: «For maximal readability, most sentences in most scientific prose should be about 15 to 20 words» (Matthews et al. 2000, p. 108).

The paragraph length amounting to 110,92 words was calculated by multiplying the number of sentences per paragraph (5.9) by the number of words per sentence (18.8). Paragraph length fell within suitable readability benchmarks: «A paragraph length of about 150 words has been judged to be optimal for a scientific article» (Matthews et al. 2000, p. 109).

The biology article scored 23.5 in terms of Flesch Reading Ease, a formula that measures the readability of English academic texts in particular. On the grounds of this score, the academic prose examined was rated as rather difficult, even for English native speakers. The difficulties were probably due to the discipline-specific content in general and in particular to the high density of subject-specific vocabulary featured in the article. As Matthews et al. (2000, pp. 107-108) suggest with regard to readability formulas:

> Difficulty of content is ignored, as is the recognition factor. Some multisyllable medical terms are easily and immediately recognized, even by general readers. Conversely, highly technical material may use very short words and be easy to read, but still be difficult for any but a few specialists to comprehend.

2.3 Vocabulary coverage in content-specific texts

Further analysis was necessary to identify both content-related features specific to the biology text and possible problems arising from the subject-specific content itself. As a result, the scientific article was investigated in terms of general, academic, and content-specific vocabulary. The analysis of the vocabulary coverage of the text was carried out using the Vocabulary Profiler (`http://www.lextutor.ca`) (Fig. 15).

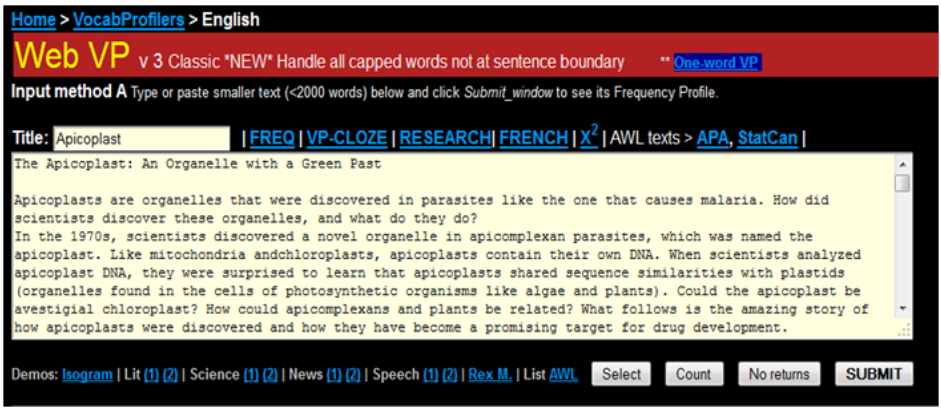


Figure 15. Vocabulary Profiler

The vocabulary profile generated (Fig. 16) showed that 62.48% of the text was made up of the first 1000 most commonly used words of general English. Out of the 62.48% mentioned above, 38.45% were function words while 24.03% were content words. These data were in keeping with research: «Corpus word counts consistently show that function words are among the most frequent in language. […] This holds true regardless of whether the discourse is general in nature, technical or academic» (Schmitt 2010, p. 54). In terms of content words, 9.26% were of Anglo-Saxon origin.

Furthermore, 2.95% of the text was made up of the 1001-2000 most commonly used words in general English; 0.62% of these words were of Anglo-Saxon origin. Overall, the first 2000 most commonly used words of general English accounted for 65.43% of the academic text examined.

Academic words, featured in Coxhead's (2000) AWL, Academic Word List (`http://www.victoria.ac.nz/lals/resources/academicwordlist`), covered 7.67% of the text; 0.44% of academic words were of Anglo-Saxon origin. The percentage of academic language featured in the article was in keeping with research: «Typically these words make up about 9/10% of the running words in an academic text» (Schmitt 2010, p. 78). Off-list words, which were likely to be mainly content-specific vocabulary items, accounted for 26.90% of the text.

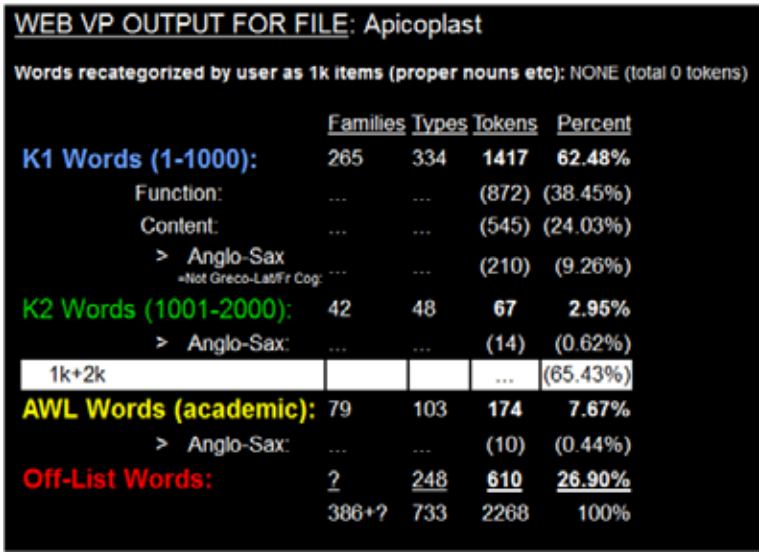


Figure 16. Vocabulary profile of the biology article

The distribution of general, academic, and off-list vocabulary across the text was also accessed through a colour-coded layout (Fig. 17).

Figure 17. Colour-coded vocabulary profile of the biology article

Red colour-coded off-list words, which could also be accessed as a list (Fig. 18), were almost exclusively subject-specific words.

Figure 18. Off-list words of the biology article

Yellow colour-coded academic words were grouped together. Furthermore, they were sub-classified according to the AWL sublist they belonged to (Fig. 19).

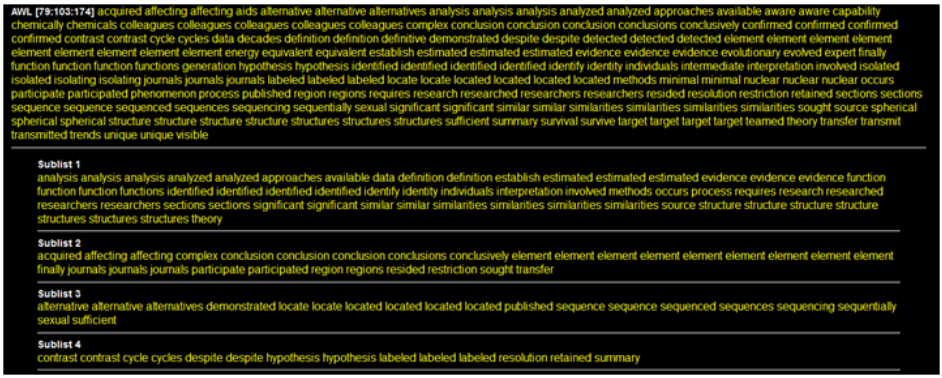


Figure 19. AWL vocabulary featured in the biology article

AWL head words retrieved from the biology text could also be accessed in the following format (Fig. 20).

AWL complete: Found headwords underlined	AWL head-words found	AWL head-words not found
1. ABANDON	1.	1. ABANDON
2. ABSTRACT	2.	2. ABSTRACT
3. ACADEMY	3.	3. ACADEMY
4. ACCESS	4.	4. ACCESS
5. ACCOMMODATE	5.	5. ACCOMMODATE
6. ACCOMPANY	6.	6. ACCOMPANY
7. ACCUMULATE	7.	7. ACCUMULATE
8. ACCURATE	8.	8. ACCURATE
9. ACHIEVE	9.	9. ACHIEVE
10. ACKNOWLEDGE	10.	10. ACKNOWLEDGE
11. ACQUIRE	11. ACQUIRE	11.
12. ADAPT	12.	12. ADAPT
13. ADEQUATE	13.	13. ADEQUATE
14. ADJACENT	14.	14. ADJACENT
15. ADJUST	15.	15. ADJUST
16. ADMINISTRATE	16.	16. ADMINISTRATE
17. ADULT	17.	17. ADULT
18. ADVOCATE	18.	18. ADVOCATE
19. AFFECT	19. AFFECT	19.
20. AGGREGATE	20.	20. AGGREGATE
21. AID	21. AID	21.
22. ALBEIT	22.	22. ALBEIT
23. ALLOCATE	23.	23. ALLOCATE
24. ALTER	24.	24. ALTER
25. ALTERNATIVE	25. ALTERNATIVE	25.
26. AMBIGUOUS	26.	26. AMBIGUOUS
27. AMEND	27.	27. AMEND
28. ANALOGY	28.	28. ANALOGY
29. ANALYSE	29. ANALYSE	29.

Figure 20. AWL head words retrieved from the biology article

Familizer (`http://www.lextutor.ca/familizer`), another tool of Compleat Lexical Tutor, was used to «build family word lists of technical terms drawn from a specialist corpus» (`http://www.lextutor.ca/familizer`). Each content-specific word was expanded into its word family; thus, family word lists of each technical lexical item were generated. Examples of the family word list of the biology article featured in the 'family in lines' format are provided below:

algae algaes algael;
bacterium bacteria bacterial;
biology biologies biologist biologists biological biologically;
biosynthetic biosynthetically biosynthesis;
cryptosporidia cryptosporidiosis cryptosporidioses;

cryptosporidium cryptosporidia;
cyanobacteria cyanobacterium;
parasite parasites parasitic parasitically parasitism;
ribosome ribosomes ribosomal.

Examples of the family word list of the content-specific text presented in the 'family in tabs' format are shown below:

chemical
chemically
chemicals
gradient
gradients
granule
granules
granular
hybrid
hybridism
hybridize
hybridizes
hybridized
hybridizing
hybridization

The analysis of the vocabulary coverage of the biology article allowed the CLIL expert to examine more thoroughly the reading materials the students were expected to study. The data gathered were then used to create course-tailored reading comprehension and language awareness activities. In particular, the findings generated through Familizer were instrumental in creating activities aimed at enhancing students' content knowledge in terms of subject-specific lexical sets.

2.4 Specialized corpora and the investigation of content-specific texts

The CLIL expert used Sketch Engine to further investigate the targeted English biology article. First of all, a Do-It-Yourself (DIY) Part Of Speech (POS) tagged specialized corpus, named 'Apicoplast', was created using the English biology text featured in the CLIL biology course reading list. The intercollocability of the content-specific text was thus thoroughly analyzed:

> Intercollocability is determined by generating, in a fully automated form, a text's collocational profile. This is done by first establishing keywords and the collocates of the keywords, and then the collocates of collocates

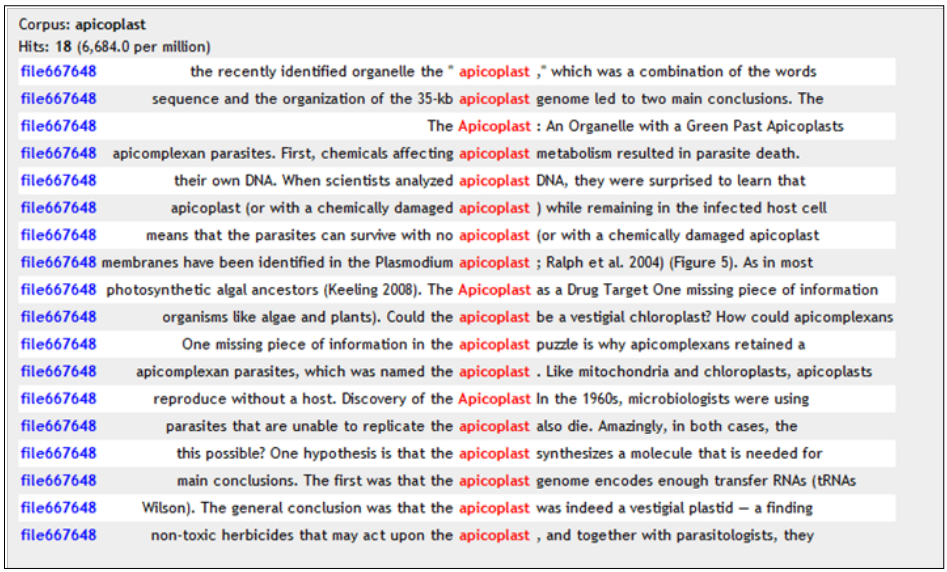
> to determine groups of words, which in turn determine what Sinclair terms the aboutness of the text. (Cheng 2009, p. 161)

A keyword list of the targeted biology article was first retrieved. In the keyword list of the biology text, the most commonly used discipline-specific lexical items were easily detected and the keyness of the article determined. As figure 21 shows, technical terms, such as 'apicoplast' and 'extrachromosomal', were ranked as the most commonly used content-specific words followed by (a) nouns such as 'plastid', 'toxoplasma', 'organelle', 'plasmodium', 'apicomplexan', 'genome', 'mitochondria', 'parasite' and (b) adjectives such as 'extrachromosomal' and 'mitochondrial'. In this context, it is important to note that most of the keywords retrieved were nouns: as previously mentioned, nominalization is a key feature of scientific writing.

	apicoplast		*British National Corpus (TreeTagger)*		
word	Freq	Freq/mill	Freq	Freq/mill	Score
apicoplast	15.0	5570.0	0	0.0	5571.0
apicoplasts	10.0	3713.3	0	0.0	3714.3
extrachromosomal	10.0	3713.3	4.0	0.0	3587.3
Kilejian	8.0	2970.7	0	0.0	2971.7
plastid	8.0	2970.7	1.0	0.0	2945.6
Toxoplasma	8.0	2970.7	5.0	0.0	2845.7
organelle	7.0	2599.3	6.0	0.1	2469.2
Plasmodium	7.0	2599.3	17.0	0.2	2260.2
apicomplexan	6.0	2228.0	0	0.0	2229.0
genomes	8.0	2970.7	40.0	0.4	2194.7
mitochondria	10.0	3713.3	83.0	0.7	2141.3
parasites	26.0	9654.7	406.0	3.6	2102.1
mitochondrial	10.0	3713.3	103.0	0.9	1943.0
apicomplexans	5.0	1856.7	0	0.0	1857.7
Eimeria	5.0	1856.7	0	0.0	1857.7
genome	11.0	4084.7	185.0	1.6	1549.1
organelles	5.0	1856.7	25.0	0.2	1521.1
DNAs	7.0	2599.3	81.0	0.7	1514.5
McFadden	5.0	1856.7	34.0	0.3	1428.0
Parasitology	4.0	1485.3	16.0	0.1	1302.0

Figure 21. Keyword list of the 'Apicoplast corpus'

The top ranking keyword 'apicoplast' was used as a query term to generate concordances. Through left-sorted concordance lines of 'apicoplast' (Fig. 22), the CLIL expert identified some domain-specific collocations, such as 'chemically damaged apicoplast', '35-kb apicoplast genome' and 'plasmodium apicoplast'.

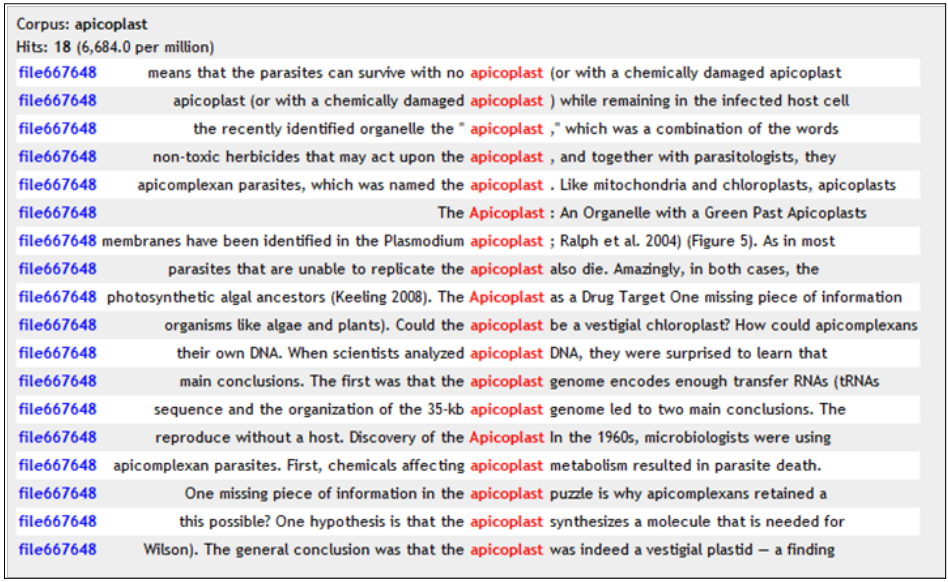


Corpus: apicoplast
Hits: 18 (6,684.0 per million)

file667648	the recently identified organelle the "	apicoplast	," which was a combination of the words
file667648	sequence and the organization of the 35-kb	apicoplast	genome led to two main conclusions. The
file667648	The	Apicoplast	: An Organelle with a Green Past Apicoplasts
file667648	apicomplexan parasites. First, chemicals affecting	apicoplast	metabolism resulted in parasite death.
file667648	their own DNA. When scientists analyzed	apicoplast	DNA, they were surprised to learn that
file667648	apicoplast (or with a chemically damaged	apicoplast	) while remaining in the infected host cell
file667648	means that the parasites can survive with no	apicoplast	(or with a chemically damaged apicoplast
file667648	membranes have been identified in the Plasmodium	apicoplast	; Ralph et al. 2004) (Figure 5). As in most
file667648	photosynthetic algal ancestors (Keeling 2008). The	Apicoplast	as a Drug Target One missing piece of information
file667648	organisms like algae and plants). Could the	apicoplast	be a vestigial chloroplast? How could apicomplexans
file667648	One missing piece of information in the	apicoplast	puzzle is why apicomplexans retained a
file667648	apicomplexan parasites, which was named the	apicoplast	. Like mitochondria and chloroplasts, apicoplasts
file667648	reproduce without a host. Discovery of the	Apicoplast	In the 1960s, microbiologists were using
file667648	parasites that are unable to replicate the	apicoplast	also die. Amazingly, in both cases, the
file667648	this possible? One hypothesis is that the	apicoplast	synthesizes a molecule that is needed for
file667648	main conclusions. The first was that the	apicoplast	genome encodes enough transfer RNAs (tRNAs
file667648	Wilson). The general conclusion was that the	apicoplast	was indeed a vestigial plastid — a finding
file667648	non-toxic herbicides that may act upon the	apicoplast	, and together with parasitologists, they

Figure 22. Left-sorted concordances of 'apicoplast'

On the other hand, through right-sorted concordances of 'apicoplast' (Fig. 23), collocations such as 'apicoplast DNA', 'apicoplast genome' and 'apicoplast metabolism' were detected. In this case, the query item, namely the noun 'apicoplast', worked as a premodifier of noun phrases, which is a very common phenomenon in academic texts: «[n]ouns as pre-modifiers [are] very common in academic prose» (Biber 2006, p. 15). Furthermore, in this context it is pivotal to mention that, as Biber (2006, p. 15) suggests, «60% of all noun phrases in academic prose have a modifier».



Corpus: apicoplast
Hits: 18 (6,684.0 per million)

file667648	means that the parasites can survive with no	apicoplast	(or with a chemically damaged apicoplast
file667648	apicoplast (or with a chemically damaged	apicoplast	) while remaining in the infected host cell
file667648	the recently identified organelle the "	apicoplast	," which was a combination of the words
file667648	non-toxic herbicides that may act upon the	apicoplast	, and together with parasitologists, they
file667648	apicomplexan parasites, which was named the	apicoplast	. Like mitochondria and chloroplasts, apicoplasts
file667648	The	Apicoplast	: An Organelle with a Green Past Apicoplasts
file667648	membranes have been identified in the Plasmodium	apicoplast	; Ralph et al. 2004) (Figure 5). As in most
file667648	parasites that are unable to replicate the	apicoplast	also die. Amazingly, in both cases, the
file667648	photosynthetic algal ancestors (Keeling 2008). The	Apicoplast	as a Drug Target One missing piece of information
file667648	organisms like algae and plants). Could the	apicoplast	be a vestigial chloroplast? How could apicomplexans
file667648	their own DNA. When scientists analyzed	apicoplast	DNA, they were surprised to learn that
file667648	main conclusions. The first was that the	apicoplast	genome encodes enough transfer RNAs (tRNAs
file667648	sequence and the organization of the 35-kb	apicoplast	genome led to two main conclusions. The
file667648	reproduce without a host. Discovery of the	Apicoplast	In the 1960s, microbiologists were using
file667648	apicomplexan parasites. First, chemicals affecting	apicoplast	metabolism resulted in parasite death.
file667648	One missing piece of information in the	apicoplast	puzzle is why apicomplexans retained a
file667648	this possible? One hypothesis is that the	apicoplast	synthesizes a molecule that is needed for
file667648	Wilson). The general conclusion was that the	apicoplast	was indeed a vestigial plastid — a finding

Figure 23. Right-sorted concordances of 'apicoplast'

To generate the lexical profiling of targeted query terms, Word Sketches, a tool of Sketch Engine, was consistently used: «Word sketches are one-page automatic, corpus-based summaries of a word's grammatical and collocational behaviour» (Kilgarriff et al. 2004). In Word Sketches, collocates are classified on the basis of their grammatical relations with the query term:

> The word sketch […] provides one list of collocates for each grammatical relation the word participates in. For a verb, the subject, the objects, the conjoined verbs (stand and deliver, hope and pray), modifying adverbs, prepositions and prepositional objects, are all presented in different lists. (Kilgarriff et al. 2004)

Word Sketches allowed the CLIL expert to retrieve important collocational distributional data which were extremely useful in materials design:

> Word Sketches (cf. Kilgariff, Rundell) produces a kind of statistical summary which reveals the grammatical and collocational behaviour of a word. For example, a Word Sketch for the verb 'exercise' reveals that the kinds of objects exercise usually takes are words like restraint, discretion, caution, and vigilance (Atkins, Rundell). (Flowerdew 2012, p. 179)

The Word Sketch for 'apicoplast' showed (Fig. 24a) that the query item worked as (a) the object of verbs such as 'damage', 'replicate', 'name'; (b)

the subject of verbs such as 'synthesize', 'die' and 'be' and (c) the premodifier of nouns such as 'genome', 'metabolism' and 'DNA'. Furthermore, the Word Sketch showed that 'apicoplast' was premodified by a noun, namely 'plasmodium': this co-occurrence is a key feature of academic prose, as previously mentioned.

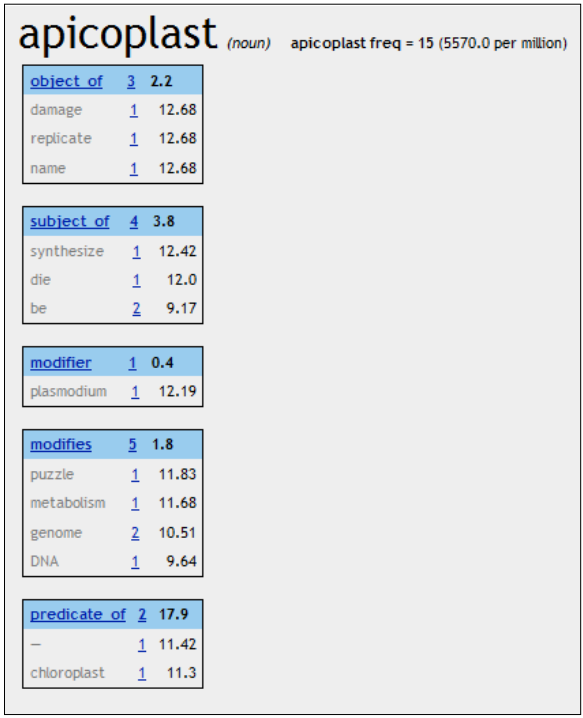


Figure 24a. Word Sketch for 'apicoplast'

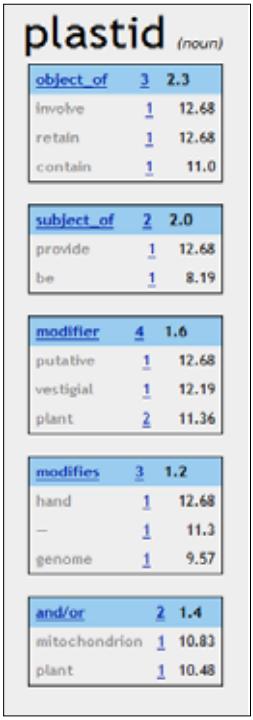


Figure 24b. Word Sketch for 'plastid'

'Plastid' – another top ranking keyword – was examined through Word Sketches (Fig. 24b). Word Sketch-retrieved data showed that 'plastid' worked as (a) the object of verbs such as 'involve', 'retain' and 'contain', (b) the subject of verbs such as 'provide' and 'be' and (c) the pre-modifier of nouns, such as 'hand' and 'genome'. Furthermore, the query term was premodified by attributive adjectives, such as 'putative' and 'vestigial', and nouns, such as 'plant'. Additionally, 'plastid' was identified as a component of a nominal binomial in connection with 'mitochondrion', namely 'mitochondrion and apicoplast'. As Biber (2006) holds, nominal binomials are most common in academic prose.

The biology article was also searched for adverbs. The data retrieved (Fig. 25) showed that overall adjectives were more widely used than adverbs. These data were in keeping with Biber's (2006, p. 47) findings

showing that, with regard to academic prose, «adjectives are used more commonly in the written registers, while adverbs are favoured in the spoken registers».

unable to replicate the apicoplast also die.	Amazingly	, in both cases, the parasites only die
can survive with no apicoplast (or with a	chemically	damaged apicoplast) while remaining in
mitochondria were not labeled. These experiments	clearly	showed that the 35-kb genome was located
another six years before the 35-kb DNA was	conclusively	shown to associate with the "spherical
apicomplexan" and "plastid" (Köhler et al. 1997).	Finally	, the new organelle had a name. Since then
genomes?" (Wilson et al. 1991). This title was	intentionally	provocative. At the time, biologists agreed
clinical cases and a million deaths in 2008,	mainly	in sub-Saharan Africa (Figure 1). In contrast
hosts simply by accident. The disease is	only	serious (even fatal) in individuals with
Köhler and colleagues proposed naming the	recently	identified organelle the "apicoplast,"
, micronemes, and dense granules), which	sequentially	secrete enzymes that allow the parasite
typically read journals on parasitology. Quite	serendipitously	, they came across Wilson's paper in Parasitology
felines, and humans are intermediate hosts	simply	by accident. The disease is only serious
when a second extrachromosomal genome — a	tandemly	arranged, 6-kb, linear DNA molecule — was
Waller were plant scientists, they did not	typically	read journals on parasitology. Quite serendipitously

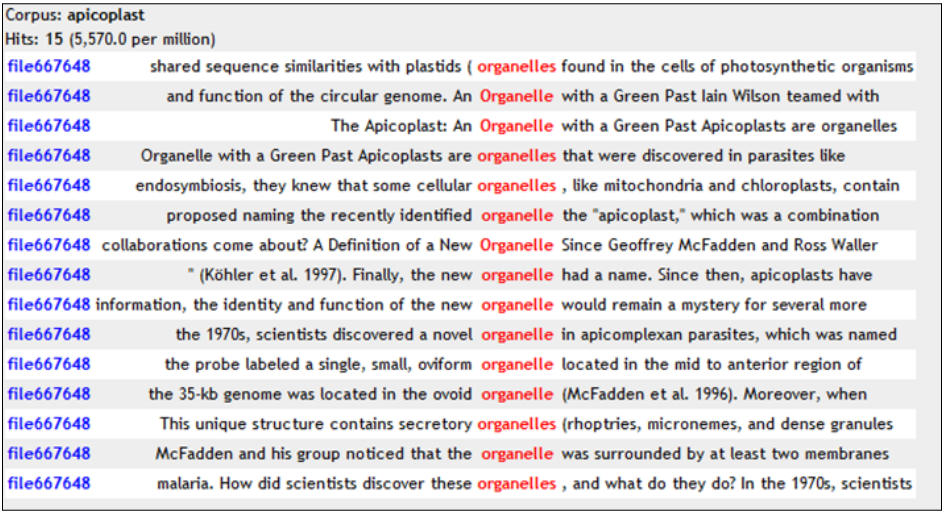
Figure 25. Adverbs retrieved from the 'Apicoplast corpus'

With regard to nominal binomials, further examples were retrieved, such as 'algae and plants' (Fig. 26).

to the parasite's survival. In algae and	plants	, plastids are not only involved in photosynthesis
photosynthetic organisms like algae and	plants	). Could the apicoplast be a vestigial chloroplast
well aware of the dogma that only algae and	plants	have three genomes. So, when they read

Figure 26. Nominal binomials

Concordances were retrieved for 'organelle', another technical term which was highly ranked in the keyword list. Left-sorted concordances of 'organelle' (Fig. 27) showed collocations such as 'cellular organelles', 'new organelle', 'novel organelle', 'oviform organelle', 'ovoid organelle' and 'secretory organelles'. The query item was thus consistently premodified by attributive adjectives. These results were in keeping with Biber's (2006) studies holding that adjectives – and attributive adjectives in particular – are distinctive of academic prose.



Corpus: apicoplast
Hits: 15 (5,570.0 per million)

file667648	shared sequence similarities with plastids (	organelles	found in the cells of photosynthetic organisms
file667648	and function of the circular genome. An	Organelle	with a Green Past Iain Wilson teamed with
file667648	The Apicoplast: An	Organelle	with a Green Past Apicoplasts are organelles
file667648	Organelle with a Green Past Apicoplasts are	organelles	that were discovered in parasites like
file667648	endosymbiosis, they knew that some cellular	organelles	, like mitochondria and chloroplasts, contain
file667648	proposed naming the recently identified	organelle	the "apicoplast," which was a combination
file667648	collaborations come about? A Definition of a New	Organelle	Since Geoffrey McFadden and Ross Waller
file667648	" (Köhler et al. 1997). Finally, the new	organelle	had a name. Since then, apicoplasts have
file667648	information, the identity and function of the new	organelle	would remain a mystery for several more
file667648	the 1970s, scientists discovered a novel	organelle	in apicomplexan parasites, which was named
file667648	the probe labeled a single, small, oviform	organelle	located in the mid to anterior region of
file667648	the 35-kb genome was located in the ovoid	organelle	(McFadden et al. 1996). Moreover, when
file667648	This unique structure contains secretory	organelles	(rhoptries, micronemes, and dense granules
file667648	McFadden and his group noticed that the	organelle	was surrounded by at least two membranes
file667648	malaria. How did scientists discover these	organelles	, and what do they do? In the 1970s, scientists

Figure 27. Left-sorted concordances of 'organelle'

On the other hand, the Word Sketch for 'organelle' (Fig. 28) showed that the query item worked as (a) the object of verbs such as 'discover', 'surround', 'find', 'identify', 'locate' and 'contain' and (b) the subject of verbs such as 'have' and 'be'. Thus, through Word Sketch-generated data, new information was retrieved.

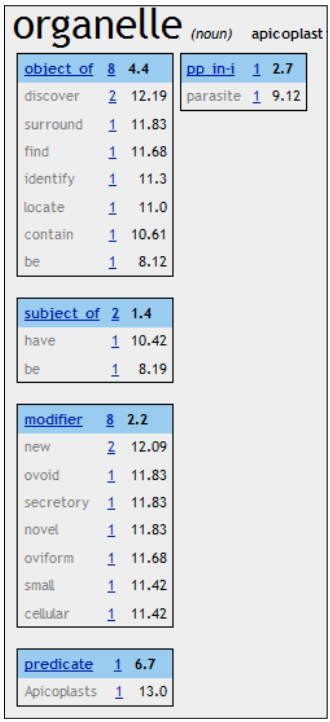


Figure 28. Word Sketch for 'organelle'

Right-sorted concordances of 'plasmodium' (Fig. 29) highlighted collocations such as 'plasmodium apicoplast', 'plasmodium parasites' and 'plasmodium species'. Nouns as premodifiers (i.e. 'plastid apicoplast'), widely used in academic prose as previously suggested, were thus consistently detected in the biology article investigated.

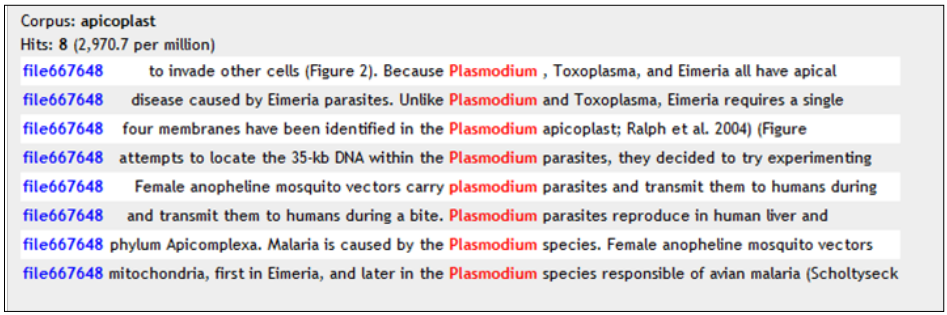
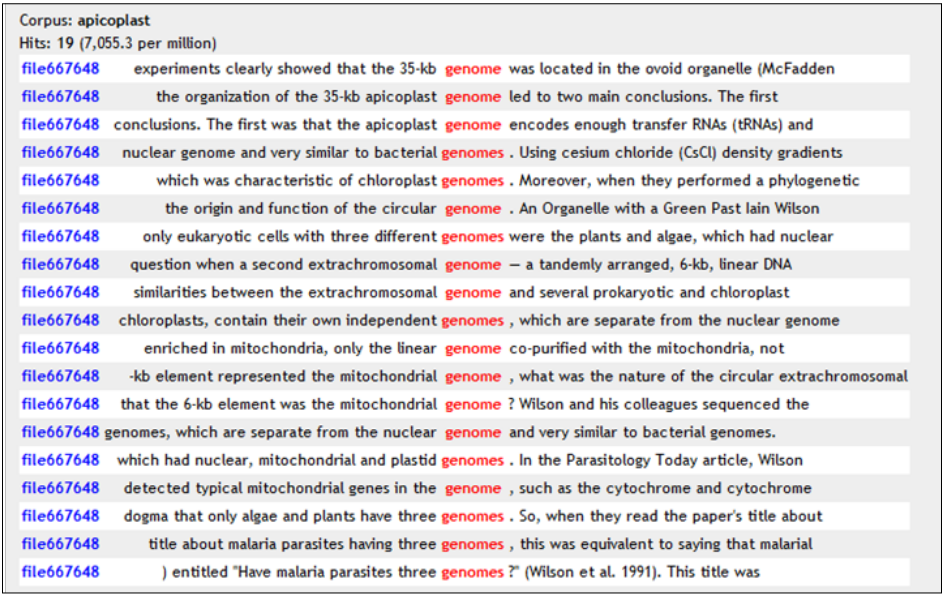


Figure 29. Right-sorted concordances of 'plasmodium'

Right-sorted concordances of ‘apicomplexan’ (Fig. 30) showed that the query item collocated mainly with ‘parasite’. Furthermore, various participial clauses working as postmodifiers in noun phrases – such as ‘apicomplexan parasites evolved from’ and ‘apicomplexan parasites studied’ – were identified. This piece of data was in keeping with Biber’s (2006, p. 17) studies stating that «[p]articipial clauses as post-modifiers in noun phrases [are] very common in academic prose».

Figure 30. Right-sorted concordances of ‘apicomplexan’

Left-sorted concordances of ‘genome’ (Fig. 31) retrieved collocations such as ‘35-kb genome’, ‘35-kb apicoplast genome’, ‘bacterial genome’, ‘chloroplast genome, circular genome’, ‘extrachromosomal genome’, ‘independent genome’, ‘linear genome’, ‘mitochondrial genome’, ‘nuclear genome’ and ‘plastid genome’. Quite a few derived adjectives ending in «-al» – such as ‘bacterial’, ‘extrachromosomal’ and ‘mitochondrial’ – used in attributive positions were identified. In this respect, it is noteworthy to mention that, as Biber (2006) suggests, derived adjectives and especially those ending in «-al» are much more widely used in academic prose. Thus, another important feature of academic writing was found in the biology text examined.


Corpus: apicoplast
Hits: 19 (7,055.3 per million)

file667648	experiments clearly showed that the 35-kb	genome	was located in the ovoid organelle (McFadden
file667648	the organization of the 35-kb apicoplast	genome	led to two main conclusions. The first
file667648	conclusions. The first was that the apicoplast	genome	encodes enough transfer RNAs (tRNAs) and
file667648	nuclear genome and very similar to bacterial	genomes	. Using cesium chloride (CsCl) density gradients
file667648	which was characteristic of chloroplast	genomes	. Moreover, when they performed a phylogenetic
file667648	the origin and function of the circular	genome	. An Organelle with a Green Past Iain Wilson
file667648	only eukaryotic cells with three different	genomes	were the plants and algae, which had nuclear
file667648	question when a second extrachromosomal	genome	— a tandemly arranged, 6-kb, linear DNA
file667648	similarities between the extrachromosomal	genome	and several prokaryotic and chloroplast
file667648	chloroplasts, contain their own independent	genomes	, which are separate from the nuclear genome
file667648	enriched in mitochondria, only the linear	genome	co-purified with the mitochondria, not
file667648	-kb element represented the mitochondrial	genome	, what was the nature of the circular extrachromosomal
file667648	that the 6-kb element was the mitochondrial	genome	? Wilson and his colleagues sequenced the
file667648	genomes, which are separate from the nuclear	genome	and very similar to bacterial genomes.
file667648	which had nuclear, mitochondrial and plastid	genomes	. In the Parasitology Today article, Wilson
file667648	detected typical mitochondrial genes in the	genome	, such as the cytochrome and cytochrome
file667648	dogma that only algae and plants have three	genomes	. So, when they read the paper's title about
file667648	title about malaria parasites having three	genomes	, this was equivalent to saying that malarial
file667648	) entitled "Have malaria parasites three	genomes	?" (Wilson et al. 1991). This title was

Figure 31. Left-sorted concordances of 'genome'

In the Word Sketch for 'genome' (Fig. 32), it emerged that the query term worked as (a) the object of verbs such as 'represent', 'co-purify', 'have', 'locate', 'contain' and 'be', and as (b) the subject of verbs such as 'encode', 'lead' and 'be'.

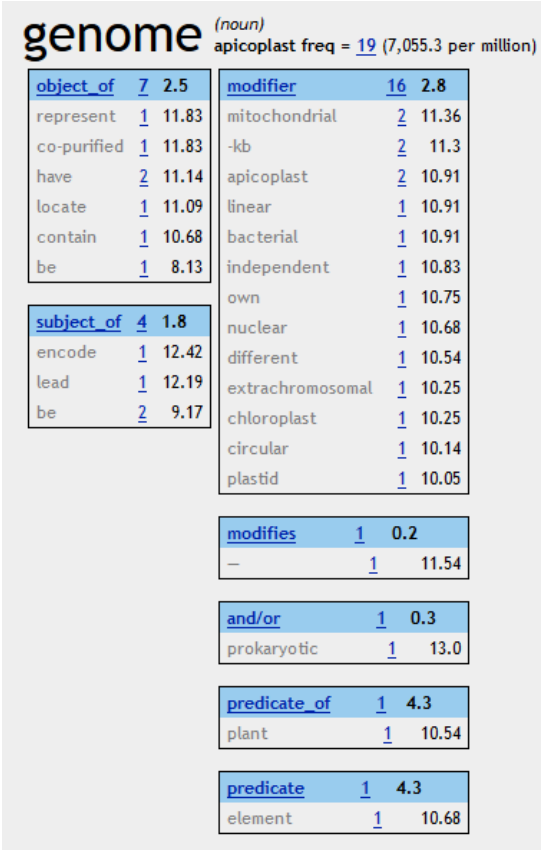
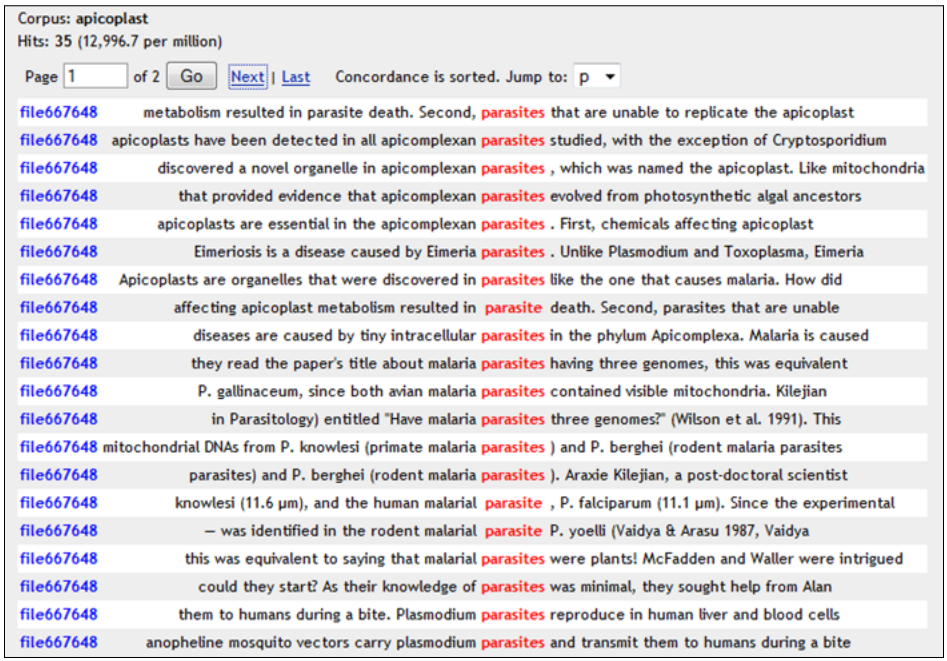


Figure 32. Word Sketch for 'genome'

Left-sorted concordances of 'parasite' (Figs. 33 and 34) highlighted collocations such as 'apicomplexan parasite', 'tiny intracellular parasites', 'avian malaria parasites', 'primate malaria parasites', 'rodent malaria parasites', 'human malarial parasites', 'rodent malarial parasites', 'plasmodium parasites', 'protozoan parasites' and 'toxoplasma parasites'. In this context, it is worth noting that quite a few 3-word units were detected, such as 'avian malaria parasites', 'primate malaria parasites', 'rodent malaria parasites', 'human malarial parasites' and 'rodent malarial parasites'. Furthermore, a wide array of nouns and adjectives working as premodifiers of the query item was identified, which was in keeping with Biber's (2006) previously mentioned research claiming that nouns are widely premodified by other nouns and attributive adjectives in academic writing.

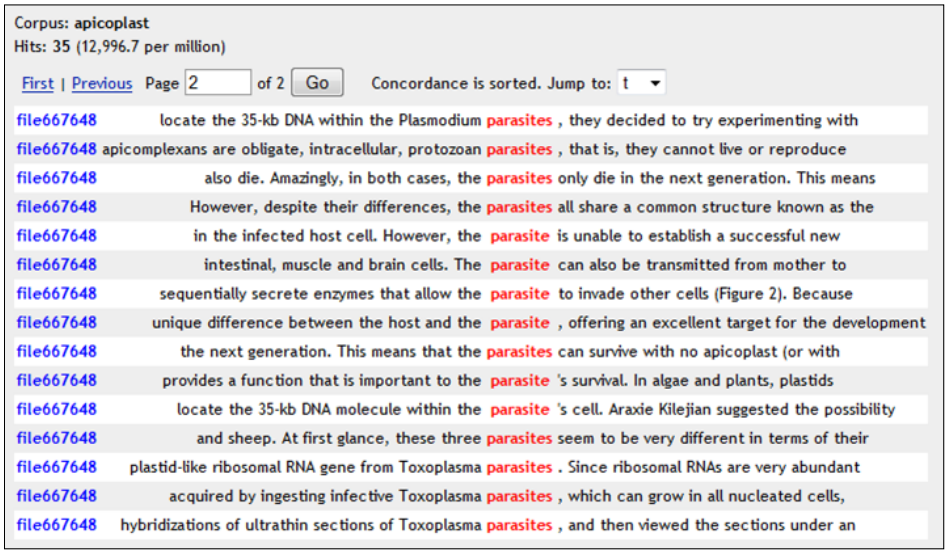


Corpus: apicoplast
Hits: 35 (12,996.7 per million)

Page 1 of 2 Go Next | Last Concordance is sorted. Jump to: p

file667648	metabolism resulted in parasite death. Second,	parasites	that are unable to replicate the apicoplast
file667648	apicoplasts have been detected in all apicomplexan	parasites	studied, with the exception of Cryptosporidium
file667648	discovered a novel organelle in apicomplexan	parasites	, which was named the apicoplast. Like mitochondria
file667648	that provided evidence that apicomplexan	parasites	evolved from photosynthetic algal ancestors
file667648	apicoplasts are essential in the apicomplexan	parasites	. First, chemicals affecting apicoplast
file667648	Eimeriosis is a disease caused by Eimeria	parasites	. Unlike Plasmodium and Toxoplasma, Eimeria
file667648	Apicoplasts are organelles that were discovered in	parasites	like the one that causes malaria. How did
file667648	affecting apicoplast metabolism resulted in	parasite	death. Second, parasites that are unable
file667648	diseases are caused by tiny intracellular	parasites	in the phylum Apicomplexa. Malaria is caused
file667648	they read the paper's title about malaria	parasites	having three genomes, this was equivalent
file667648	P. gallinaceum, since both avian malaria	parasites	contained visible mitochondria. Kilejian
file667648	in Parasitology) entitled "Have malaria	parasites	three genomes?" (Wilson et al. 1991). This
file667648	mitochondrial DNAs from P. knowlesi (primate malaria	parasites	) and P. berghei (rodent malaria parasites
file667648	parasites) and P. berghei (rodent malaria	parasites	). Araxie Kilejian, a post-doctoral scientist
file667648	knowlesi (11.6 µm), and the human malarial	parasite	, P. falciparum (11.1 µm). Since the experimental
file667648	— was identified in the rodent malarial	parasite	P. yoelli (Vaidya & Arasu 1987, Vaidya
file667648	this was equivalent to saying that malarial	parasites	were plants! McFadden and Waller were intrigued
file667648	could they start? As their knowledge of	parasites	was minimal, they sought help from Alan
file667648	them to humans during a bite. Plasmodium	parasites	reproduce in human liver and blood cells
file667648	anopheline mosquito vectors carry plasmodium	parasites	and transmit them to humans during a bite

Figure 33. Left-sorted concordances of 'parasite' [a]



Corpus: apicoplast
Hits: 35 (12,996.7 per million)

First | Previous Page 2 of 2 Go Concordance is sorted. Jump to: t

file667648	locate the 35-kb DNA within the Plasmodium	parasites	, they decided to try experimenting with
file667648	apicomplexans are obligate, intracellular, protozoan	parasites	, that is, they cannot live or reproduce
file667648	also die. Amazingly, in both cases, the	parasites	only die in the next generation. This means
file667648	However, despite their differences, the	parasites	all share a common structure known as the
file667648	in the infected host cell. However, the	parasite	is unable to establish a successful new
file667648	intestinal, muscle and brain cells. The	parasite	can also be transmitted from mother to
file667648	sequentially secrete enzymes that allow the	parasite	to invade other cells (Figure 2). Because
file667648	unique difference between the host and the	parasite	, offering an excellent target for the development
file667648	the next generation. This means that the	parasites	can survive with no apicoplast (or with
file667648	provides a function that is important to the	parasite	's survival. In algae and plants, plastids
file667648	locate the 35-kb DNA molecule within the	parasite	's cell. Araxie Kilejian suggested the possibility
file667648	and sheep. At first glance, these three	parasites	seem to be very different in terms of their
file667648	plastid-like ribosomal RNA gene from Toxoplasma	parasites	. Since ribosomal RNAs are very abundant
file667648	acquired by ingesting infective Toxoplasma	parasites	, which can grow in all nucleated cells,
file667648	hybridizations of ultrathin sections of Toxoplasma	parasites	, and then viewed the sections under an

Figure 34. Left-sorted concordances of 'parasite' [b]

Through the Word Sketch for parasite (Fig. 35), other important distributional patterns emerged. In particular, the data retrieved showed that parasite worked as the object of verbs such as evolve and carry, the subject of verbs such as reproduce, die, cause, and contain, and as the premodifier of death.

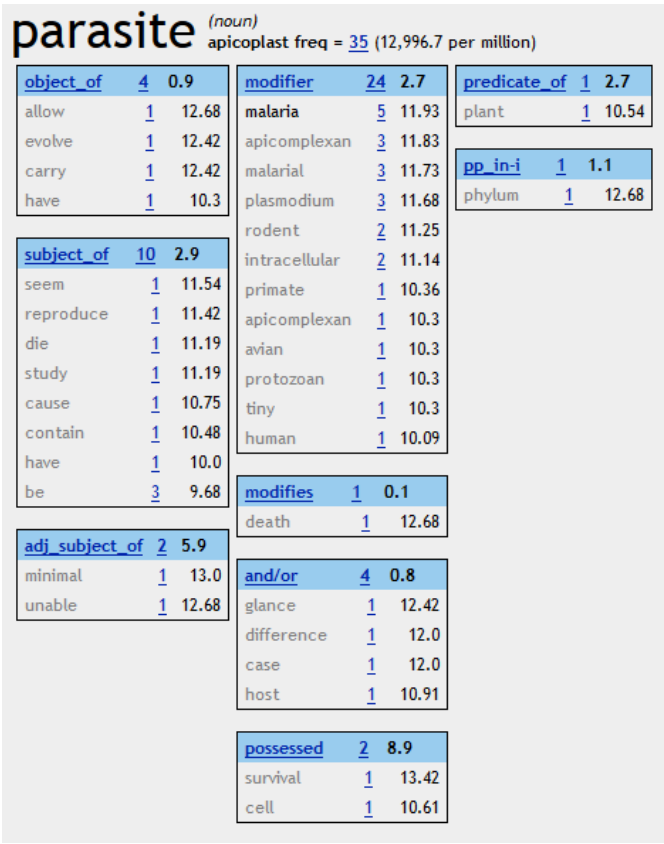


Figure 35. Word Sketch for 'parasite'

As for the query term 'mitochondrial', right-sorted concordances (Fig. 36) showed collocations such as 'mitochondrial genomes', 'mitochondrial DNA' and 'mitochondrial genes'. Furthermore, 'mitochondrial' was identified as part of an adjectival binomial expression, namely 'mitochondrial and plastic genomes'.

Corpus: **apicoplast**
Hits: 10 (3,713.3 per million)

file667648	the plants and algae, which had nuclear, **mitochondrial** and plastid genomes. In the Parasitology
file667648	confirmed that the 6-kb element was the real **mitochondrial** DNA. How did they prove that the 6-kb element
file667648	the circular, extrachromosomal DNAs were **mitochondrial** DNA until late in the 1980s. In 1987, the
file667648	Kilejian mistook these circular structures for **mitochondrial** DNAs. Later, several groups confirmed her
file667648	scientists isolated what they believed to be **mitochondrial** DNAs from P. knowlesi (primate malaria
file667648	intrigued by the results. She wondered why **mitochondrial** DNAs had not been detected in P. lophurae
file667648	the ring" dogma was that most, if not all, **mitochondrial** DNAs were circular, so Kilejian mistook
file667648	sequenced the element and detected typical **mitochondrial** genes in the genome, such as the cytochrome
file667648	. Since the 6-kb element represented the **mitochondrial** genome, what was the nature of the circular
file667648	they prove that the 6-kb element was the **mitochondrial** genome? Wilson and his colleagues sequenced

Figure 36. Right-sorted concordances of mitochondrial

Sketch Differences is a tool of Sketch Engine «which specif[ies], for two semantically related words, what behaviour they share and how they differ» (Kilgarriff et al. 2004). A Sketch Difference for 'organelle' and 'mitochondrion' was generated (Fig. 37).

organelle/mitochondrion

apicoplast freqs = 12 /11

organelle	6.0	4.0	2.0	0	-2.0	-4.0	-6.0	mitochondrion

object_of	8	2	4.4	2.0
label	0	1	0.0	12.4
contain	1	1	10.6	11.1
be	1	0	8.1	0.0
locate	1	0	11.0	0.0
identify	1	0	11.3	0.0
find	1	0	11.7	0.0
surround	1	0	11.8	0.0
discover	2	0	12.2	0.0

subject_of	2	2	1.4	2.6
be	1	2	8.2	9.2
have	1	0	10.4	0.0

modifier	8	1	2.2	0.5
visible	0	1	0.0	14.0
cellular	1	0	11.4	0.0
small	1	0	11.4	0.0
oviform	1	0	11.7	0.0
novel	1	0	11.8	0.0
secretory	1	0	11.8	0.0
ovoid	1	0	11.8	0.0
new	2	0	12.1	0.0

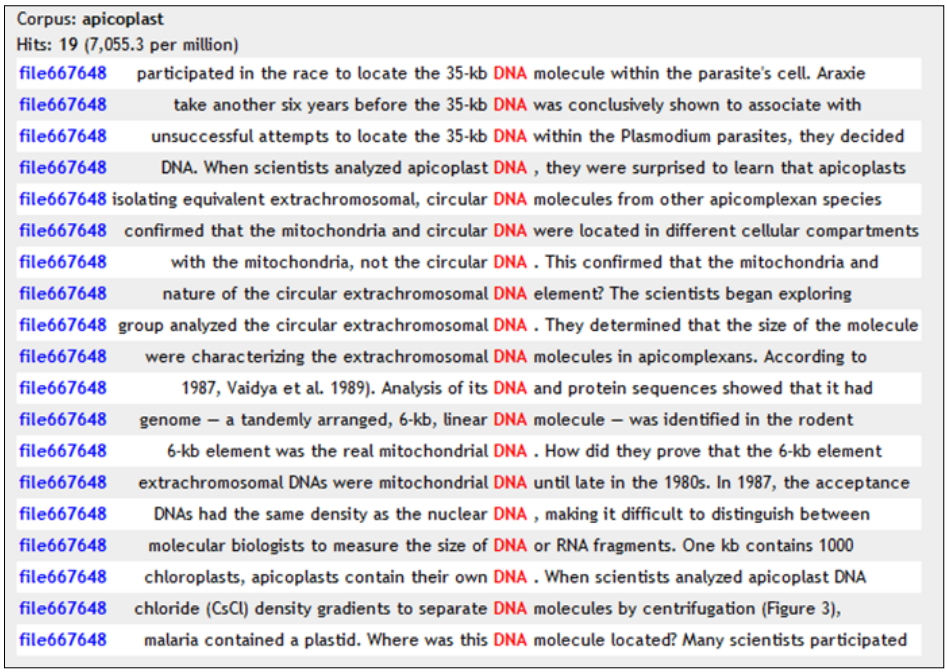
Figure 37. Sketch Difference for 'organelle' and 'mitochondrion'

The Sketch Difference for ‘organelle’ and ‘mitochondrion’ showed that both query items worked as objects of verbs such as ‘contain’ and ‘be’ (Fig. 38).

Figure 38. Sketch Difference-retrieved data

On the other hand, the adjective ‘visible’ was more likely to occur as a premodifier of ‘mitochondrion’ while various adjectives – such as ‘cellular’, ‘small’, ‘oviform’, ‘novel’, ‘secretory’, ‘ovoid’ and ‘new’ – were more likely to be used as premodifiers of ‘organelle’.

Left-sorted concordances of ‘DNA’ (Fig. 39) showed collocations such as ‘apicoplast DNA’, ‘circular DNA molecules’, ‘extrachromosomal DNA element’, ‘tandemly arranged 6-kb linear DNA molecule’, ‘mitochondrial DNA’ and ‘nuclear DNA’. In particular, 3-word units –such as ‘circular DNA molecules’ and ‘extrachromosomal DNA element’– and one 6-word unit, namely ‘tandemly arranged 6-kb linear DNA molecule’, were identified. Lexical bundles were thus consistently retrieved from the text. Furthermore, in keeping with the academic prose features previously mentioned, it emerged that the query item was premodified mainly by attributive adjectives, some of which ended in «-al» and nouns.



Corpus: apicoplast
Hits: 19 (7,055.3 per million)

file667648	participated in the race to locate the 35-kb	DNA	molecule within the parasite's cell. Araxie
file667648	take another six years before the 35-kb	DNA	was conclusively shown to associate with
file667648	unsuccessful attempts to locate the 35-kb	DNA	within the Plasmodium parasites, they decided
file667648	DNA. When scientists analyzed apicoplast	DNA	, they were surprised to learn that apicoplasts
file667648	isolating equivalent extrachromosomal, circular	DNA	molecules from other apicomplexan species
file667648	confirmed that the mitochondria and circular	DNA	were located in different cellular compartments
file667648	with the mitochondria, not the circular	DNA	. This confirmed that the mitochondria and
file667648	nature of the circular extrachromosomal	DNA	element? The scientists began exploring
file667648	group analyzed the circular extrachromosomal	DNA	. They determined that the size of the molecule
file667648	were characterizing the extrachromosomal	DNA	molecules in apicomplexans. According to
file667648	1987, Vaidya et al. 1989). Analysis of its	DNA	and protein sequences showed that it had
file667648	genome — a tandemly arranged, 6-kb, linear	DNA	molecule — was identified in the rodent
file667648	6-kb element was the real mitochondrial	DNA	. How did they prove that the 6-kb element
file667648	extrachromosomal DNAs were mitochondrial	DNA	until late in the 1980s. In 1987, the acceptance
file667648	DNAs had the same density as the nuclear	DNA	, making it difficult to distinguish between
file667648	molecular biologists to measure the size of	DNA	or RNA fragments. One kb contains 1000
file667648	chloroplasts, apicoplasts contain their own	DNA	. When scientists analyzed apicoplast DNA
file667648	chloride (CsCl) density gradients to separate	DNA	molecules by centrifugation (Figure 3),
file667648	malaria contained a plastid. Where was this	DNA	molecule located? Many scientists participated

Figure 39. Left-sorted concordances of DNA

Word Sketch-retrieved data (Fig. 40) showed that DNA worked as (a) the object of verbs such as ‘analyze’, ‘locate’, ‘show’, ‘contain’ and ‘be’, (b) the subject of the verb ‘be’, and (c) the premodifier of nouns such as ‘molecule’ and ‘element’.

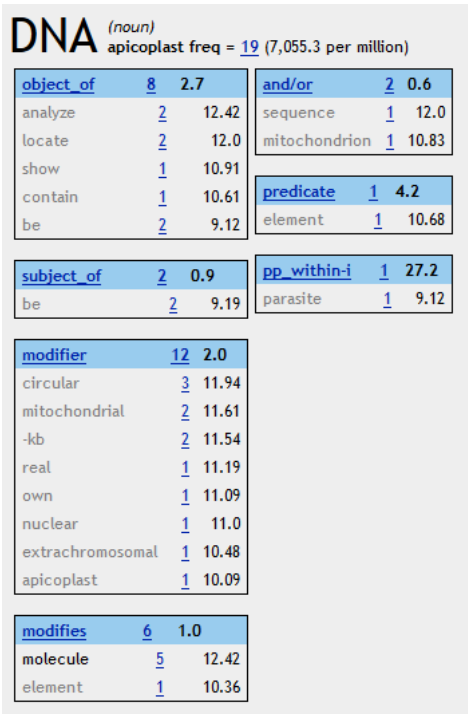


Figure 40. Word Sketch for 'DNA'

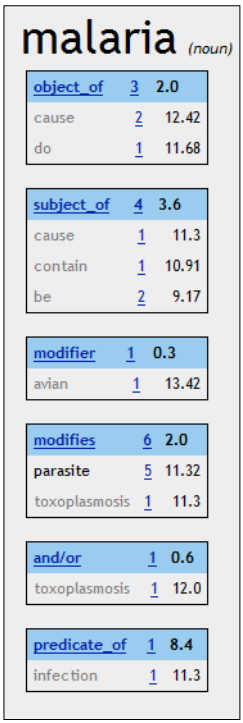


Figure 41. Word Sketch for 'malaria'

The Word Sketch for 'malaria' (Fig. 41) showed that the query item worked as (a) the object of verbs such as 'cause' and 'do', (b) the subject of verbs such as 'cause', 'contain' and 'be', and (c) the premodifier of nouns such as 'parasite' and 'toxoplasmosis'. Furthermore, 'malaria' was premodified by the attributive adjective 'avian'.

A nominal trinomial also emerged from the Word Sketch for 'malaria' (fig. 42).

Figure 42. Nominal trinomials

The Word Sketch for 'chloroplast' (Fig. 43) showed that the query item worked as (a) the object of the verb 'be', and (b) the premodifier of nouns such as 'counterpart', 'apicoplast' and 'genome'. The attributive adjective 'vestigial' was identified as a premodifier of 'chloroplast'; as previously mentioned, derived adjectives ending in «-al» are a main characteristic of scientific writing.

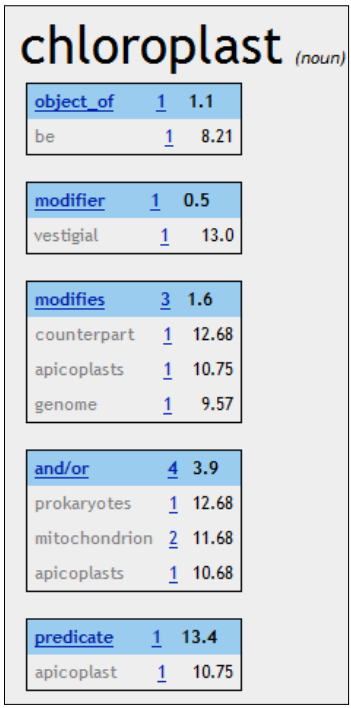


Figure 43. Word Sketch for 'chloroplast'

Nominal binomials were also detected in the Word Sketch for 'chloroplast' (Fig. 44).

named the apicoplast. Like *mitochondria* and chloroplasts , apicoplasts contain their own DNA. When
cellular organelles, like *mitochondria* and chloroplasts , contain their own independent genomes,
apicoplasts share evolutionary similarities with chloroplasts and *prokaryotes* (cyanobacteria), they stand

Figure 44. Nominal binomials

The analysis of concordance output (Fig. 45) showed that prepositional phrases as postmodifiers were very common in the academic writing investigated. Examples, such as the following, were retrieved: 'discovery of the apicoplast', 'the identity and function of the new organelle', 'the acceptance of the rule', 'the nature of the circular extrachromosomal DNA element', 'the size of the molecule', and 'the organization of the 35-kb apicoplast genome'. As Biber's (2006) studies suggest, prepositional phrases as postmodifiers in noun phrases are extensively used in academic prose. In particular, «of-phrases» working as postmodifiers emerged as a key feature of the biology article examined (Fig. 45). This finding was in keeping

with research stating that «of-phrases» are much more widely used in scientific prose (Biber 2006). In general, a wide use of prepositional phrases as postmodifiers was detected in the scientific prose investigated, which was in keeping with Biber et al.'s (2002, p. 269) research highlighting the function performed in this case by prepositional phrases, which «allow a very dense packaging of information in a text».

developmental malformations. An estimated one third	of the	world's human population has been infected
have apical complexes, they are members	of the	same club-the phylum Apicomplexa. All apicomplexan
live or reproduce without a host. Discovery	of the	Apicoplast In the 1960s, microbiologists
more information, the identity and function	of the	new organelle would remain a mystery for
pairs. At the time, the prevailing "rule	of the	ring" dogma was that most, if not all,
late in the 1980s. In 1987, the acceptance	of the	"rule of the ring" dogma was called into
1980s. In 1987, the acceptance of the "rule	of the	ring" dogma was called into question when
mitochondrial genome, what was the nature	of the	circular extrachromosomal DNA element?
explanations for the origin and function	of the	circular genome. An Organelle with a Green
extrachromosomal DNA. They determined that the size	of the	molecule was 35 kb using restriction mapping
when they performed a phylogenetic analysis	of the	conserved regions of the small subunit
phylogenetic analysis of the conserved regions	of the	small subunit rRNA, they noticed similarities
As plant scientists, they were well aware	of the	dogma that only algae and plants have three
organelle located in the mid to anterior region	of the	cell, which was anterior to the nucleus
the "apicoplast," which was a combination	of the	words "apicomplexan" and "plastid" (Köhler
the genomic sequence and the organization	of the	35-kb apicoplast genome led to two main
for a minimal, but sufficient, translation	of the	protein-encoding genes present in the element

Figure 45. Concordances of the prepositional phrase 'of the'

In the same concordance output (Fig. 45), examples of complex postmodification were spotted, such as 'discovery of the apicoplast in the 1960s' and 'analysis of the conserved regions of the small subunit'. As Biber et al. (2002, p. 269) suggest, «[p]repositional phrases as postmodifiers [...] often occur in extremely dense, embedded sequences».

Overall, in the biology article examined, prepositional phrases were more common than relative clauses as postmodifiers. These findings were in keeping with Biber et al.'s (2002) studies stating that in scientific writing, postmodification is mainly conveyed through prepositional phrases, and to a much lesser extent, through relative clauses.

In the biology article investigated, the relative pronouns 'which' (Fig. 46) and 'that' (Fig. 47) were almost equally distributed. These findings were not in line with Biber's (2006) research reporting instead that 'which' is mainly used in academic writing.

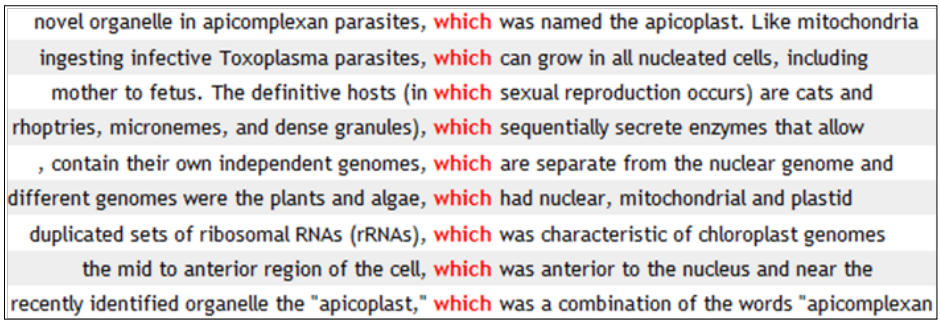


Figure 46. Concordances of the relative pronoun 'which'

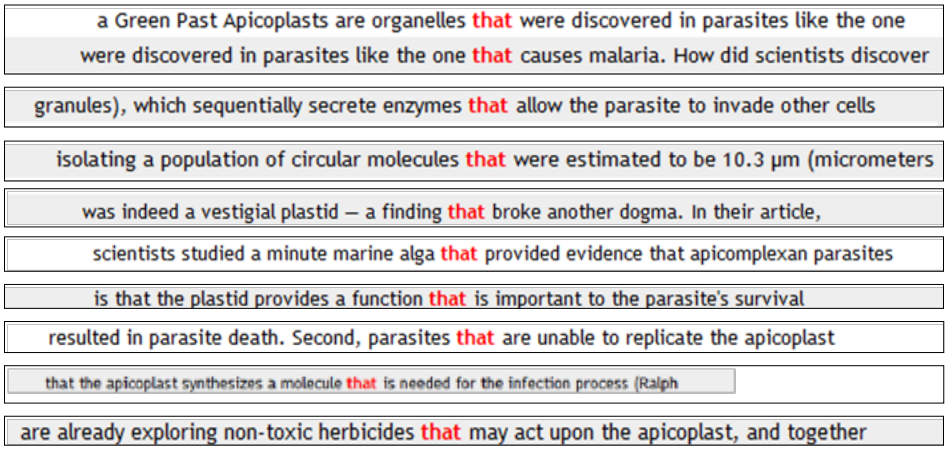


Figure 47. Concordances of the relative pronoun 'that'

The role of modification, in terms of premodification and postmodification, surfaced as a pivotal characteristic of the biology article examined. As Biber et al. (2002, p. 268) suggest with regard to modification in academic texts:

> noun phrases in academic writing usually have premodifiers or postmodifiers (or both). […] [In] academic prose, […] a majority of all noun phrases have some modifications. In fact, much of the new information in academic texts occurs in the modifiers in noun phrases, resulting in a very high density of information.

Concordance output (Fig. 48) showed that the academic prose examined featured many verbs starting with «re-». This phenomenon was in keeping

with Biber's (2006, p. 16) studies reporting that derived verbs are «[m]ost common in academic prose, especially verbs formed with *re-*».

chloroplast? How could apicomplexans and plants be	related	? What follows is the amazing story of how
humans during a bite. Plasmodium parasites	reproduce	in human liver and blood cells, as well
fetus. The definitive hosts (in which sexual	reproduction	occurs) are cats and other felines, and
Unlike Plasmodium and Toxoplasma, Eimeria	requires	a single host to complete its life cycle
parasites, that is, they cannot live or	reproduce	without a host. Discovery of the Apicoplast
malariologist, and Naomi Lang-Unnasch, who	researched	Toxoplasma. After several unsuccessful
the apicoplast puzzle is why apicomplexans	retained	a vestigial plastid despite losing the
chemicals affecting apicoplast metabolism	resulted	in parasite death. Second, parasites that
death. Second, parasites that are unable to	replicate	the apicoplast also die. Amazingly, in
with a chemically damaged apicoplast) while	remaining	in the infected host cell. However, the

Figure 48. Verbs formed with «re-»

Concordance output (Fig. 49) also revealed that derived verbs ending in «-ize» were especially used in the text investigated. As Biber (2006, p. 16) suggests in this respect: «Derived verbs [...] [are] [m]ost common in academic prose, especially verbs formed with [...] *-ize*».

apicoplasts contain their own DNA. When scientists	analyzed	apicoplast DNA, they were surprised to
Parasitology Today article, Wilson and his group	analyzed	the circular extrachromosomal DNA. They
added to the puzzle by scientists who were	characterizing	the extrachromosomal DNA molecules in apicomplexans
these extrachromosomal molecules. She soon	realized	that P. lophurae's extrachromosomal DNAs
possible? One hypothesis is that the apicoplast	synthesizes	a molecule that is needed for the infection

Figure 49. Verbs formed with «-ize»

A word list retrieved from the 'Apicoplast corpus' (Fig. 50) showed a rather large number of nouns ending in «-tion», such as 'function', 'infection', 'combination' and 'hybridization'. This piece of data was in keeping with Biber's (2006, p. 15) research: «Nominalizations [are] much more common in academic prose, especially nouns formed with *-tion*».

word	Freq
function	3
infection	3
organization	2
question	2
information	2
population	2
generation	1
translation	1
exception	1
combination	1
high-resolution	1
hybridization	1
Definition	1
restriction	1
interpretation	1
centrifugation	1
production	1
association	1
reproduction	1
Organization	1

Figure 50. Word list of nouns in «-tion»

word	Freq
density	2
capability	1
possibility	1
University	1
identity	1
severity	1

Figure 51. Word list of nouns in «-ity»

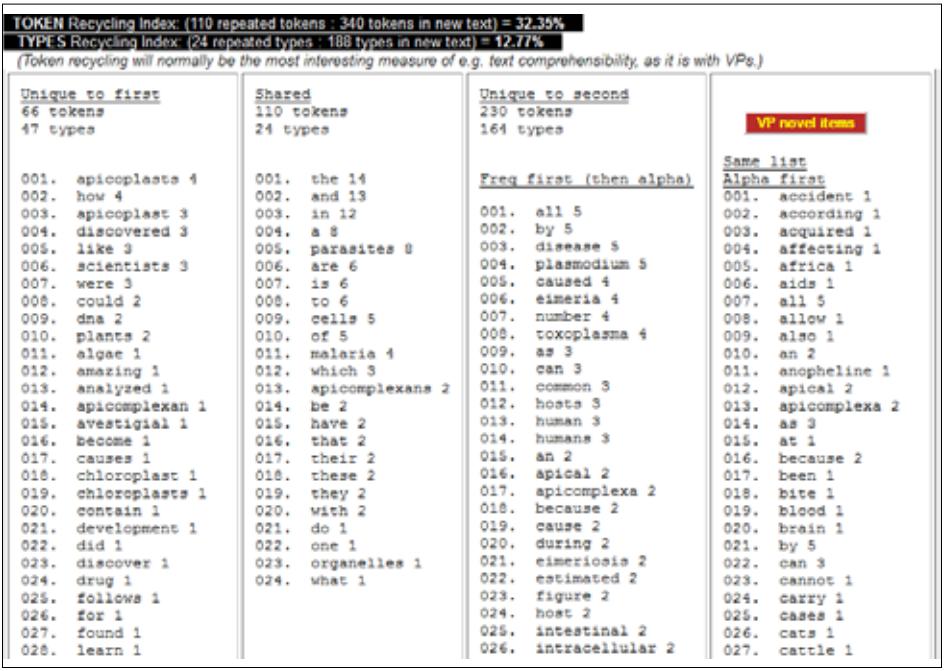
On the other hand, not so many nouns formed with «-ity» (Fig. 51) were found in the corpus in comparison with those formed with «-tion», which was slightly in contrast with Biber's (2006) findings claiming that nouns ending with «-ity» are particularly used in academic prose.

2.5 Content-specific vocabulary distribution

To analyze the biology article more thoroughly and thus improve further course-tailored activities, Text-Lex Compare (`http://www.lextutor.ca/text_lex_compare`), another tool made available by Compleat Lexical Tutor, was used. With Text-Lex Compare, the CLIL expert calculated how many discipline-specific vocabulary items were introduced for the first time in each paragraph of the article and to what extent technical terms were recycled in the following paragraphs.

The biology article was divided into seven sections; the division was made on the grounds of the article paragraphs. First of all, the first para-

graph was compared with the second one (Fig. 52). The data showed that content-specific words such as 'apicoplast', 'DNA', 'apicomplexan', 'chloroplast' and 'plastid' were introduced in the first paragraph while words such as 'plasmodium', 'toxoplasma' and 'intracellular' were introduced for the first time in the second paragraph.

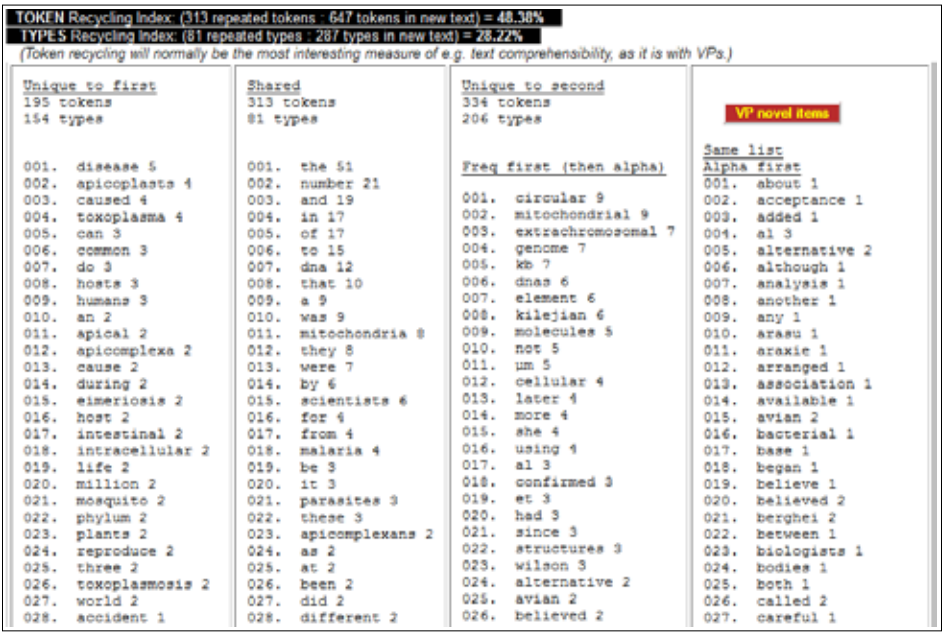


TOKEN Recycling Index: (110 repeated tokens : 340 tokens in new text) = **32.35%**
TYPES Recycling Index: (24 repeated types : 188 types in new text) = **12.77%**
(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first 66 tokens 47 types	Shared 110 tokens 24 types	Unique to second 230 tokens 164 types	VP novel items
		Freq first (then alpha)	Same list Alpha first
001. apicoplasts 4	001. the 14	001. all 5	001. accident 1
002. how 4	002. and 13	002. by 5	002. according 1
003. apicoplast 3	003. in 12	003. disease 5	003. acquired 1
004. discovered 3	004. a 8	004. plasmodium 5	004. affecting 1
005. like 3	005. parasites 8	005. caused 4	005. africa 1
006. scientists 3	006. are 6	006. eimeria 4	006. aids 1
007. were 3	007. is 6	007. number 4	007. all 5
008. could 2	008. to 6	008. toxoplasma 4	008. allow 1
009. dna 2	009. cells 5	009. as 3	009. also 1
010. plants 2	010. of 5	010. can 3	010. an 2
011. algae 1	011. malaria 4	011. common 3	011. anopheline 1
012. amazing 1	012. which 3	012. hosts 3	012. apical 2
013. analyzed 1	013. apicomplexans 2	013. human 3	013. apicomplexa 2
014. apicomplexan 1	014. be 2	014. humans 3	014. as 3
015. avestigial 1	015. have 2	015. an 2	015. at 1
016. become 1	016. that 2	016. apical 2	016. because 2
017. causes 1	017. their 2	017. apicomplexa 2	017. been 1
018. chloroplast 1	018. these 2	018. because 2	018. bite 1
019. chloroplasts 1	019. they 2	019. cause 2	019. blood 1
020. contain 1	020. with 2	020. during 2	020. brain 1
021. development 1	021. do 1	021. eimeriosis 2	021. by 5
022. did 1	022. one 1	022. estimated 2	022. can 3
023. discover 1	023. organelles 1	023. figure 2	023. cannot 1
024. drug 1	024. what 1	024. host 2	024. carry 1
025. follows 1		025. intestinal 2	025. cases 1
026. for 1		026. intracellular 2	026. cats 1
027. found 1			027. cattle 1
028. learn 1			

Figure 52. Comparison of the first paragraph with the second paragraph of the biology article

Afterwards, the first and second paragraphs were compared with the third one (Fig. 53). The list revealed that the following content-specific lexical items were introduced, for example, for the first time in the third paragraph: 'circular', 'mitochondrial', 'extrachromosomal', 'genome', 'molecule', 'cellular' and 'avian'.



TOKEN Recycling Index: (313 repeated tokens : 647 tokens in new text) = **48.38%**

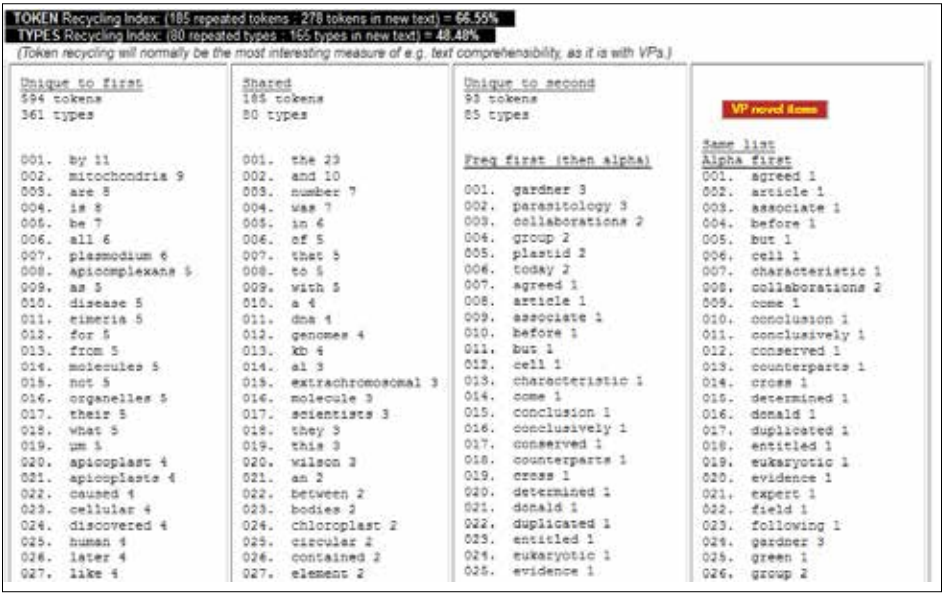
TYPES Recycling Index: (81 repeated types : 287 types in new text) = **28.22%**

(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first 195 tokens 154 types		Shared 313 tokens 81 types		Unique to second 334 tokens 206 types		VP novel items	
				Freq first (then alpha)		Same list Alpha first	
001.	disease 5	001.	the 51	001.	circular 9	001.	about 1
002.	apicoplasts 4	002.	number 21	002.	mitochondrial 9	002.	acceptance 1
003.	caused 4	003.	and 19	003.	extrachromosomal 7	003.	added 1
004.	toxoplasma 4	004.	in 17	004.	genome 7	004.	al 3
005.	can 3	005.	of 17	005.	kb 7	005.	alternative 2
006.	common 3	006.	to 15	006.	dnas 6	006.	although 1
007.	do 3	007.	dna 12	007.	element 6	007.	analysis 1
008.	hosts 3	008.	that 10	008.	kilejian 6	008.	another 1
009.	humans 3	009.	a 9	009.	molecules 5	009.	any 1
010.	an 2	010.	was 9	010.	not 5	010.	arasu 1
011.	apical 2	011.	mitochondria 8	011.	um 5	011.	araxie 1
012.	apicomplexa 2	012.	they 8	012.	cellular 4	012.	arranged 1
013.	cause 2	013.	were 7	013.	later 4	013.	association 1
014.	during 2	014.	by 6	014.	more 4	014.	available 1
015.	eimeriosis 2	015.	scientists 6	015.	she 4	015.	avian 2
016.	host 2	016.	for 4	016.	using 4	016.	bacterial 1
017.	intestinal 2	017.	from 4	017.	al 3	017.	base 1
018.	intracellular 2	018.	malaria 4	018.	confirmed 3	018.	began 1
019.	life 2	019.	be 3	019.	et 3	019.	believe 1
020.	million 2	020.	it 3	020.	had 3	020.	believed 2
021.	mosquito 2	021.	parasites 3	021.	since 3	021.	berghei 2
022.	phylum 2	022.	these 3	022.	structures 3	022.	between 1
023.	plants 2	023.	apicomplexans 2	023.	wilson 3	023.	biologists 1
024.	reproduce 2	024.	as 2	024.	alternative 2	024.	bodies 1
025.	three 2	025.	at 2	025.	avian 2	025.	both 1
026.	toxoplasmosis 2	026.	been 2	026.	believed 2	026.	called 2
027.	world 2	027.	did 2			027.	careful 1
028.	accident 1	028.	different 2				

Figure 53. Comparison of the first and second paragraphs with the third paragraph

Then, the first three paragraphs were compared with the fourth paragraph (Fig. 54). The data showed that discipline-specific words such as ‘parasitology’ and ‘eukaryotic’ appeared in the fourth paragraph for the first time.

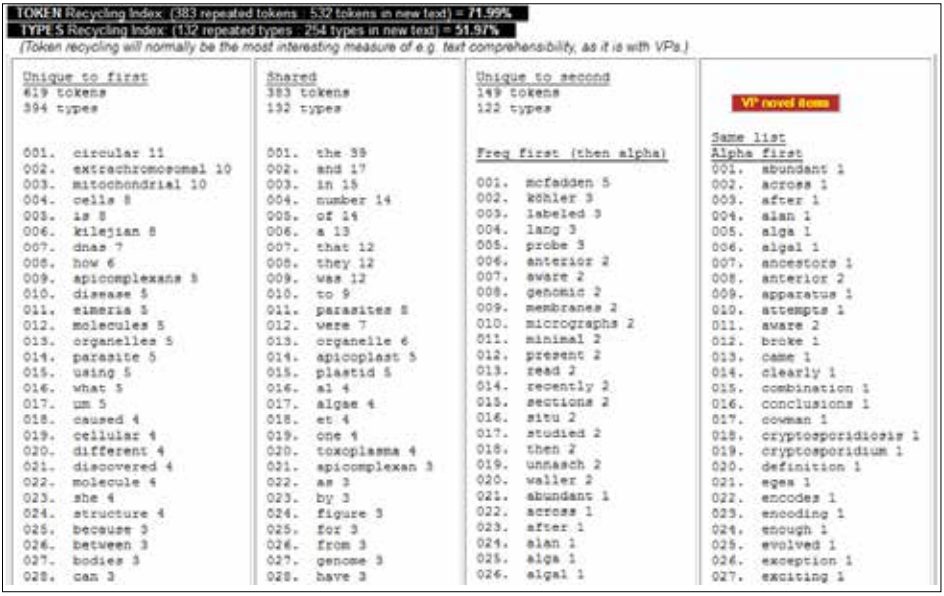


TOKEN Recycling Index: (185 repeated tokens : 278 tokens in new text) = 66.55%
TYPES Recycling Index: (80 repeated types : 165 types in new text) = 48.48%
(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first	Shared	Unique to second	
594 tokens	185 tokens	93 tokens	VP novel items
361 types	80 types	85 types	
001. by 11	001. the 23	Freq first (then alpha)	Same list
002. mitochondria 9	002. and 10		Alpha first
003. are 8	003. number 7	001. gardner 3	001. agreed 1
004. is 8	004. was 7	002. parasitology 3	002. article 1
005. be 7	005. in 6	003. collaborations 2	003. associate 1
006. all 6	006. of 5	004. group 2	004. before 1
007. plasmodium 6	007. that 5	005. plastid 2	005. but 1
008. apicomplexans 5	008. to 5	006. today 2	006. cell 1
009. as 5	009. with 5	007. agreed 1	007. characteristic 1
010. disease 5	010. a 4	008. article 1	008. collaborations 2
011. eimeria 5	011. dna 4	009. associate 1	009. come 1
012. for 5	012. genomes 4	010. before 1	010. conclusion 1
013. from 5	013. kb 4	011. but 1	011. conclusively 1
014. molecules 5	014. al 3	012. cell 1	012. conserved 1
015. not 5	015. extrachromosomal 3	013. characteristic 1	013. counterparts 1
016. organelles 5	016. molecule 3	014. come 1	014. cross 1
017. their 5	017. scientists 3	015. conclusion 1	015. determined 1
018. what 5	018. they 3	016. conclusively 1	016. donald 1
019. um 5	019. this 3	017. conserved 1	017. duplicated 1
020. apicoplast 4	020. wilson 3	018. counterparts 1	018. entitled 1
021. apicoplasts 4	021. an 2	019. cross 1	019. eukaryotic 1
022. caused 4	022. between 2	020. determined 1	020. evidence 1
023. cellular 4	023. bodies 2	021. donald 1	021. expert 1
024. discovered 4	024. chloroplast 2	022. duplicated 1	022. field 1
025. human 4	025. circular 2	023. entitled 1	023. following 1
026. later 4	026. contained 2	024. eukaryotic 1	024. gardner 3
027. like 4	027. element 2	025. evidence 1	025. green 1
			026. group 2

Figure 54. Comparison of the first three paragraphs with the fourth paragraph

Afterwards, the first four paragraphs were compared with the fifth paragraph (Fig. 55). The list generated showed that technical terms such as ‘genomic’ and ‘membrane’ were featured in the fifth paragraph for the first time.

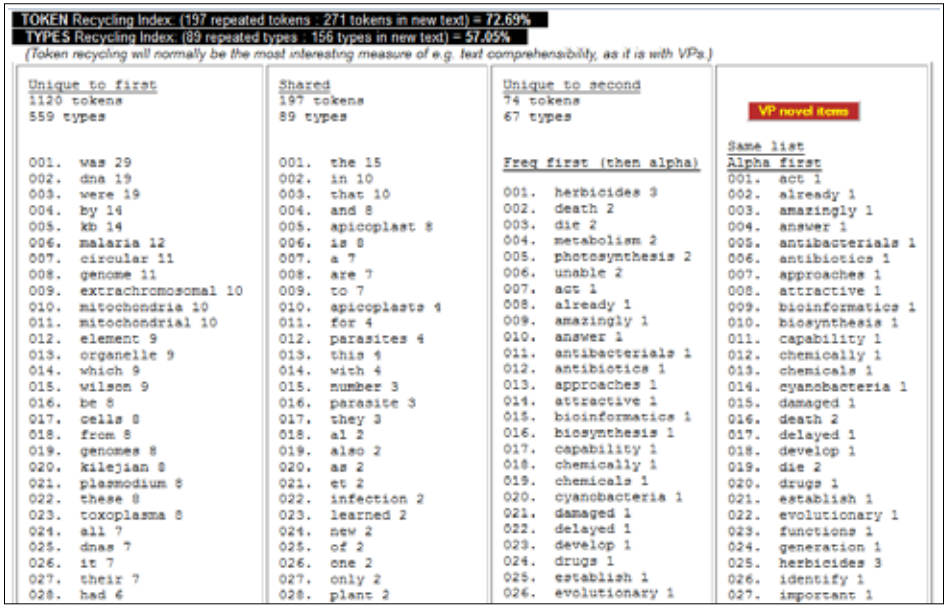


TOKEN Recycling Index: (383 repeated tokens : 532 tokens in new text) = 71.99%
TYPES Recycling Index: (132 repeated types : 254 types in new text) = 51.97%
(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first 619 tokens 394 types		Shared 383 tokens 132 types		Unique to second 149 tokens 122 types		VP novel items	
				Freq first (then alpha)		Same list Alpha first	
001.	circular 11	001.	the 39	001.	mcfadden 5	001.	abundant 1
002.	extrachromosomal 10	002.	and 17	002.	köhler 3	002.	across 1
003.	mitochondrial 10	003.	in 15	003.	labeled 3	003.	after 1
004.	cells 8	004.	number 14	004.	lang 3	004.	alan 1
005.	is 8	005.	of 14	005.	probe 3	005.	alga 1
006.	kilejian 8	006.	a 13	006.	anterior 2	006.	algal 1
007.	dnas 7	007.	that 12	007.	aware 2	007.	ancestors 1
008.	how 6	008.	they 12	008.	genomic 2	008.	anterior 2
009.	apicomplexans 5	009.	was 12	009.	membranes 2	009.	apparatus 1
010.	disease 5	010.	to 9	010.	micrographs 2	010.	attempts 1
011.	eimeria 5	011.	parasites 8	011.	minimal 2	011.	aware 2
012.	molecules 5	012.	were 7	012.	present 2	012.	broke 1
013.	organelles 5	013.	organelle 6	013.	read 2	013.	came 1
014.	parasite 5	014.	apicoplast 5	014.	recently 2	014.	clearly 1
015.	using 5	015.	plastid 5	015.	sections 2	015.	combination 1
016.	what 5	016.	al 4	016.	situ 2	016.	conclusions 1
017.	um 5	017.	algae 4	017.	studied 2	017.	cowman 1
018.	caused 4	018.	et 4	018.	then 2	018.	cryptosporidiosis 1
019.	cellular 4	019.	one 4	019.	unnasch 2	019.	cryptosporidium 1
020.	different 4	020.	toxoplasma 4	020.	waller 2	020.	definition 1
021.	discovered 4	021.	apicomplexan 3	021.	abundant 1	021.	ages 1
022.	molecule 4	022.	as 3	022.	across 1	022.	encodes 1
023.	she 4	023.	by 3	023.	after 1	023.	encoding 1
024.	structure 4	024.	figure 3	024.	alan 1	024.	enough 1
025.	because 3	025.	for 3	025.	alga 1	025.	evolved 1
026.	between 3	026.	from 3	026.	algal 1	026.	exception 1
027.	bodies 3	027.	genome 3			027.	exciting 1
028.	can 3	028.	have 3				

Figure 55. Comparison of the first four paragraphs with the fifth paragraph

Later on, the first five paragraphs were compared with the sixth paragraph (Fig. 56). The subject-specific words introduced in the sixth paragraph were, for example, ‘metabolism’, ‘photosynthesis’, ‘antibacterial’, ‘biosynthesis’ and ‘cyanobacteria’.

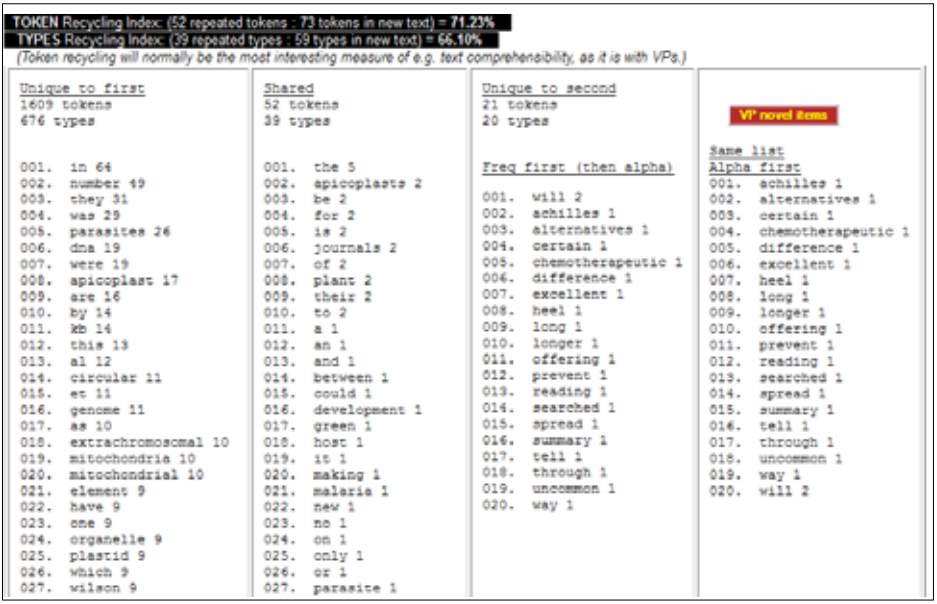


TOKEN Recycling Index: (197 repeated tokens : 271 tokens in new text) = **72.69%**
TYPES Recycling Index: (89 repeated types : 156 types in new text) = **57.05%**
(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first 1120 tokens 559 types	Shared 197 tokens 89 types	Unique to second 74 tokens 67 types	VP novel items
			Same list
		Freq first (then alpha)	Alpha first
001. was 29	001. the 15	001. herbicides 3	001. act 1
002. dna 19	002. in 10	002. death 2	002. already 1
003. were 19	003. that 10	003. die 2	003. amazingly 1
004. by 14	004. and 8	004. metabolism 2	004. answer 1
005. kb 14	005. apicoplast 8	005. photosynthesis 2	005. antibacterials 1
006. malaria 12	006. is 8	006. unable 2	006. antibiotics 1
007. circular 11	007. a 7	007. act 1	007. approaches 1
008. genome 11	008. are 7	008. already 1	008. attractive 1
009. extrachromosomal 10	009. to 7	009. amazingly 1	009. bioinformatics 1
010. mitochondria 10	010. apicoplasts 4	010. answer 1	010. biosynthesis 1
011. mitochondrial 10	011. for 4	011. antibacterials 1	011. capability 1
012. element 9	012. parasites 4	012. antibiotics 1	012. chemically 1
013. organelle 9	013. this 4	013. approaches 1	013. chemicals 1
014. which 9	014. with 4	014. attractive 1	014. cyanobacteria 1
015. wilson 9	015. number 3	015. bioinformatics 1	015. damaged 1
016. be 8	016. parasite 3	016. biosynthesis 1	016. death 2
017. cells 8	017. they 3	017. capability 1	017. delayed 1
018. from 8	018. al 2	018. chemically 1	018. develop 1
019. genomes 8	019. also 2	019. chemicals 1	019. die 2
020. kilejian 8	020. as 2	020. cyanobacteria 1	020. drugs 1
021. plasmodium 8	021. et 2	021. damaged 1	021. establish 1
022. these 8	022. infection 2	022. delayed 1	022. evolutionary 1
023. toxoplasma 8	023. learned 2	023. develop 1	023. functions 1
024. all 7	024. new 2	024. drugs 1	024. generation 1
025. dnas 7	025. of 2	025. establish 1	025. herbicides 3
026. it 7	026. one 2	026. evolutionary 1	026. identify 1
027. their 7	027. only 2		027. important 1
028. had 6	028. plant 2		

Figure 56. Comparison of the first five paragraphs with the sixth paragraph

The data retrieved in the summary section (Fig. 57), corresponding to the seventh paragraph of the text, showed that only one new content-specific word was introduced for the first time, namely ‘chemotherapeutic’.



TOKEN Recycling Index: (52 repeated tokens : 73 tokens in new text) = 71.23%
TYPES Recycling Index: (39 repeated types : 59 types in new text) = 66.10%
(Token recycling will normally be the most interesting measure of e.g. text comprehensibility, as it is with VPs.)

Unique to first 1609 tokens 676 types	Shared 52 tokens 39 types	Unique to second 21 tokens 20 types	VP novel items
			Same list Alpha first
001. in 64	001. the 5	Freq first (then alpha)	001. achilles 1
002. number 49	002. apicoplasts 2		002. alternatives 1
003. they 31	003. be 2	001. will 2	003. certain 1
004. was 29	004. for 2	002. achilles 1	004. chemotherapeutic 1
005. parasites 26	005. is 2	003. alternatives 1	005. difference 1
006. dna 19	006. journals 2	004. certain 1	006. excellent 1
007. were 19	007. of 2	005. chemotherapeutic 1	007. heel 1
008. apicoplast 17	008. plant 2	006. difference 1	008. long 1
009. are 16	009. their 2	007. excellent 1	009. longer 1
010. by 14	010. to 2	008. heel 1	010. offering 1
011. kb 14	011. a 1	009. long 1	011. prevent 1
012. this 13	012. an 1	010. longer 1	012. reading 1
013. al 12	013. and 1	011. offering 1	013. searched 1
014. circular 11	014. between 1	012. prevent 1	014. spread 1
015. et 11	015. could 1	013. reading 1	015. summary 1
016. genome 11	016. development 1	014. searched 1	016. tell 1
017. as 10	017. green 1	015. spread 1	017. through 1
018. extrachromosomal 10	018. host 1	016. summary 1	018. uncommon 1
019. mitochondria 10	019. it 1	017. tell 1	019. way 1
020. mitochondrial 10	020. making 1	018. through 1	020. will 2
021. element 9	021. malaria 1	019. uncommon 1	
022. have 9	022. new 1	020. way 1	
023. one 9	023. no 1		
024. organelle 9	024. on 1		
025. plastid 9	025. only 1		
026. which 9	026. or 1		
027. wilson 9	027. parasite 1		

Figure 57. Comparison of the first six paragraphs with the seventh paragraph

The information retrieved with Text-Lex Compare was used to create pre- and while-reading comprehension activities focusing especially on newly introduced content-specific keywords. The data collected through Text-Lex Compare were also instrumental in targeting lexical priming while devising CLIL teaching materials.

2.6 Online corpus-informed activities

The intercollocability of the biology article mapped through concordance outputs, Word Sketches and Sketch Differences, represented the basis for CLIL materials design in terms of both reading comprehension and language awareness activities. The data retrieved through Text-Lex Compare were also extensively used.

Corpus-driven findings were used to create different kinds of teaching materials. In the pre-reading phase, for example, students were expected to do matching and multiple choice activities aimed at introducing key content-specific vocabulary items (Fig. 58) and collocations (Fig. 59) previously identified through keyword lists and concordance output. The activities were created with Learnclick (`www.learnclick.com`), an online tool, and embedded in the CLIL website.

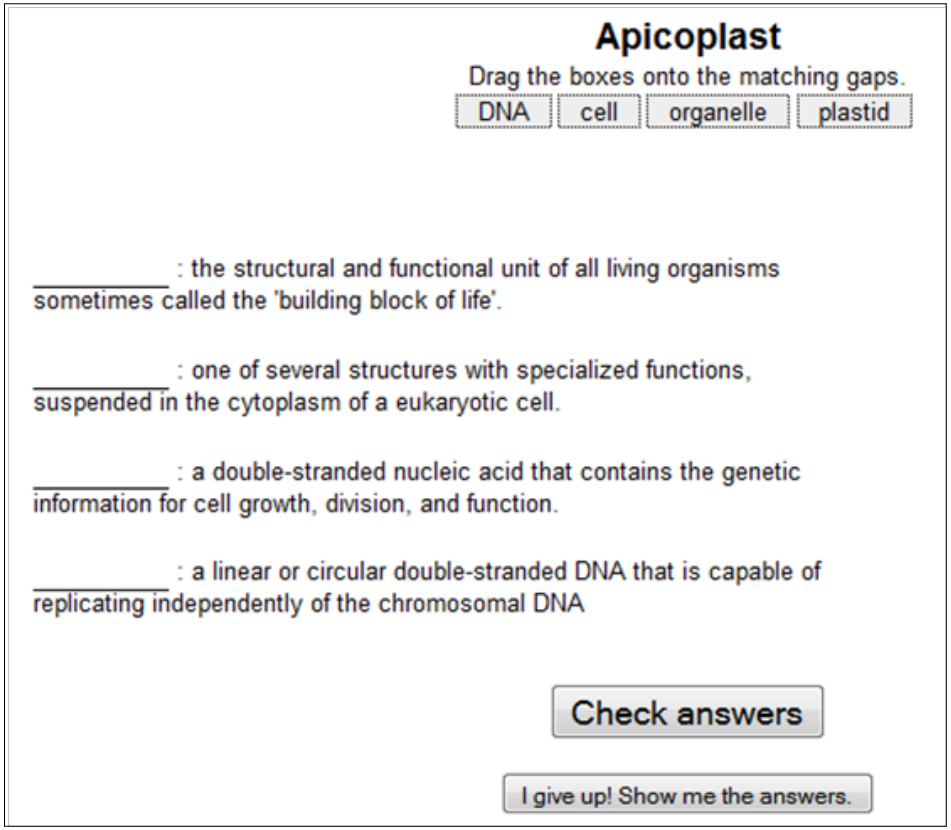


Figure 58. Learnclick-generated matching activity

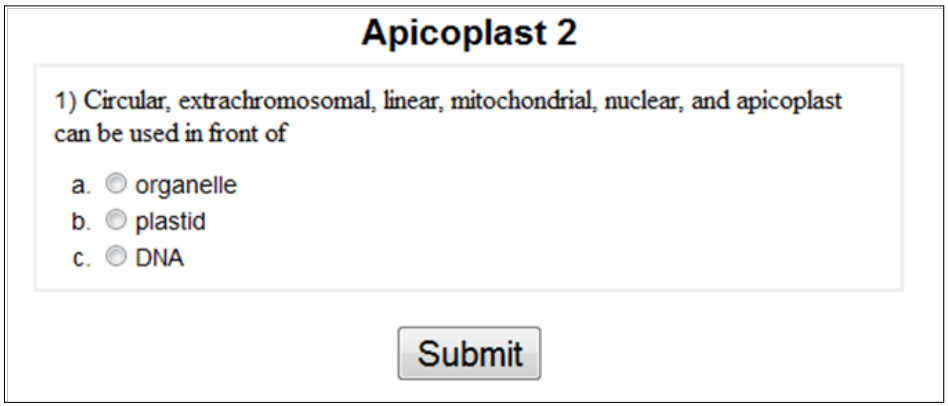


Figure 59. Learnclick-generated matching activity

Upon completion of the matching activities described above, learners were usually expected to carry out a pre-reading, subject-specific cloze exercise as well. To this purpose, Textcompactor (`http://textcompactor.com`), a free online automatic text summarization tool previously mentioned, was extensively used to summarize parts of the targeted content-specific text. In Texcompactor, language items are kept or erased on the grounds of their frequency scores:

> After text is placed on the page, the web app [i.e. Textcompactor] calculates the frequency of each word in the passage. Then, a score is calculated for each sentence based on the frequency count associated with the words it contains. The most important sentence is deemed to be the sentence with the highest frequency count. (`http://textcompactor.com/about`)

Texcompactor-generated summaries were usually checked and edited to different extents, if necessary, by an English native-speaker expert specialized in scientific subjects to guarantee the quality of the text produced and to make the summary suitable for the activity planned.

The pre-reading Learnclick-generated cloze exercise mentioned above, which targets content-specific vocabulary items and collocations, was devised on the grounds of a Texcompactor-generated summary amounting to 40% of the first paragraph of the biology article examined (Fig. 60). Drop down clues were provided to help students' hypothesis formation process. Learners were then asked to check their hypotheses while reading the first paragraph of the subject-specific text. All Learnclick-generated activities were also provided in PDF format (Fig. 61).

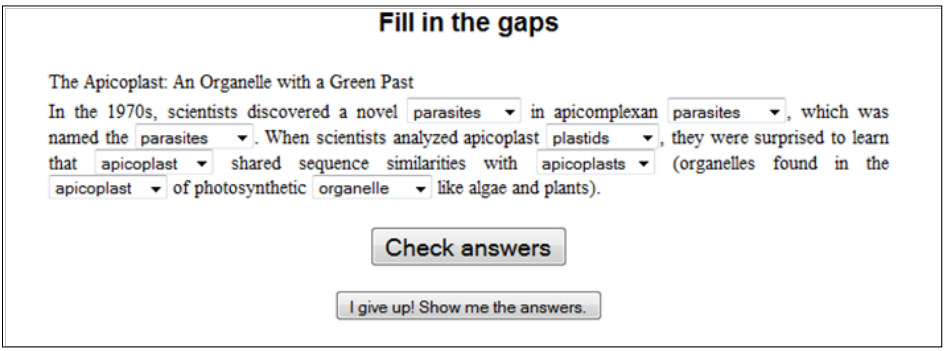
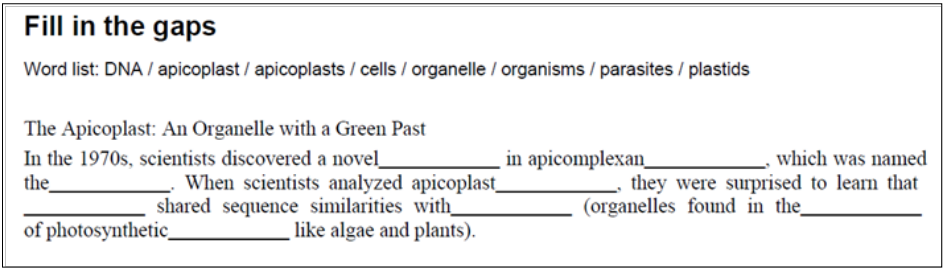


Figure 60. Learnclick-generated cloze passage with dropdowns



Fill in the gaps

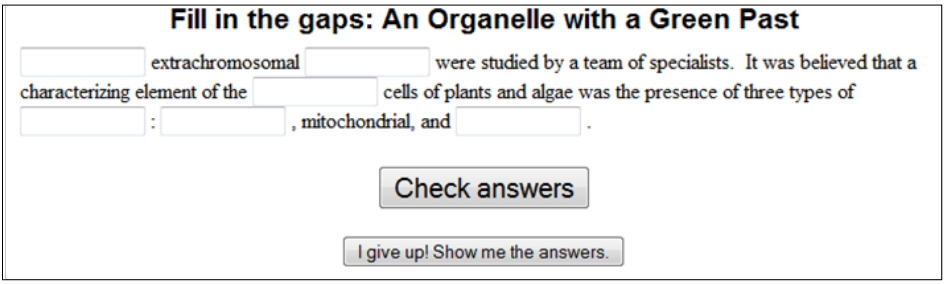
Word list: DNA / apicoplast / apicoplasts / cells / organelle / organisms / parasites / plastids

The Apicoplast: An Organelle with a Green Past

In the 1970s, scientists discovered a novel_____ in apicomplexan_____, which was named the_____. When scientists analyzed apicoplast_____, they were surprised to learn that _____ shared sequence similarities with_____ (organelles found in the_____ of photosynthetic_____ like algae and plants).

Figure 61. Learnclick-generated cloze passage in PDF format

While-reading, corpus-informed, Learnclick-generated cloze activities targeting content-specific concepts and collocations were also devised. To this purpose, for example, a Texcompactor-generated summary amounting to 40% of the first part of the fourth paragraph of the biology article was employed. The summary was slightly edited by an English native-speaker specialist to differentiate it from the original text. This online while-reading, content-specific, Learnclick-generated activity was provided in two different formats (Figs. 62 and 63) to suit a wider range of students' learning styles as well as levels of content and language proficiency.



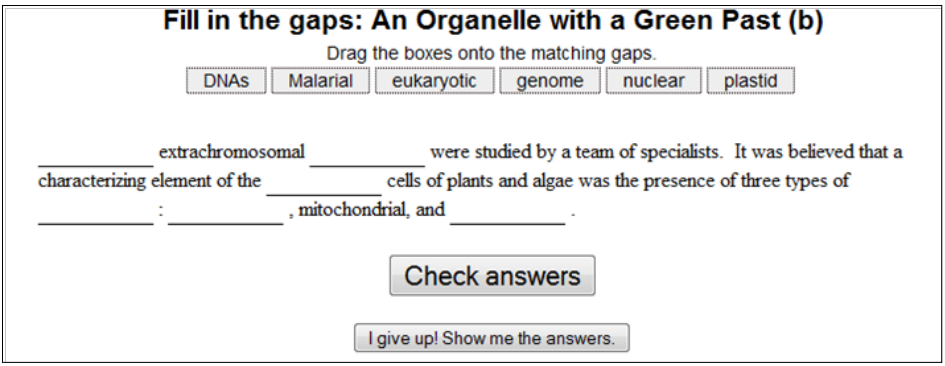
Fill in the gaps: An Organelle with a Green Past

_____ extrachromosomal _____ were studied by a team of specialists. It was believed that a characterizing element of the _____ cells of plants and algae was the presence of three types of _____ : _____ , mitochondrial, and _____ .

Check answers

I give up! Show me the answers.

Figure 62. Learnclick-generated cloze passage with blank boxes



Fill in the gaps: An Organelle with a Green Past (b)

Drag the boxes onto the matching gaps.

DNAs | Malarial | eukaryotic | genome | nuclear | plastid

_____ extrachromosomal _____ were studied by a team of specialists. It was believed that a characterizing element of the _____ cells of plants and algae was the presence of three types of _____ : _____ , mitochondrial, and _____ .

Check answers

I give up! Show me the answers.

Figure 63. Learnclick-generated drag and drop cloze passage

Texcompactor-generated summaries – partly edited by the CLIL expert and an English native-speaker specialist on the grounds of the triangulation of the data retrieved with Sketch Engine, Vocabulary Profiler, and Tex-Lex Compare – were also used to devise especially customized content-specific vocabulary targeted cloze passages. The comprehension of new content concepts was thereby enhanced. The while-reading discipline-specific cloze exercises were generated with Learnclick and embedded in the CLIL website (Fig. 64); they were also made available in PDF format.

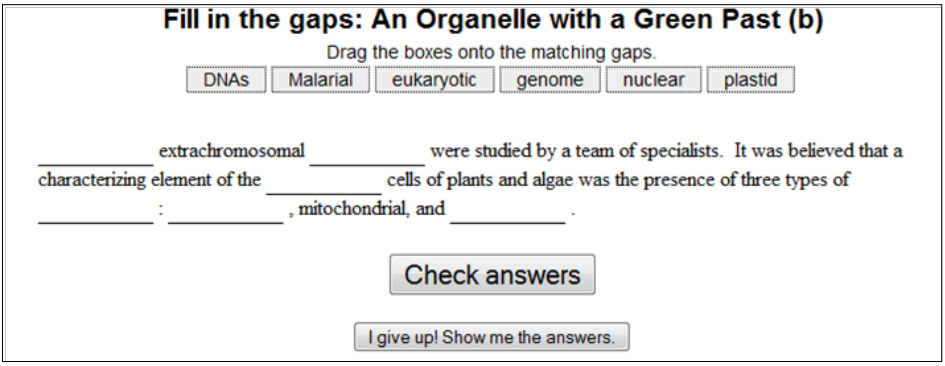


Figure 64. Learnclick-generated cloze exercise

Overall, learners were provided with a wide array of online course-tailored corpus-informed activities, featuring authenticity as well as real usage, as added educational values:

> corpus-informed materials are genuinely special in the following ways:
> - they are based on actual usage;
> - the examples used in them, although they may sometimes be edited or adapted, are a reflection of real usage. They are not invented;
> - the syllabus (the items to be taught as well as the sequence in which they will be presented) is informed by frequency information: [f]or instance, we can prioritize grammar and vocabulary that is most frequent and most useful. (McCarthy 2004, pp. 15-16)

Authenticity and actual usage are also basic tenets of CLIL. Corpus-informed activities thus seemed to be especially suitable to cater to CLIL students.

Learners were expected to carry out activities in a self-determined way. The acquisition of subject-specific concepts was thus enhanced along with the development of autonomous learning. These two combined processes are specific to CLIL: «CLIL as a fused subject provides a learning environment which promotes the capacity for self-organisation» (Wolff 2011, p. 73).

2.7 Lexical priming: technology-enhanced activities

In keeping with the theoretical principles of lexical priming introduced earlier in this work, the data retrieved through concordance outputs and Word Sketches from the 'Apicoplast corpus', as well as by means of Text-Lex Compare, were used to create course-tailored activities aimed to foster lexical priming. To this purpose, subject-specific collocations and colligations were used to design pre- and while-reading comprehension activities.

Furthermore, activities working on reaction time of content-specific word recognition were created with Compleat Reaction Time (`http://www.lextutor.ca/rt`), a tool of Compleat lexical Tutor, to promote discipline-specific lexical priming. Activities working on priming semantic associations as well as discipline-specific collocations were devised. As Hoey (2005, p. 63) suggests, «Words can be primed for collocation, semantic association and colligation».

To create the activities, non-words retrieved from the list elaborated by Paul Meara and his team[2] were used. An example of how a Compleat Reaction Time activity priming semantic associations was devised is provided below (Fig. 65).

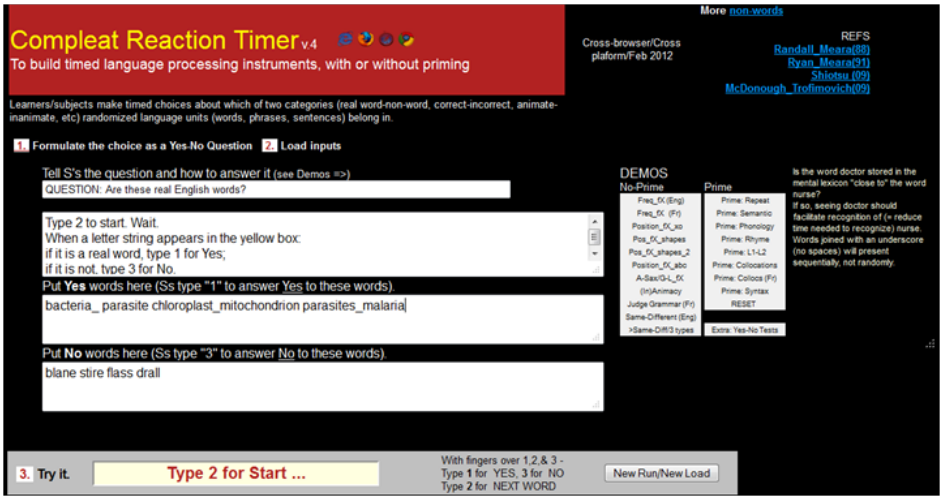


Figure 65. Compleat Reaction Time activity targeting semantic priming

2 «Plausible Non-Words (PNWs) as devised by Paul Meara & colleagues for use in Yes-No Tests» (`http://www.lextutor.ca/freq/lists_download/pnwords.html`).

An example of how a Compleat Reaction Time activity priming content-specific collocations was created is also provided (Fig. 66) along with learners' interface (Fig. 67).

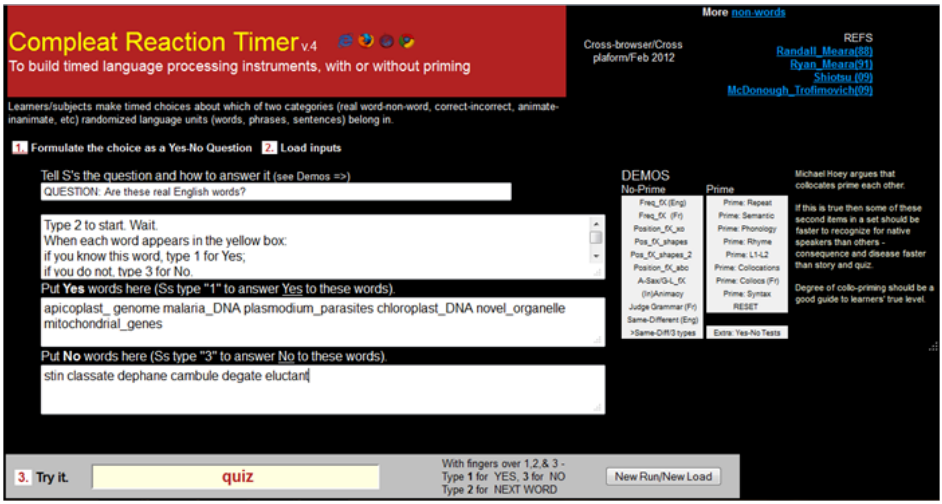


Figure 66. Compleat Reaction Time activity priming content-specific collocations

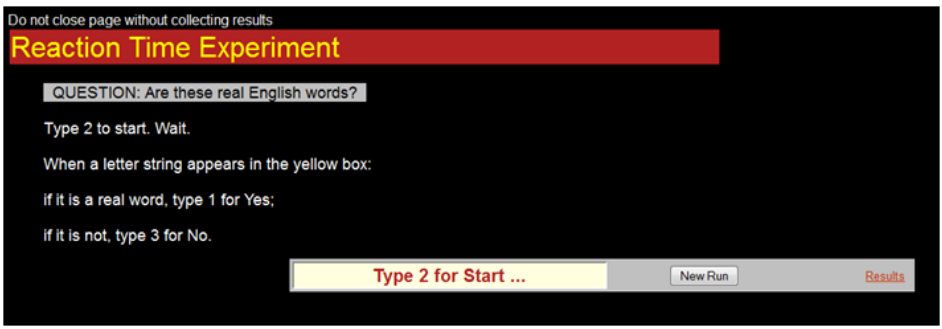


Figure 67. Students' interface of Compleat Reaction Time activities

2.8 ConcGrams in a biology text

ConcGram was used by the CLIL expert to develop the phraseology profile of the biology article examined. Examples of 3-, 4-, and 5-word concgrams retrieved from the biology article are provided below (Fig. 68).

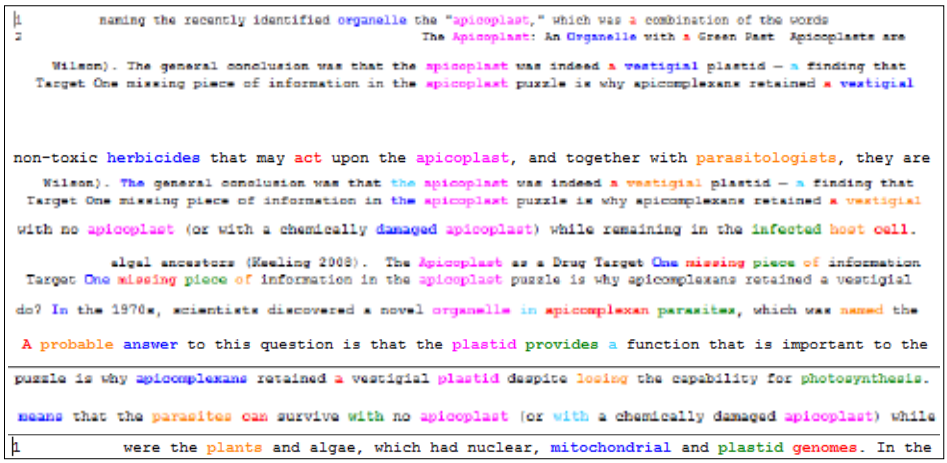


Figure 68. 3-, 4-, and 5-word concgrams

A summary of positional variations of the concgram 'circular/mitochondrial' retrieved from the biology article is given below (Fig. 69).

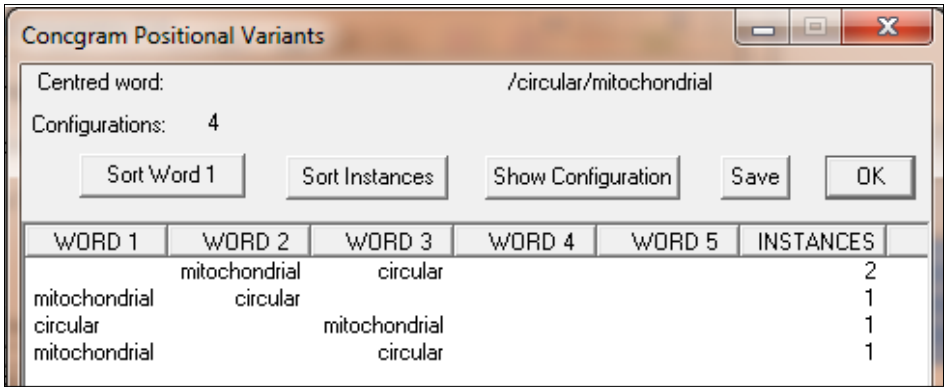


Figure 69. A summary of positional variations of the concgram 'circular/mitochondrial'

Examples of positional variations of the concgram 'circular/mitochondrial' – such as 'circular, extrachromosomal DNAs were mitochondrial DNA' (line 1) and 'mitochondrial DNAs were circular' (line 2) – can be observed in the concordance output provided (Fig. 70).

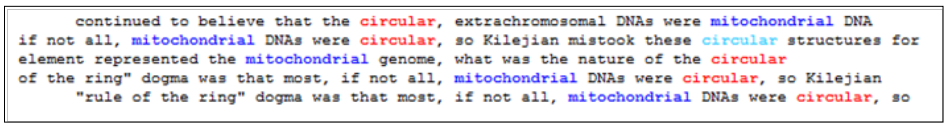


Figure 70. Concordance output of the concgram 'circular/mitochondrial'

Examples of positional variations of the concgram 'organelle/located' – such as 'organelle located' and 'located in the ovoid organelle'– are also given (Fig. 71).

These experiments clearly showed that the 35-kb genome was located in the ovoid organelle
showed that the probe labeled a single, small, oviform organelle located in the mid to anterior

Figure 71. Concordance output of the concgram 'organelle/located'

A summary of constituency variations of the concgram 'parasites/reproduce' retrieved from the biology article is provided (Fig. 72).

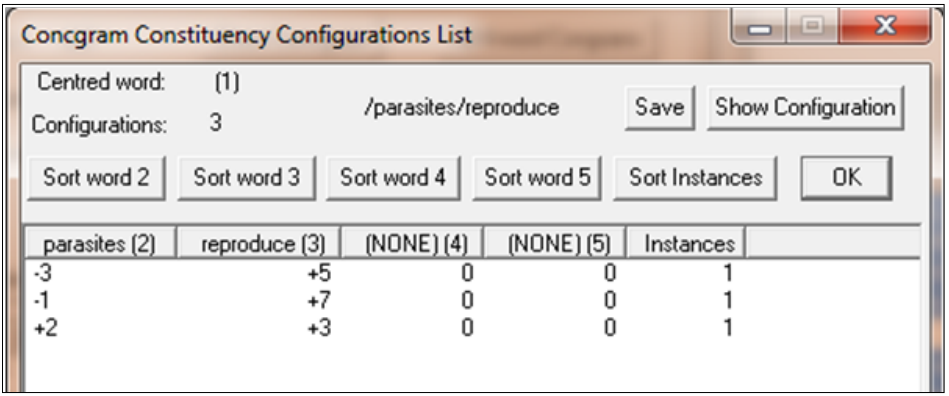


Figure 72. A summary of constituency variations of the concgram 'parasites/reproduce'

Examples of constituency variations of the concgram 'parasites/reproduce' – such as 'parasites reproduce' (line 2) and 'parasites, that is, they cannot live or reproduce' (line 1), where intervening elements occur between 'parasites' and 'reproduce' – can be observed in the concordance output below (Fig. 73).

are obligate, intracellular, protozoan parasites, that is, they cannot live or reproduce without a
and transmit them to humans during a bite. Plasmodium parasites reproduce in human liver and

Figure 73. Concordance output of the concgram 'parasites/reproduce'

In a CLIL learning environment, it is pivotal to detect the most widely used content-specific contiguous and non-contiguous multi-word units, also featuring constituency and positional variations characteristic of English academic writing. In a CLIL setting, it is especially important for students to be sensitized to the role played by content-specific, non-contiguous multi-word units. Activities targeted to promoting the acquisition of the multi-word units retrieved from scientific texts need to be implemented. To this end, the aboutgrams retrieved from the biology article were first used to create language awareness activities. Then, the acquisition of aboutgrams was enhanced by means of content-specific cloze exercises; here, learners were

expected to guess elements of the discipline-specific aboutgrams provided. Afterwards, aboutgrams were employed to enhance semantic associations, and therefore lexical priming in reading comprehension activities.

2.9 The Academic Formulas List in a biology text

The CLIL expert investigated the Sketch Engine-generated 'Apicoplast corpus' to identify the Academic Formulas it contained (Figs. 74-79).

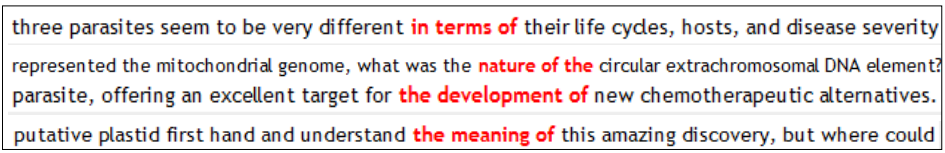


Figure 74. Academic Formulas working as intangible framing attributes[3] in the biology article

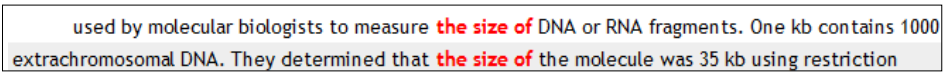


Figure 75. Academic Formulas working as tangible framing attributes in the biology article

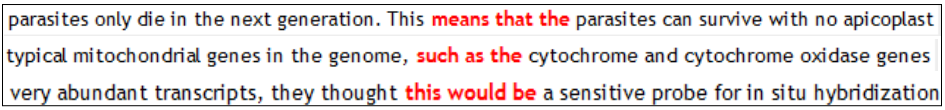


Figure 76. Academic Formulas signaling identification and focus in the biology article

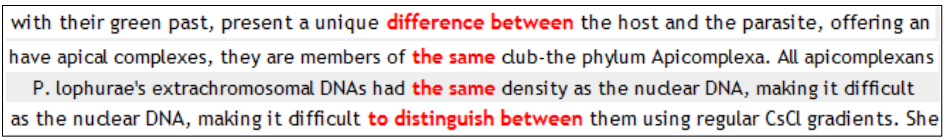


Figure 77. Academic Formulas signaling contrast and comparison in the biology article

Figure 78. Academic Formulas conveying ability and possibility in the biology article

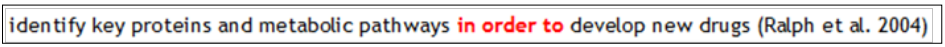


Figure 79. Academic Formulas signaling cause and effect in the biology article

3 The definitions used to classify the Academic Formulas are drawn from Simpson-Vlach and Ellis's (2009) work.

The Academic Formulas retrieved from the 'Apicoplast corpus' are summarized in the grid provided below (Tab. 1).

Table 1. The Academic Formulas featured in the biology article

A. Referential Expressions	
1. Specification of attributes	
a) Intangible framing attributes	in terms of, nature of the, the development of, the meaning of
b) Tangible framing attributes	the size of
c) Quantity specification	
2. Identification & focus	means that the, such as the, this would be
3. Contrast & comparison	difference between, to distinguish between, the same
4. Deictics & locatives	
5. Vagueness markers	

B. Stance Expressions	
1. Hedges	
2. Epistemic stance	
3. Obligation & directive	
4. Expressions of ability & possibility	can also be
5. Evaluation	
6. Intention/volition, prediction	

C. Discourse Organizing Functions	
1. Metadiscourse & textual reference	
2. Topic introduction & focus	
3. Topic elaboration	
a) Non-causal	
b) Topic elaboration: cause & effect	in order to
4. Discourse markers	

Activities were created to help students notice and acquire the Academic Formulas featured in the scientific article investigated. Furthermore, students were usually asked to identify the category the Academic Formulas belonged to by means of multiple choice activities; classifying concepts is likely to foster the acquisition of targeted language items. Language awareness was thus implemented in a contextualized way, and likewise, active learning was consistently enhanced.

2.10 Data-Driven Learning

Corpus-driven findings were deployed to implement DDL (Data-Driven Learning). In DDL, learners are provided with sets of concordance lines, which are a «listing of all the occurrences in the corpus of the query item, together with some surrounding context in the form of words to the left and right» (Hoffmann et al. 2008, p. 264). Browsing the paradigmatic and syntagmatic axes of the concordances, students have to infer the content-specific query item missing. DDL thus assumes «three stages of inductive reasoning […]: observation (of concordanced evidence), classification (of salient features) and generalization (of rules)» (McEnery et al. 2006, p. 99). It follows that DDL is instrumental in fostering inferring skills:

> The assumption that underlies this [data-driven learning approach] is that effective language learning is itself a form of linguistic research, and that the concordance printout offers a unique resource for the stimulation of inductive learning strategies – in particular the strategies of perceiving similarities and differences and of hypothesis formation and testing. (Johansson 2009, p. 37)

In DDL, active learning is thus enhanced through discovery learning: «By exposing students to authentic data in the form of concordance lines, corpus-driven exercises allow for a more student-centred approach to learning, giving them the opportunity to evaluate data and draw their conclusions» (Oksefjell Ebeling 2009, p. 81).

In a CLIL learning environment, DDL materials are created to address targeted discipline-specific language items. Content-specific DDL can be instrumental in fostering both language awareness and foreign language acquisition. In DDL, concgrams can be a valuable resource to enhance lexical priming through the activation of semantic associations; guessing the domain-specific node word deleted from the concordance lines provided can therefore be scaffolded through lexical priming.

In DDL, noticing is effectively enhanced along with inferring, and the combination of these two cognitive processes is likely to foster intake of content-specific vocabulary: «if students can notice and find more examples for themselves, they have more chance of acquiring new language» (McCarthy et al. 2010, p. 36). Reading concordances syntagmatically and paradigmatically enhances learners' contact with the language units targeted: «Cobb (1997) points to the potential of DDL to provide multiple contextual encounters for the acquisition of new vocabulary» (O'Keeffe et al. 2007, p. 24). Cobb's position is further supported by Johansson (2009, p. 39) who holds that: «the use of corpora will strengthen natural processes of language acquisition in that they make it easier for learners to notice and experience repeated instances of use».

In DDL, students are provided with authentic materials, which is in keeping with CLIL tenets. In CLIL, content-specific DDL is customized for each course. Discipline-specific DDL:

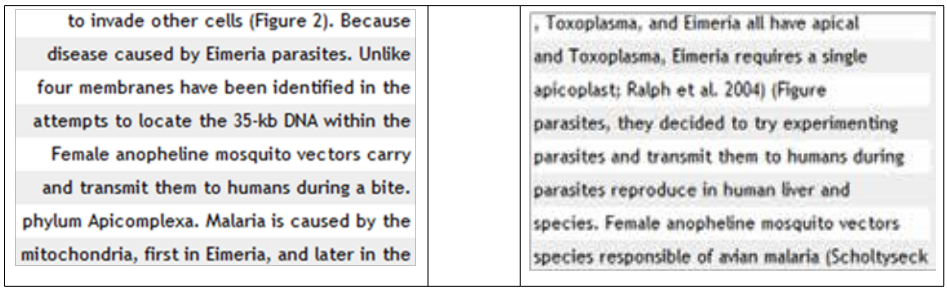
> has the advantage that the information is based on what the students are interested in finding out about, and it gives the teacher an additional opportunity to show students how corpora can provide information beyond what we find in reference books. (Granath 2009, p. 53)

The discovery learning specific to DDL is in keeping with the inquiry-based approach Coyle et al. (2010) hold to be a paradigm of CLIL. In DDL, the integration of content and language is consistently implemented as CLIL requires. Furthermore, through hypothesis formation and testing, DDL allows students to get acquainted with a descriptive view of the English grammar in contrast with the prescriptive view they are usually accustomed to: «The learner should be aware of grammar as a method rather than as a set of facts; that grammar is about knowing how to observe and how to interpret observations, rather than knowing what other people have observed» (Flowerdew 2012, p. 229).

Implementing content-specific DDL, the CLIL expert operationalized the use of corpora advocated by Gavioli and Aston (cited in McEnery et al. 2006, p. 195): «corpora should not only be viewed as resources which help teachers to decide what to teach, they should also be viewed as resources from which learners may learn directly». Students were expected to carry out DDL activities autonomously, and accordingly, enhance learner autonomy, which is a tenet of both DDL and CLIL.

In general, the corpus-driven data retrieved from the 'Apicoplast corpus' allowed the CLIL expert to identify key issues, such as subject-specific collocations and colligations, multiword units, and other lexico-grammatical features, to be focused on in DDL course-tailored language awareness activities. In discipline-specific DDL, lexical priming was especially enhanced through co-texts featuring concgrams, which were instrumental in activating semantic associations and thereby triggering target words. Examples of content-specific DDL activities created for the biology article are provided here (Figs. 80-82).

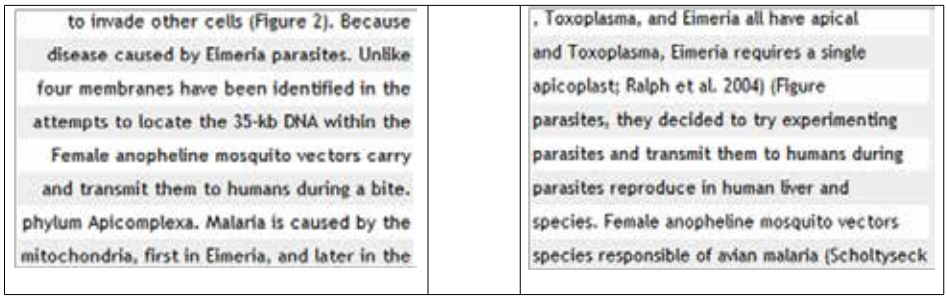
Read the concordance lines and write the missing content-specific word: _____



to invade other cells (Figure 2). Because		, Toxoplasma, and Eimeria all have apical
disease caused by Eimeria parasites. Unlike		and Toxoplasma, Eimeria requires a single
four membranes have been identified in the		apicoplast; Ralph et al. 2004) (Figure
attempts to locate the 35-kb DNA within the		parasites, they decided to try experimenting
Female anopheline mosquito vectors carry		parasites and transmit them to humans during
and transmit them to humans during a bite.		parasites reproduce in human liver and
phylum Apicomplexa. Malaria is caused by the		species. Female anopheline mosquito vectors
mitochondria, first in Eimeria, and later in the		species responsible of avian malaria (Scholtyseck

Figure 80. DDL activity

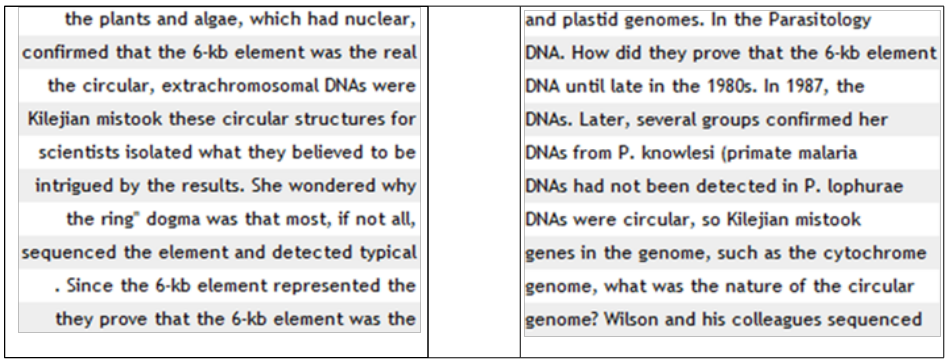
Read the concordance lines and write the missing content-specific word: _____



to invade other cells (Figure 2). Because		, Toxoplasma, and Eimeria all have apical
disease caused by Eimeria parasites. Unlike		and Toxoplasma, Eimeria requires a single
four membranes have been identified in the		apicoplast; Ralph et al. 2004) (Figure
attempts to locate the 35-kb DNA within the		parasites, they decided to try experimenting
Female anopheline mosquito vectors carry		parasites and transmit them to humans during
and transmit them to humans during a bite.		parasites reproduce in human liver and
phylum Apicomplexa. Malaria is caused by the		species. Female anopheline mosquito vectors
mitochondria, first in Eimeria, and later in the		species responsible of avian malaria (Scholtyseck

Figure 81. DDL activity

Read the concordance lines and write the missing content-specific word: _____



the plants and algae, which had nuclear,		and plastid genomes. In the Parasitology
confirmed that the 6-kb element was the real		DNA. How did they prove that the 6-kb element
the circular, extrachromosomal DNAs were		DNA until late in the 1980s. In 1987, the
Kilejian mistook these circular structures for		DNAs. Later, several groups confirmed her
scientists isolated what they believed to be		DNAs from P. knowlesi (primate malaria
intrigued by the results. She wondered why		DNAs had not been detected in P. lophurae
the ring" dogma was that most, if not all,		DNAs were circular, so Kilejian mistook
sequenced the element and detected typical		genes in the genome, such as the cytochrome
. Since the 6-kb element represented the		genome, what was the nature of the circular
they prove that the 6-kb element was the		genome? Wilson and his colleagues sequenced

Figure 82. DDL activity

Through DDL, students were sensitized to the specific features of English academic prose which were detected in their biology reading material. For instance, activities were devised to help students notice that verbs beginning with 're-' and verbs ending in '-ize' are recurrent in English scientific articles (Figs. 83 and 84).

Look at the following concordance lines. What do the red colour-coded verbs have in common? They all _____ with _____.

chloroplast? How could apicomplexans and plants be	related	? What follows is the amazing story of how
humans during a bite. Plasmodium parasites	reproduce	in human liver and blood cells, as well
Unlike Plasmodium and Toxoplasma, Eimeria	requires	a single host to complete its life cycle
parasites, that is, they cannot live or	reproduce	without a host. Discovery of the Apicoplast
malariologist, and Naomi Lang-Unnasch, who	researched	Toxoplasma. After several unsuccessful
the apicoplast puzzle is why apicomplexans	retained	a vestigial plastid despite losing the
chemicals affecting apicoplast metabolism	resulted	in parasite death. Second, parasites that
death. Second, parasites that are unable to	replicate	the apicoplast also die. Amazingly, in
with a chemically damaged apicoplast) while	remaining	in the infected host cell. However, the

Figure 83. DDL activity

Look at the following concordance lines. What do the base forms of the red colour-coded verbs have in common? They all _____ with _____.

apicoplasts contain their own DNA. When scientists	analyzed	apicoplast DNA, they were surprised to
Parasitology Today article, Wilson and his group	analyzed	the circular extrachromosomal DNA. They
added to the puzzle by scientists who were	characterizing	the extrachromosomal DNA molecules in apicomplexans
these extrachromosomal molecules. She soon	realized	that P. lophurae's extrachromosomal DNAs
possible? One hypothesis is that the apicoplast	synthesizes	a molecule that is needed for the infection

Figure 84. DDL activity

Furthermore, activities were devised to guide students to notice that nouns ending in «-tion» and «-ity» are a key feature of English academic prose (Figs. 85 and 86).

Look at the following word list. What do the nouns have in common? They all _____ with _____.

word	Freq
function	3
infection	3
organization	2
question	2
information	2
population	2
generation	1
translation	1
exception	1
combination	1
high-resolution	1
hybridization	1
Definition	1
restriction	1
interpretation	1
centrifugation	1
production	1
association	1
reproduction	1
Organization	1

Figure 85. Activity targeted to nouns formed with «-tion»

Look at the following word list. What do the nouns have in common? They all _____ with _____.

word	Freq
density	2
capability	1
possibility	1
University	1
identity	1
severity	1

Figure 86. Activity targeted to nouns formed with «-ity»

Through the Sketch Engine-retrieved corpus-driven data previously presented, the role of pre- and postmodifiers emerged as a pivotal element worth being dealt with in biology course-tailored language awareness materials. In this respect, the role of nouns (Fig. 87) serving as premodifiers was targeted in DDL language awareness activities.

In the following concordance lines, what premodifies red colour-coded nouns?

(a) Adjectives
(b) Nouns

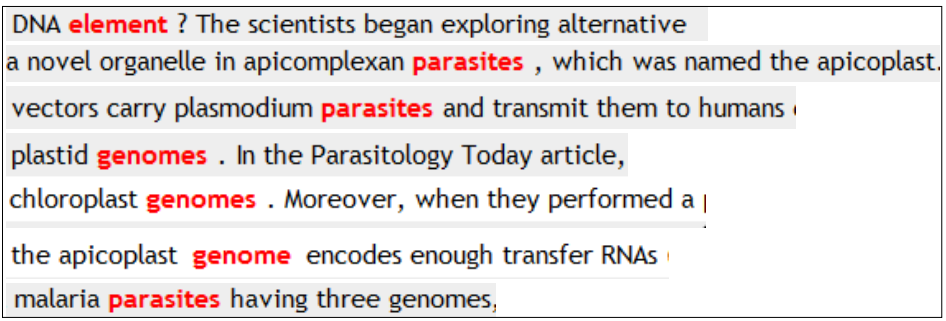


Figure 87. DDL activity

The DDL activity which follows (Fig. 88) was aimed at making students aware that attributive adjectives formed with «-al» are especially present in academic prose.

What do the adjectives used before the red colour-coded words have in common? They all _____ with _____.

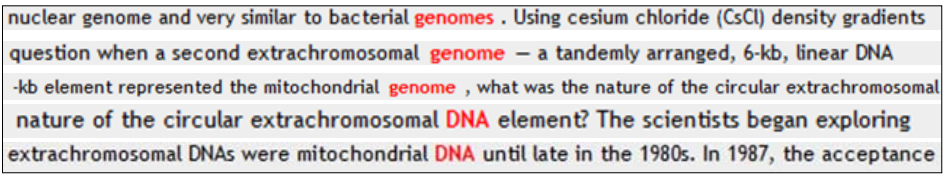


Figure 88. DDL activity

Prepositional postmodifiers were also focused on. 'Of-phrases', which were extensively used in the biology text, were targeted through DDL (Fig. 89).

What is the function of the red colour-coded 'of-phrases' in the following concordance lines?

(a) They premodify noun phrases
(b) They postmodify noun phrases

How do 'of-phrases' affect the amount of information provided?

(a) They help to package information densely
(b) They help to dilute information

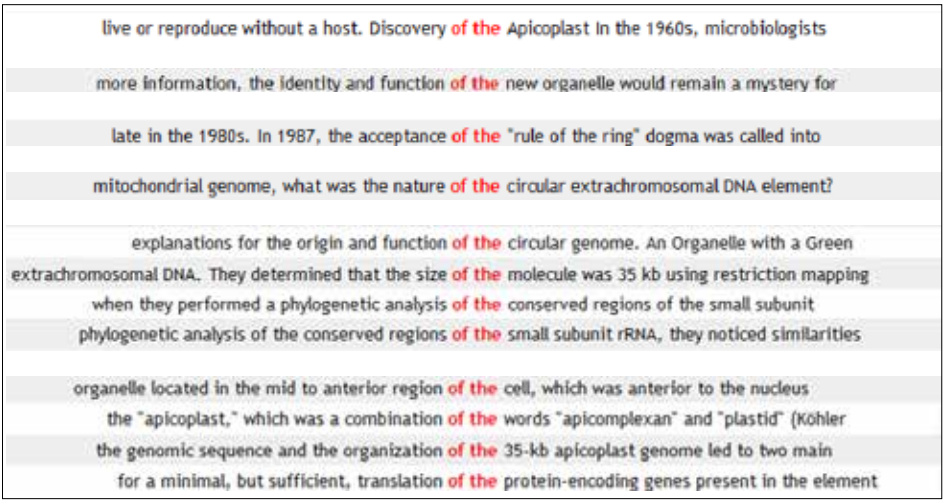


Figure 89. DDL activity

DDL was also used to sensitize students to the Academic Formulas featured in the biology article (Fig. 90).

Read the concordance lines and write the missing content-specific word:_____

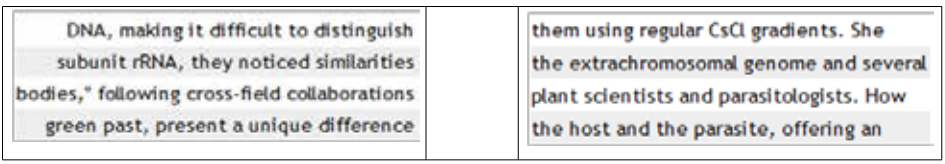


Figure 90. DDL activity

Upon completion of DDL activities targeting Academic Formulas, students were usually asked to classify the formulas just identified (Fig. 91).

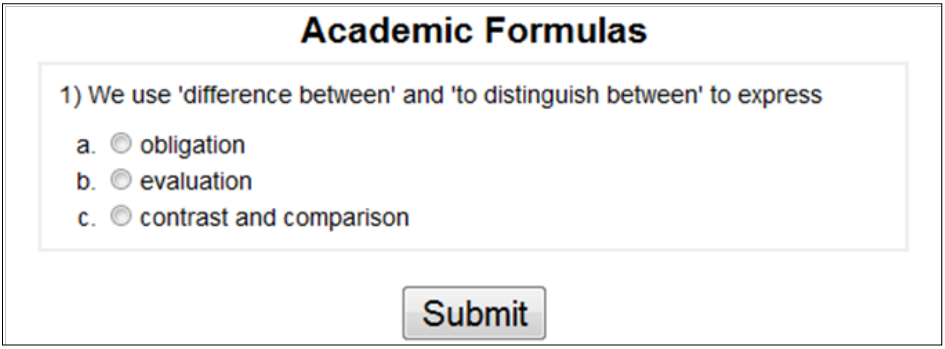


Figure 91. Learnclick-generated activity targeted to Academic Formulas

To help students work effectively in DDL, at the beginning the degree of inferring process required was decreased by providing learners in advance with key discipline-specific lexical items. As Flowerdew (2012, p. 206) suggests:

> Many practitioners seem to agree that induction-type tasks can present difficulties for students who are sometimes at a loss as to how to interpret the concordance lines. With this in mind, L. Flowerdew proposes consciousness-raising pedagogic mediation tasks in the form of hints to lead students toward interpretation of the corpus data.

To this end, before browsing concordance lines, students were introduced to key content-specific vocabulary by means of interactive glossaries as well as multiple choice and matching activities. Detecting recurring patterns in concordance outputs was thus made easier.

Overall, on the grounds of what has been presented thus far, the combination of DDL and CLIL seems especially effective:

> corpora are a way of enhancing and focusing the input to the student. They provide authentic data. They encourage reflection. They are well suited for consciousness-raising activities and for the training of inferencing. They stimulate the student to work actively and independently, and in this way they probably increase both the motivation of the student and the learning effect. (Johansson 2009, p. 38)

2.11 Reading academic prose: online tools

CLIL students can use various online resources to read articles in the target language. For example, if they so desire, learners can listen to online content-specific articles read in English using SoundGecko (`http://`

soundgecko.com), a free online tool which converts html articles into podcasts. To this purpose, students can also use NaturalReader (http://www.naturalreaders.com/index.htm), free text-to-speech software which can convert a wide range of texts, such as PDF files, MS Word files, and web pages into spoken words and into downloadable MP3 or WAV audio files. Besides English, the tool provides its service in various languages, such as French, German, and Spanish.

While reading articles online, students can use ProfessorWord (http://www.professorword.com), a free online tool which automatically highlights SAT (Scholastic Assessment Test) and ACT (American College Test) words – that is, the vocabulary especially tested in admission tests students need to take if they want to attend US universities – featured in html texts. Clicking on the highlighted words, a definition is provided in pop up windows (Fig. 92).

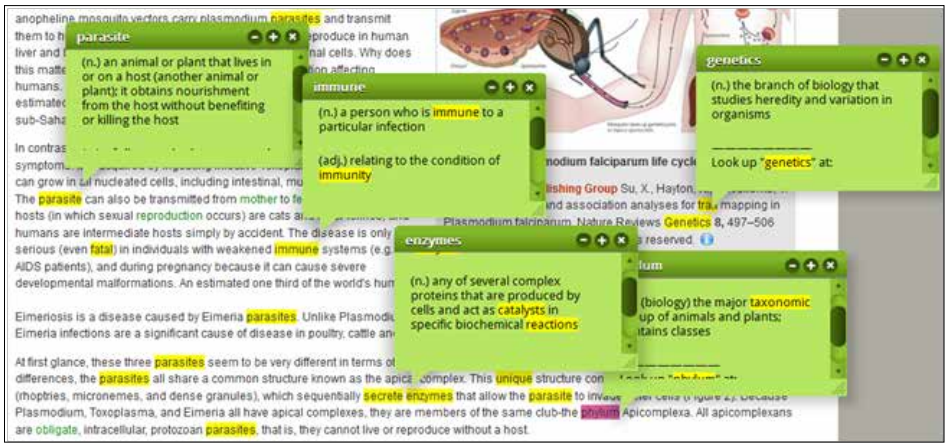


Figure 92. ProfessorWord pop-up definitions

Even words which are not highlighted can be searched, such as the subject-specific vocabulary item ‘organelle’ (Fig. 93).

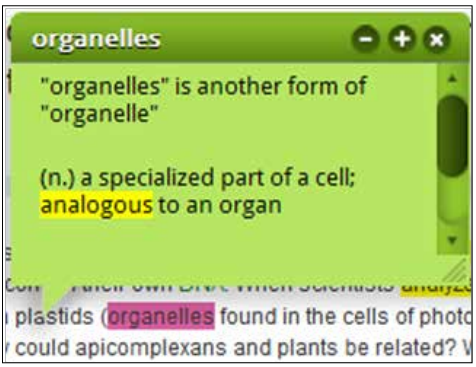


Figure 93. ProfessorWord pop-up definitions for non-SAT/ACT words

Definitions of words contained in pop-ups can also be generated (Fig. 94).

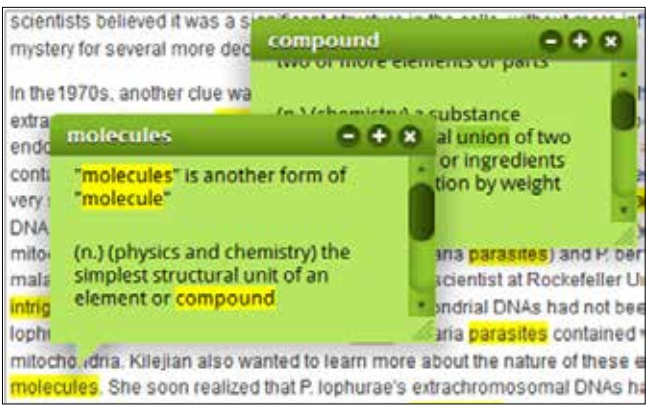


Figure 94. ProfessorWord pop-up definition of a word appearing in a pop up

Learners can save the ProfessorWord bookmarklet on their toolbar and simply click on it to activate the tool. If students print the page, all definitions of the words which they look up are printed as footnotes. Furthermore, learners can save all the words they look up as well as their definitions in a personalized word list. CLIL students are thus encouraged to work on subject-specific vocabulary building in a self-directed way.

Learners can also use Visuwords (`http://www.visuwords.com`), a free online graphic dictionary, to look up words – even technical ones, such as 'organelle' (Fig. 95). Words and concepts associated with the node word are retrieved and arranged in a circular netlike shape. By clicking on each word a definition is provided. Single words can be moved around to organize the concepts better.

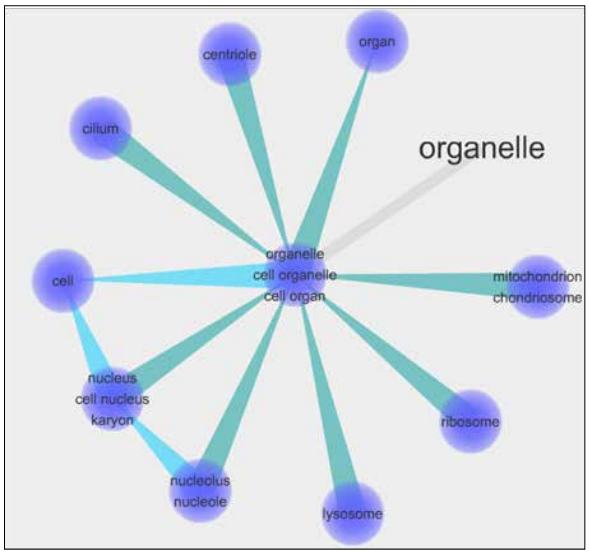


Figure 95. Circular, netlike output of Visuwords

Furthermore, the conceptual and grammatical connections between the different words are visualized by means of a wide array of linking segments (Fig. 96).

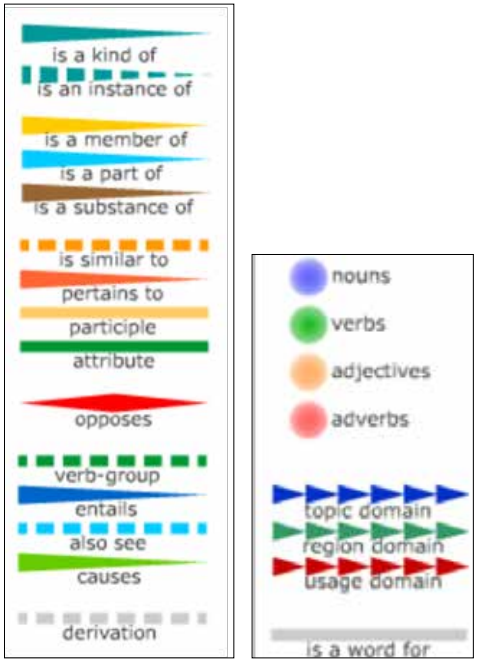


Figure 96. List of conceptual and grammatical connections provided in Visuwords

Visuwords is based on the open source database WordNet (http://wordnetweb.princeton.edu/perl/webwn) developed at Princeton University:

> WordNet is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. […] WordNet's structure makes it a useful tool for computational linguistics and natural language processing. (http://wordnet.princeton.edu)

Using this free online dictionary, CLIL students can develop their content-specific vocabulary autonomously.

2.12 Corpus-informed materials and academic skills

CLIL learners need to be trained to develop academic listening skills. The lecture used to this purpose was divided into six chapters. To this purpose, Open Educational Resources (OER) have been used:

> The Hewlett Foundation defines OER as:
> Teaching, learning, and research resources that reside in the public domain or have been released under an intellectual property license that permits their free use or re-purposing by others.
> Whilst OECD defines them as:
> Digitised materials offered freely and openly for educators, students and self-learners to use and reuse for teaching, learning and research. (OECD 2007, p. 133). (Conole 2012, p. 226)

A lecture on genomic conflict, part of an evolutionary biology series, was chosen to promote CLIL biology students' academic listening skills. The video lecture selected represented perfectly what OER are expected to be: «OER are intended to make 'high-quality educational material freely available worldwide in many languages'» (Conole 2012, p. 228).

Learners could watch the video lecture on genomic conflict in a self-directed way and at their own convenience. The genome lecture was divided into six chapters: namely, introduction, hierarchical selection and conflicts, segregation distortion, reproductive conflicts, reproductive conflict and mental disorders and evolutionary principles of conflict revolution. This clear-cut division of the input was likely to help students activate prior knowledge in terms of content and formal schemata before watching each subsequent part of the lecture:

> Called alternatively 'frames', 'scripts', or 'schemata', this background knowledge makes it possible to anticipate incoming information, relate to previous knowledge and thus make global sense of the text as it unfolds. Fillmore has distinguished three kinds of schemata: text schemata that deal with grammatical and cohesion structures; genre schemata that pertain to the rhetorical structures of different text genres (fairy tales, letters, newspaper articles, etc.); and content schemata that refer to the topic. (Kramsch 1993, p. 124)

Through the activation of genre and content schemata, the integration of top-down and bottom-up processes is enhanced, which is pivotal for scaffolding learners' processing of oral incoming input effectively. Furthermore, it is worth noting that in the OER video selected, the instructor uses a Power Point presentation during lecture delivery, which makes it easier for foreign language students to understand the input provided.

To create pre- and while-reading materials, various tools were used to analyze the lecture thoroughly. The transcript of the lecture was investigated using Sketch Engine. In particular, a DIY POS tagged corpus, named 'Genome corpus', was compiled using the genome conflict transcript and the intercollocability of the lecture text was analyzed. To this purpose, keywords were first retrieved from the 'Genome corpus' (Fig. 97).

Word list

Corpus: **Genome**

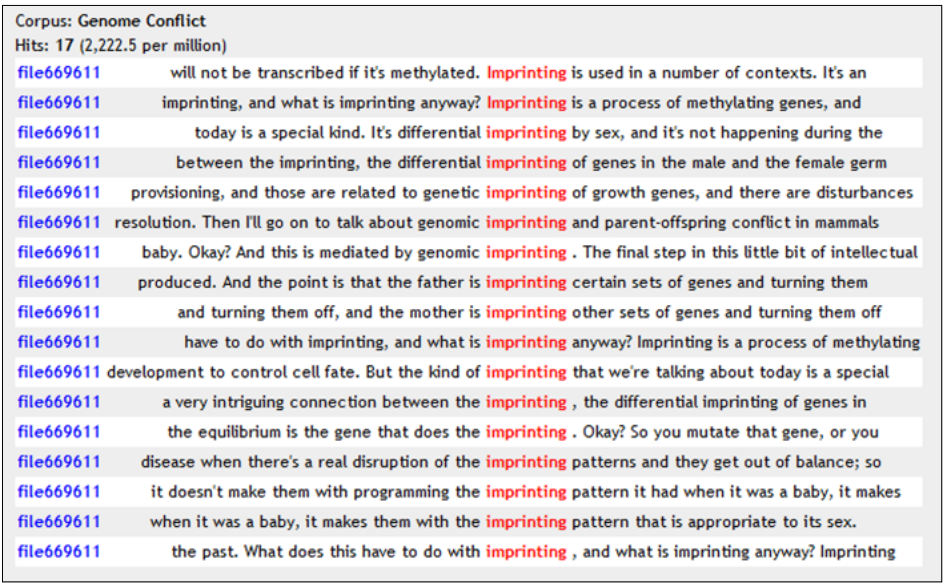
Reference corpus: **British National Corpus**

Page 1 Go Next >

	Genome		*British National Corpus*		
word	Freq	Freq/mill	Freq	Freq/mill	Score
imprinting	15.0	1943.0	59.0	0.5	1274.0
imprinted	13.0	1683.9	104.0	0.9	874.4
Wolbachia	7.0	906.7	5.0	0.0	869.0
organelles	8.0	1036.3	25.0	0.2	848.2
genome	16.0	2072.5	182.0	1.6	790.7
parent-offspring	6.0	777.2	1.0	0.0	771.3
Crespi	6.0	777.2	11.0	0.1	708.7
Badcock	6.0	777.2	14.0	0.1	691.9
placenta	8.0	1036.3	82.0	0.7	599.2
meiotic	6.0	777.2	38.0	0.3	581.3
Trivers	6.0	777.2	38.0	0.3	581.3
pathogen	6.0	777.2	46.0	0.4	551.9
cytoplasmic	11.0	1424.9	183.0	1.6	541.9
autism	4.0	518.1	25.0	0.2	424.5
segregation	14.0	1813.5	398.0	3.5	399.0
over-expressed	3.0	388.6	1.0	0.0	386.2
genomes	4.0	518.1	40.0	0.4	382.7
mitochondria	5.0	647.7	83.0	0.7	372.8
plasmids	6.0	777.2	125.0	1.1	368.1
fetus	6.0	777.2	128.0	1.1	363.5

Figure 97. Keyword list of the lecture

'Imprinting', the top ranking content-specific keyword, was investigated. Browsing left-sorted concordances of 'imprinting' (Fig. 98), the following discipline-specific collocations were detected: 'differential imprinting', 'genetic imprinting' and 'genomic imprinting'. Attributive derived adjectives, such as 'differential' and 'genomic', working as premodifiers of noun phrases, were identified in keeping with Biber's (2006) research on academic prose previously mentioned.

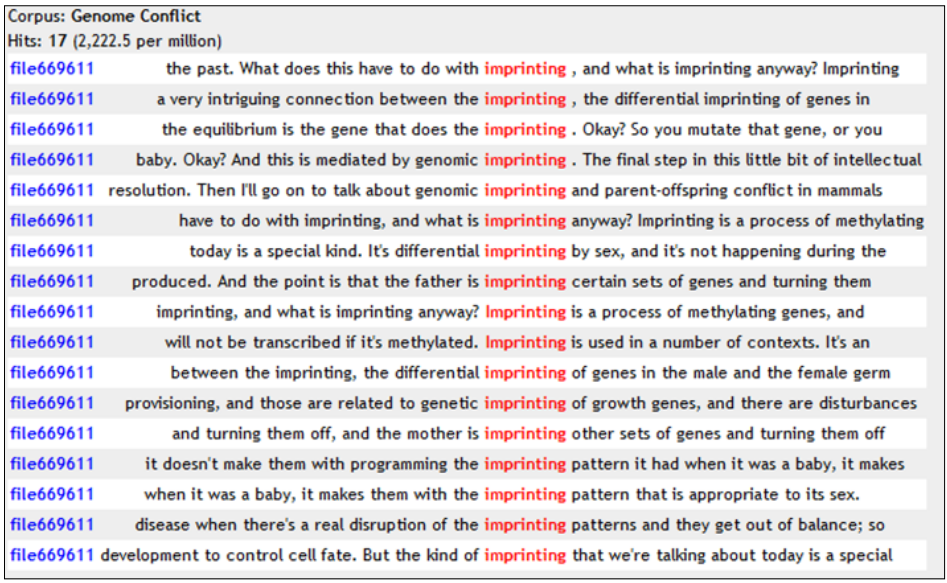


Corpus: Genome Conflict
Hits: 17 (2,222.5 per million)

file669611	will not be transcribed if it's methylated.	Imprinting	is used in a number of contexts. It's an
file669611	imprinting, and what is imprinting anyway?	Imprinting	is a process of methylating genes, and
file669611	today is a special kind. It's differential	imprinting	by sex, and it's not happening during the
file669611	between the imprinting, the differential	imprinting	of genes in the male and the female germ
file669611	provisioning, and those are related to genetic	imprinting	of growth genes, and there are disturbances
file669611	resolution. Then I'll go on to talk about genomic	imprinting	and parent-offspring conflict in mammals
file669611	baby. Okay? And this is mediated by genomic	imprinting	. The final step in this little bit of intellectual
file669611	produced. And the point is that the father is	imprinting	certain sets of genes and turning them
file669611	and turning them off, and the mother is	imprinting	other sets of genes and turning them off
file669611	have to do with imprinting, and what is	imprinting	anyway? Imprinting is a process of methylating
file669611	development to control cell fate. But the kind of	imprinting	that we're talking about today is a special
file669611	a very intriguing connection between the	imprinting	, the differential imprinting of genes in
file669611	the equilibrium is the gene that does the	imprinting	. Okay? So you mutate that gene, or you
file669611	disease when there's a real disruption of the	imprinting	patterns and they get out of balance; so
file669611	it doesn't make them with programming the	imprinting	pattern it had when it was a baby, it makes
file669611	when it was a baby, it makes them with the	imprinting	pattern that is appropriate to its sex.
file669611	the past. What does this have to do with	imprinting	, and what is imprinting anyway? Imprinting

Figure 98. Left-sorted concordances of 'imprinting

On the other hand, in right-sorted concordances of 'imprinting' (Fig. 99), the collocation 'imprinting' pattern was located.

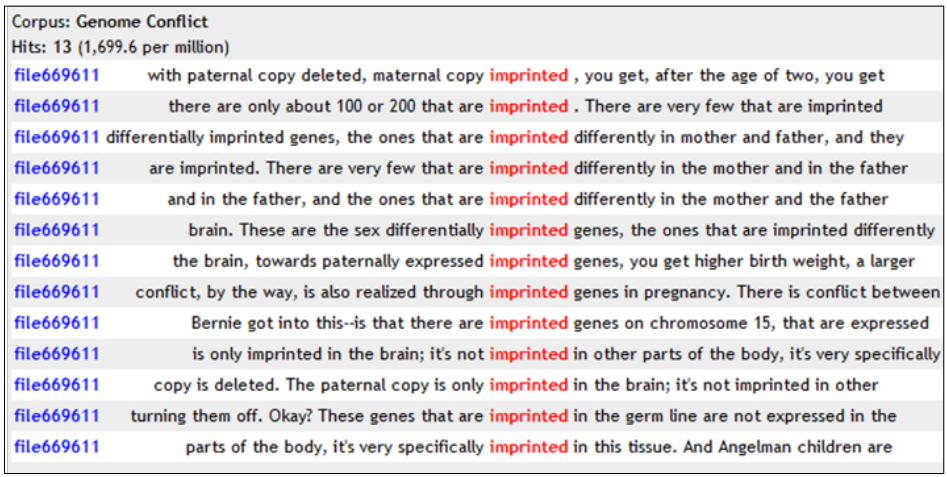


Corpus: Genome Conflict
Hits: 17 (2,222.5 per million)

file669611	the past. What does this have to do with	imprinting	, and what is imprinting anyway? Imprinting
file669611	a very intriguing connection between the	imprinting	, the differential imprinting of genes in
file669611	the equilibrium is the gene that does the	imprinting	. Okay? So you mutate that gene, or you
file669611	baby. Okay? And this is mediated by genomic	imprinting	. The final step in this little bit of intellectual
file669611	resolution. Then I'll go on to talk about genomic	imprinting	and parent-offspring conflict in mammals
file669611	have to do with imprinting, and what is	imprinting	anyway? Imprinting is a process of methylating
file669611	today is a special kind. It's differential	imprinting	by sex, and it's not happening during the
file669611	produced. And the point is that the father is	imprinting	certain sets of genes and turning them
file669611	imprinting, and what is imprinting anyway?	Imprinting	is a process of methylating genes, and
file669611	will not be transcribed if it's methylated.	Imprinting	is used in a number of contexts. It's an
file669611	between the imprinting, the differential	imprinting	of genes in the male and the female germ
file669611	provisioning, and those are related to genetic	imprinting	of growth genes, and there are disturbances
file669611	and turning them off, and the mother is	imprinting	other sets of genes and turning them off
file669611	it doesn't make them with programming the	imprinting	pattern it had when it was a baby, it makes
file669611	when it was a baby, it makes them with the	imprinting	pattern that is appropriate to its sex.
file669611	disease when there's a real disruption of the	imprinting	patterns and they get out of balance; so
file669611	development to control cell fate. But the kind of	imprinting	that we're talking about today is a special

Figure 99. Right-sorted concordances of 'imprinting'

Right-sorted concordances of 'imprinted' (Fig. 100) highlighted collocations such as 'imprinted genes'. Interesting lexical bundles, such as 'sex differentially imprinted genes' and 'paternally expressed imprinted genes', were also detected.

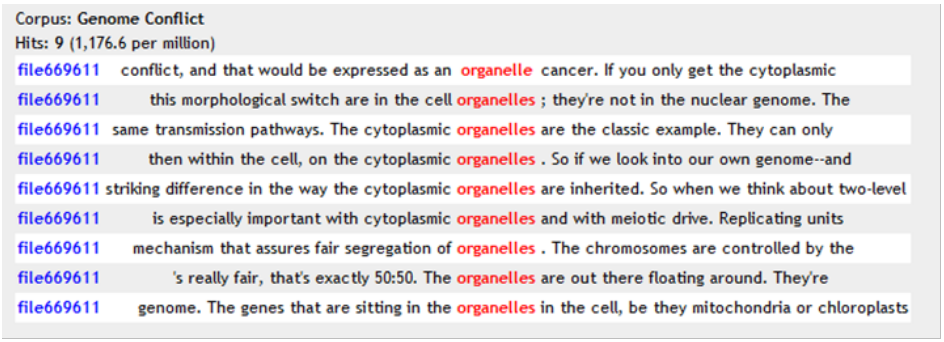


Corpus: Genome Conflict
Hits: 13 (1,699.6 per million)

file669611	with paternal copy deleted, maternal copy	imprinted	, you get, after the age of two, you get
file669611	there are only about 100 or 200 that are	imprinted	. There are very few that are imprinted
file669611	differentially imprinted genes, the ones that are	imprinted	differently in mother and father, and they
file669611	are imprinted. There are very few that are	imprinted	differently in the mother and in the father
file669611	and in the father, and the ones that are	imprinted	differently in the mother and the father
file669611	brain. These are the sex differentially	imprinted	genes, the ones that are imprinted differently
file669611	the brain, towards paternally expressed	imprinted	genes, you get higher birth weight, a larger
file669611	conflict, by the way, is also realized through	imprinted	genes in pregnancy. There is conflict between
file669611	Bernie got into this--is that there are	imprinted	genes on chromosome 15, that are expressed
file669611	is only imprinted in the brain; it's not	imprinted	in other parts of the body, it's very specifically
file669611	copy is deleted. The paternal copy is only	imprinted	in the brain; it's not imprinted in other
file669611	turning them off. Okay? These genes that are	imprinted	in the germ line are not expressed in the
file669611	parts of the body, it's very specifically	imprinted	in this tissue. And Angelman children are

Figure 100. Right-sorted concordances of 'imprinted'

Left-sorted concordances of 'organelle' (Fig. 101) retrieved collocations such as 'cell organelles' and 'cytoplasmic organelles'. In this case, one noun and one attributive derived adjective emerged as premodifiers of noun phrases.


Corpus: Genome Conflict
Hits: 9 (1,176.6 per million)

file669611	conflict, and that would be expressed as an	organelle	cancer. If you only get the cytoplasmic
file669611	this morphological switch are in the cell	organelles	; they're not in the nuclear genome. The
file669611	same transmission pathways. The cytoplasmic	organelles	are the classic example. They can only
file669611	then within the cell, on the cytoplasmic	organelles	. So if we look into our own genome--and
file669611	striking difference in the way the cytoplasmic	organelles	are inherited. So when we think about two-level
file669611	is especially important with cytoplasmic	organelles	and with meiotic drive. Replicating units
file669611	mechanism that assures fair segregation of	organelles	. The chromosomes are controlled by the
file669611	's really fair, that's exactly 50:50. The	organelles	are out there floating around. They're
file669611	genome. The genes that are sitting in the	organelles	in the cell, be they mitochondria or chloroplasts

Figure 101. Left-sorted concordances of 'organelle'

The Word Sketch for 'organelle' (Fig. 102) showed that the query term worked as (a) the object of the verb 'inherit', (b) the subject of the verbs 'float' and 'be', and (c) the modifier of 'cancer'. Furthermore, 'organelle' was premodified by the attributive derived adjective 'cytoplasmic' and the noun 'cell'. These data further confirmed the general collocational tendency already mentioned above and in line with Biber's (2006) studies.

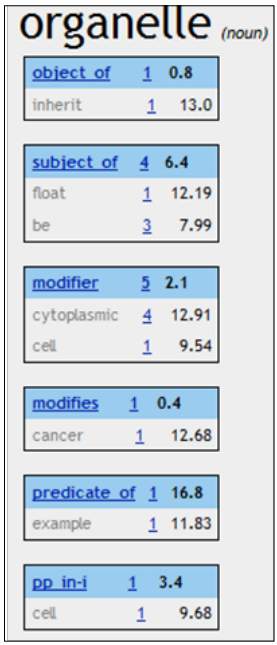
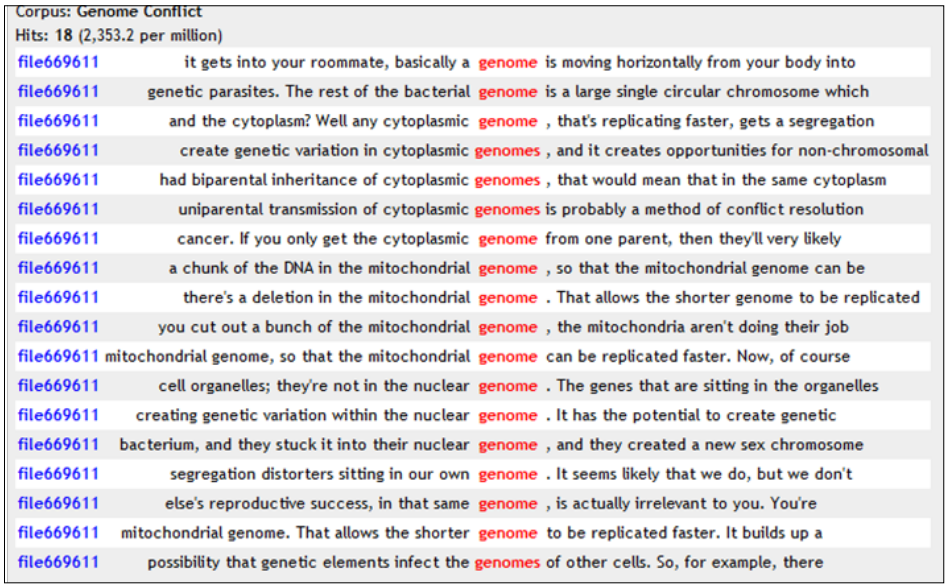


Figure 102. Word Sketch for 'organelle'

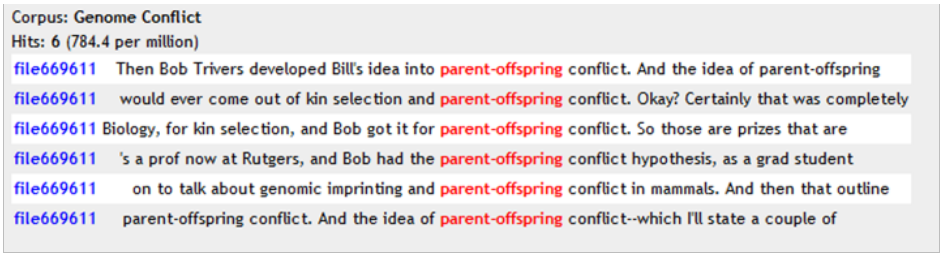
Left-sorted concordances of 'genome' (Fig. 103) generated collocations such as 'bacterial genome', 'cytoplasmic genome', 'mitochondrial genome' and 'nuclear genome'. In this case, premodifiers of the noun phrase were all attributive derived adjectives, a phenomenon which is specific of academic language as previously indicated.


Corpus: Genome Conflict
Hits: 18 (2,353.2 per million)

file669611	it gets into your roommate, basically a	genome	is moving horizontally from your body into
file669611	genetic parasites. The rest of the bacterial	genome	is a large single circular chromosome which
file669611	and the cytoplasm? Well any cytoplasmic	genome	, that's replicating faster, gets a segregation
file669611	create genetic variation in cytoplasmic	genomes	, and it creates opportunities for non-chromosomal
file669611	had biparental inheritance of cytoplasmic	genomes	, that would mean that in the same cytoplasm
file669611	uniparental transmission of cytoplasmic	genomes	is probably a method of conflict resolution
file669611	cancer. If you only get the cytoplasmic	genome	from one parent, then they'll very likely
file669611	a chunk of the DNA in the mitochondrial	genome	, so that the mitochondrial genome can be
file669611	there's a deletion in the mitochondrial	genome	. That allows the shorter genome to be replicated
file669611	you cut out a bunch of the mitochondrial	genome	, the mitochondria aren't doing their job
file669611	mitochondrial genome, so that the mitochondrial	genome	can be replicated faster. Now, of course
file669611	cell organelles; they're not in the nuclear	genome	. The genes that are sitting in the organelles
file669611	creating genetic variation within the nuclear	genome	. It has the potential to create genetic
file669611	bacterium, and they stuck it into their nuclear	genome	, and they created a new sex chromosome
file669611	segregation distorters sitting in our own	genome	. It seems likely that we do, but we don't
file669611	else's reproductive success, in that same	genome	, is actually irrelevant to you. You're
file669611	mitochondrial genome. That allows the shorter	genome	to be replicated faster. It builds up a
file669611	possibility that genetic elements infect the	genomes	of other cells. So, for example, there

Figure 103. Left-sorted concordances of 'genome'

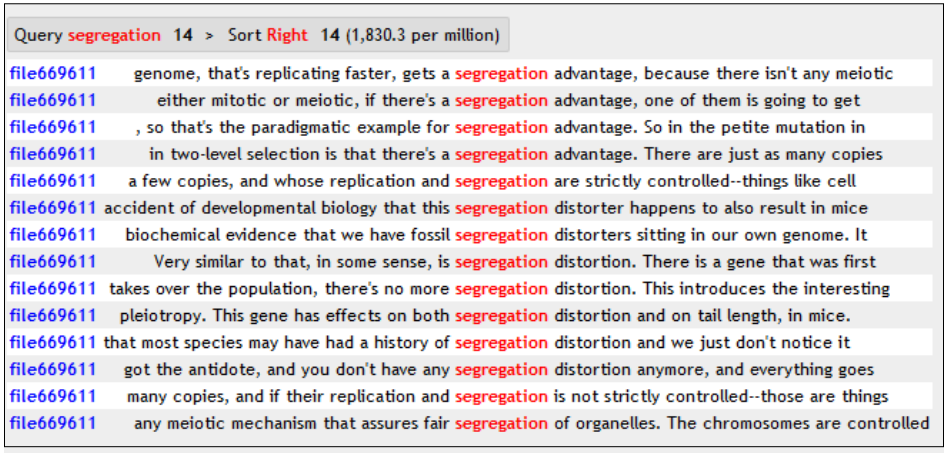
Right-sorted concordances of 'parent-offspring' (Fig. 104) retrieved the collocation 'parent-offspring conflict'.


Corpus: Genome Conflict
Hits: 6 (784.4 per million)

file669611	Then Bob Trivers developed Bill's idea into	parent-offspring	conflict. And the idea of parent-offspring
file669611	would ever come out of kin selection and	parent-offspring	conflict. Okay? Certainly that was completely
file669611	Biology, for kin selection, and Bob got it for	parent-offspring	conflict. So those are prizes that are
file669611	's a prof now at Rutgers, and Bob had the	parent-offspring	conflict hypothesis, as a grad student
file669611	on to talk about genomic imprinting and	parent-offspring	conflict in mammals. And then that outline
file669611	parent-offspring conflict. And the idea of	parent-offspring	conflict--which I'll state a couple of

Figure 104. Right-sorted concordances of 'parent-offspring'

Right-sorted concordances of 'segregation' (Fig. 105) showed collocations such as 'segregation advantage', 'segregation distorter' and 'segregation distortion'. In this case, collocations were characterized by nouns as pre-modifiers of the node word.

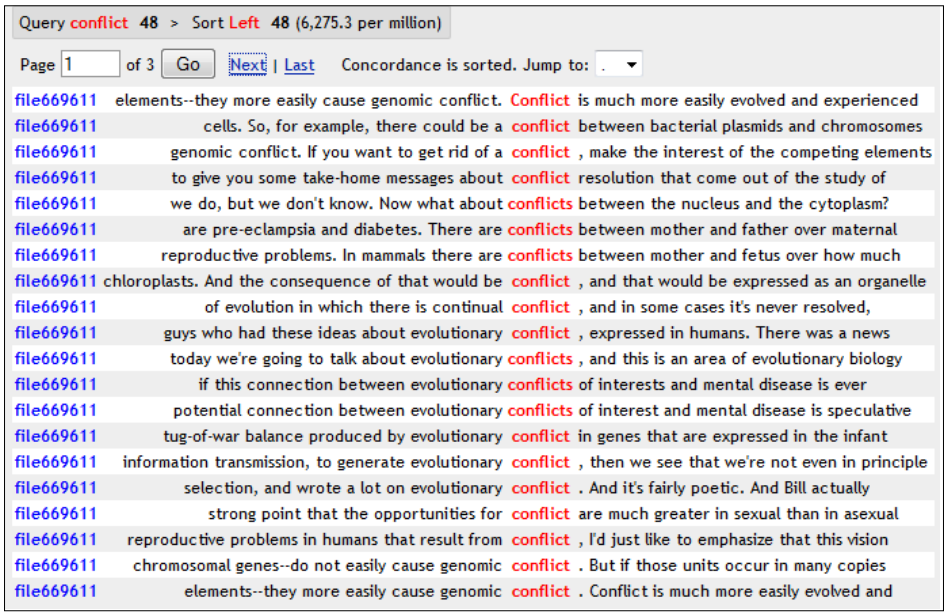


Query segregation 14 > Sort Right 14 (1,830.3 per million)

file669611	genome, that's replicating faster, gets a	segregation	advantage, because there isn't any meiotic
file669611	either mitotic or meiotic, if there's a	segregation	advantage, one of them is going to get
file669611	, so that's the paradigmatic example for	segregation	advantage. So in the petite mutation in
file669611	in two-level selection is that there's a	segregation	advantage. There are just as many copies
file669611	a few copies, and whose replication and	segregation	are strictly controlled--things like cell
file669611	accident of developmental biology that this	segregation	distorter happens to also result in mice
file669611	biochemical evidence that we have fossil	segregation	distorters sitting in our own genome. It
file669611	Very similar to that, in some sense, is	segregation	distortion. There is a gene that was first
file669611	takes over the population, there's no more	segregation	distortion. This introduces the interesting
file669611	pleiotropy. This gene has effects on both	segregation	distortion and on tail length, in mice.
file669611	that most species may have had a history of	segregation	distortion and we just don't notice it
file669611	got the antidote, and you don't have any	segregation	distortion anymore, and everything goes
file669611	many copies, and if their replication and	segregation	is not strictly controlled--those are things
file669611	any meiotic mechanism that assures fair	segregation	of organelles. The chromosomes are controlled

Figure 105. Right-sorted concordances of 'segregation'

Left-sorted concordances of 'conflict' (Figs. 106 and 107) retrieved collocations such as 'continual conflict', 'evolutionary conflict', 'genomic conflict', 'inter-genomic conflict' and 'parent-offspring conflict'. Overall, mostly attributive derived adjectives were detected as premodifiers of conflict.

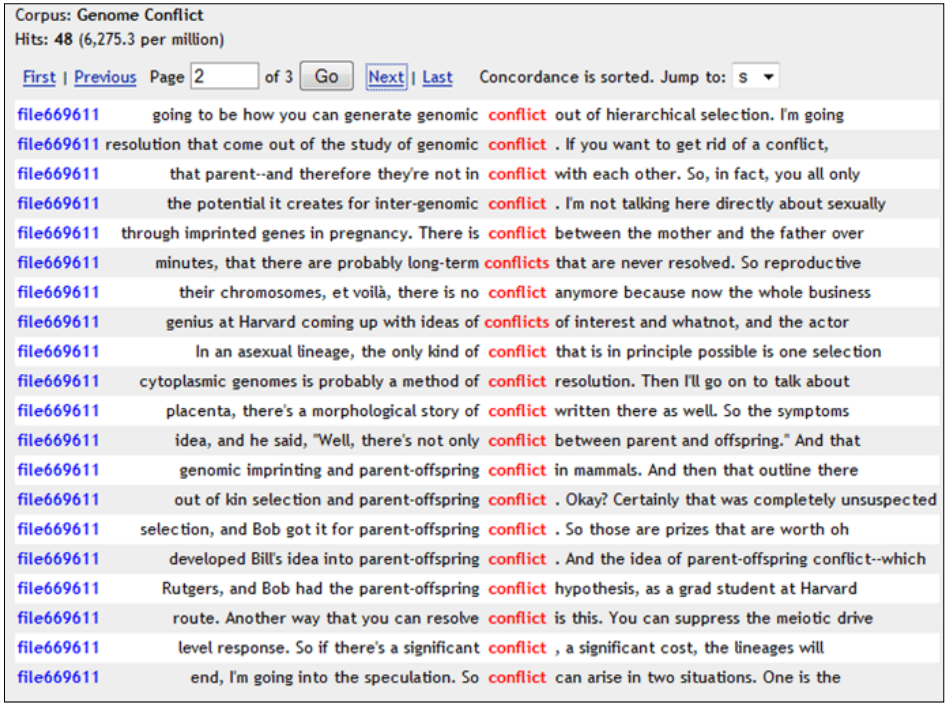


Query conflict 48 > Sort Left 48 (6,275.3 per million)

Page 1 of 3 Go Next | Last Concordance is sorted. Jump to:

file669611	elements--they more easily cause genomic	Conflict	is much more easily evolved and experienced
file669611	cells. So, for example, there could be a	conflict	between bacterial plasmids and chromosomes
file669611	genomic conflict. If you want to get rid of a	conflict	, make the interest of the competing elements
file669611	to give you some take-home messages about	conflict	resolution that come out of the study of
file669611	we do, but we don't know. Now what about	conflicts	between the nucleus and the cytoplasm?
file669611	are pre-eclampsia and diabetes. There are	conflicts	between mother and father over maternal
file669611	reproductive problems. In mammals there are	conflicts	between mother and fetus over how much
file669611	chloroplasts. And the consequence of that would be	conflict	, and that would be expressed as an organelle
file669611	of evolution in which there is continual	conflict	, and in some cases it's never resolved,
file669611	guys who had these ideas about evolutionary	conflict	, expressed in humans. There was a news
file669611	today we're going to talk about evolutionary	conflicts	, and this is an area of evolutionary biology
file669611	if this connection between evolutionary	conflicts	of interests and mental disease is ever
file669611	potential connection between evolutionary	conflicts	of interest and mental disease is speculative
file669611	tug-of-war balance produced by evolutionary	conflict	in genes that are expressed in the infant
file669611	information transmission, to generate evolutionary	conflict	, then we see that we're not even in principle
file669611	selection, and wrote a lot on evolutionary	conflict	. And it's fairly poetic. And Bill actually
file669611	strong point that the opportunities for	conflict	are much greater in sexual than in asexual
file669611	reproductive problems in humans that result from	conflict	, I'd just like to emphasize that this vision
file669611	chromosomal genes--do not easily cause genomic	conflict	. But if those units occur in many copies
file669611	elements--they more easily cause genomic	conflict	. Conflict is much more easily evolved and

Figure 106. Left-sorted concordances of 'conflict'



Corpus: Genome Conflict
Hits: 48 (6,275.3 per million)

First | Previous Page 2 of 3 Go Next | Last Concordance is sorted. Jump to: s

file669611	going to be how you can generate genomic	conflict	out of hierarchical selection. I'm going
file669611	resolution that come out of the study of genomic	conflict	. If you want to get rid of a conflict,
file669611	that parent--and therefore they're not in	conflict	with each other. So, in fact, you all only
file669611	the potential it creates for inter-genomic	conflict	. I'm not talking here directly about sexually
file669611	through imprinted genes in pregnancy. There is	conflict	between the mother and the father over
file669611	minutes, that there are probably long-term	conflicts	that are never resolved. So reproductive
file669611	their chromosomes, et voilà, there is no	conflict	anymore because now the whole business
file669611	genius at Harvard coming up with ideas of	conflicts	of interest and whatnot, and the actor
file669611	In an asexual lineage, the only kind of	conflict	that is in principle possible is one selection
file669611	cytoplasmic genomes is probably a method of	conflict	resolution. Then I'll go on to talk about
file669611	placenta, there's a morphological story of	conflict	written there as well. So the symptoms
file669611	idea, and he said, "Well, there's not only	conflict	between parent and offspring." And that
file669611	genomic imprinting and parent-offspring	conflict	in mammals. And then that outline there
file669611	out of kin selection and parent-offspring	conflict	. Okay? Certainly that was completely unsuspected
file669611	selection, and Bob got it for parent-offspring	conflict	. So those are prizes that are worth oh
file669611	developed Bill's idea into parent-offspring	conflict	. And the idea of parent-offspring conflict--which
file669611	Rutgers, and Bob had the parent-offspring	conflict	hypothesis, as a grad student at Harvard
file669611	route. Another way that you can resolve	conflict	is this. You can suppress the meiotic drive
file669611	level response. So if there's a significant	conflict	, a significant cost, the lineages will
file669611	end, I'm going into the speculation. So	conflict	can arise in two situations. One is the

Figure 107. Left-sorted concordances of 'conflict'

The Word Sketch for 'conflict' (Fig. 108) showed that the query term worked as (a) the object of verbs such as 'parent-offspring', 'generate' and 'cause', (b) the subject of verbs such as 'play', 'produce' and 'be' and (c) the premodifier of 'resolution' and 'hypothesis'.

Figure 108. Word Sketch for 'conflict'

Right-sorted concordances of 'cytoplasmic' (Fig. 109) generated collocations such as 'cytoplasmic bacterium', 'cytoplasmic genetic elements', 'cytoplasmic genome' and 'cytoplasmic organelles'.

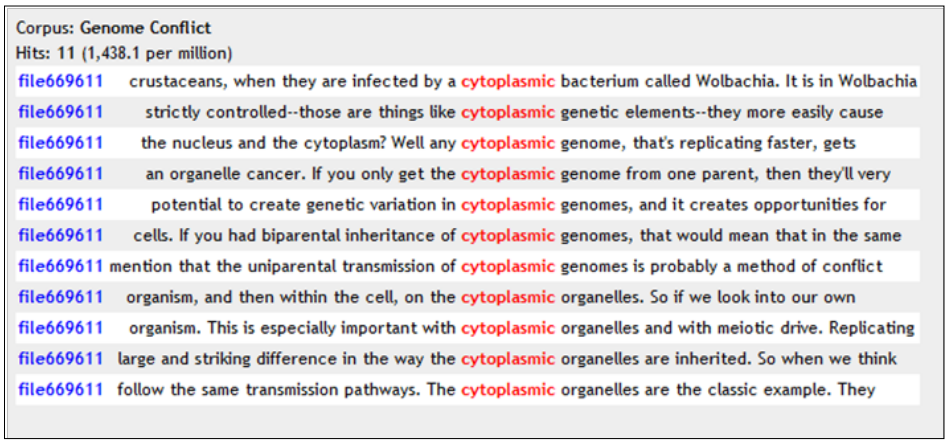
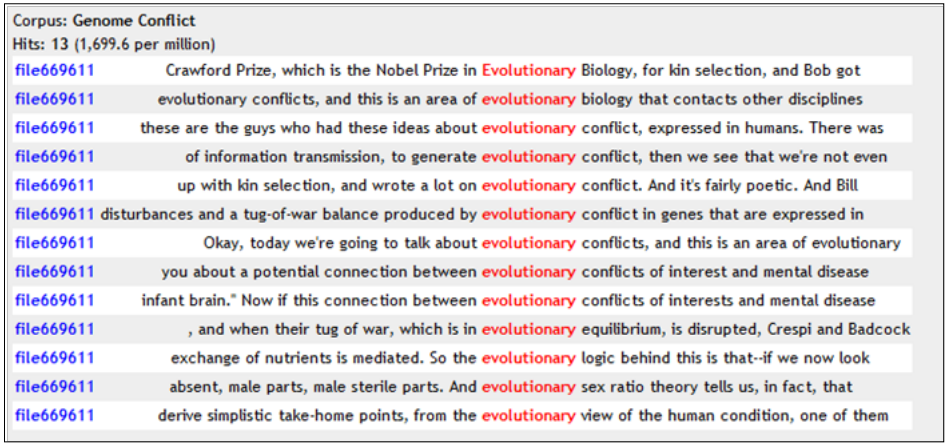


Figure 109. Right-sorted concordances of 'cytoplasmic'

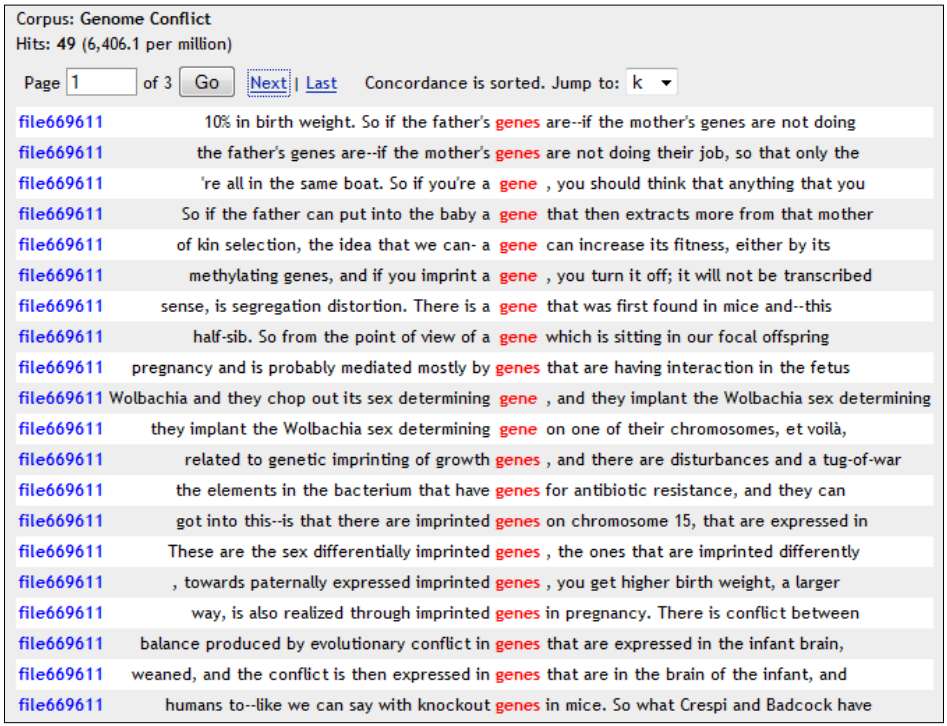
Right-sorted concordances of 'evolutionary' (Fig. 110) retrieved collocations such as 'evolutionary biology', 'evolutionary conflict', 'evolutionary equilibrium', 'evolutionary logic', 'evolutionary sex ratio theory' and 'evolutionary view'.


Corpus: Genome Conflict
Hits: 13 (1,699.6 per million)

file669611	Crawford Prize, which is the Nobel Prize in Evolutionary Biology, for kin selection, and Bob got
file669611	evolutionary conflicts, and this is an area of evolutionary biology that contacts other disciplines
file669611	these are the guys who had these ideas about evolutionary conflict, expressed in humans. There was
file669611	of information transmission, to generate evolutionary conflict, then we see that we're not even
file669611	up with kin selection, and wrote a lot on evolutionary conflict. And it's fairly poetic. And Bill
file669611	disturbances and a tug-of-war balance produced by evolutionary conflict in genes that are expressed in
file669611	Okay, today we're going to talk about evolutionary conflicts, and this is an area of evolutionary
file669611	you about a potential connection between evolutionary conflicts of interest and mental disease
file669611	infant brain." Now if this connection between evolutionary conflicts of interests and mental disease
file669611	, and when their tug of war, which is in evolutionary equilibrium, is disrupted, Crespi and Badcock
file669611	exchange of nutrients is mediated. So the evolutionary logic behind this is that--if we now look
file669611	absent, male parts, male sterile parts. And evolutionary sex ratio theory tells us, in fact, that
file669611	derive simplistic take-home points, from the evolutionary view of the human condition, one of them

Figure 110. Right-sorted concordances of 'evolutionary'

Left-sorted concordances of 'gene' (Fig. 111) produced collocations such as 'sex determining gene', 'growth genes', 'imprinted genes' and 'knockout genes'.


Corpus: Genome Conflict
Hits: 49 (6,406.1 per million)

Page 1 of 3 Go Next | Last Concordance is sorted. Jump to: k

file669611	10% in birth weight. So if the father's genes are--if the mother's genes are not doing
file669611	the father's genes are--if the mother's genes are not doing their job, so that only the
file669611	're all in the same boat. So if you're a gene , you should think that anything that you
file669611	So if the father can put into the baby a gene that then extracts more from that mother
file669611	of kin selection, the idea that we can- a gene can increase its fitness, either by its
file669611	methylating genes, and if you imprint a gene , you turn it off; it will not be transcribed
file669611	sense, is segregation distortion. There is a gene that was first found in mice and--this
file669611	half-sib. So from the point of view of a gene which is sitting in our focal offspring
file669611	pregnancy and is probably mediated mostly by genes that are having interaction in the fetus
file669611	Wolbachia and they chop out its sex determining gene , and they implant the Wolbachia sex determining
file669611	they implant the Wolbachia sex determining gene on one of their chromosomes, et voilà,
file669611	related to genetic imprinting of growth genes , and there are disturbances and a tug-of-war
file669611	the elements in the bacterium that have genes for antibiotic resistance, and they can
file669611	got into this--is that there are imprinted genes on chromosome 15, that are expressed in
file669611	These are the sex differentially imprinted genes , the ones that are imprinted differently
file669611	, towards paternally expressed imprinted genes , you get higher birth weight, a larger
file669611	way, is also realized through imprinted genes in pregnancy. There is conflict between
file669611	balance produced by evolutionary conflict in genes that are expressed in the infant brain,
file669611	weaned, and the conflict is then expressed in genes that are in the brain of the infant, and
file669611	humans to--like we can say with knockout genes in mice. So what Crespi and Badcock have

Figure 111. Left-sorted concordances of 'gene'

The Word Sketch for 'gene' (Fig. 112) showed that it worked as (a) the object of verbs such as 'imprint', 'determine' and 'mutate', (b) the subject of verbs such as 'exist' and 'mediate', and (c) the premodifier of 'expression'. Furthermore, 'gene' was modified by 'nuclear', 'knockout', 'reporter' and 'growth'.

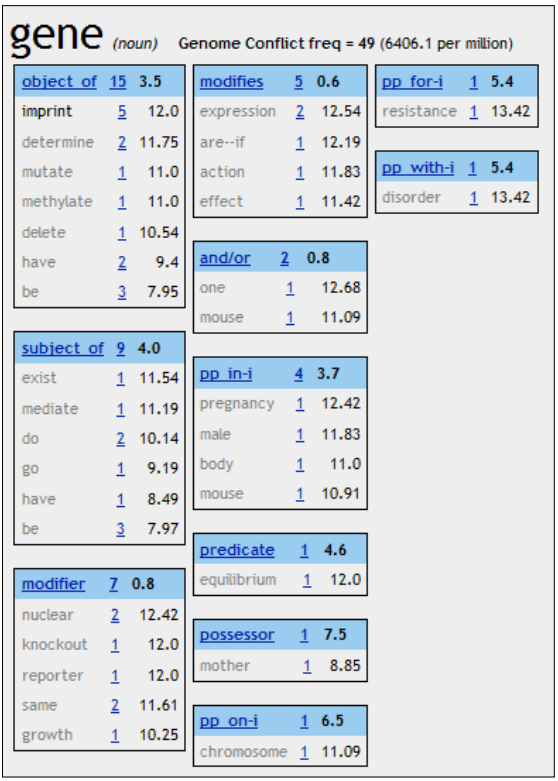


Figure 112. Word Sketch for 'gene'

The Word Sketch for 'cell' (Fig. 113) showed that the query term worked as (a) the object of verbs such as 'addict', 'become', 'destroy' and 'protect', (b) the subject of the verbs 'become' and 'divide', and (c) the premodifier of 'nucleus', 'lineage' and 'organelle'. Moreover, 'cell' was premodified by 'bacterial', 'liver' and 'brain'.

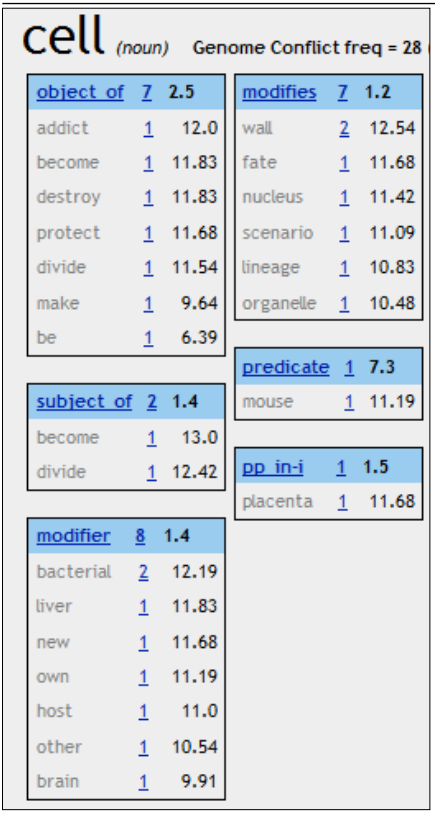


Figure 113. Word Sketch for 'cell'

The corpus-driven data retrieved were used, to a wide extent, to identify the content-specific vocabulary items and collocations which needed to be introduced before students watched the whole genome conflict lecture. As previously mentioned, to promote processing of subject-specific lexicon during discipline-specific lectures, learners need to be familiarized with key content-specific vocabulary items and collocations in advance. To this purpose, pre-viewing matching activities aimed to introduce key technical terms and their definitions were usually devised with Word Dynamo (Fig. 114).

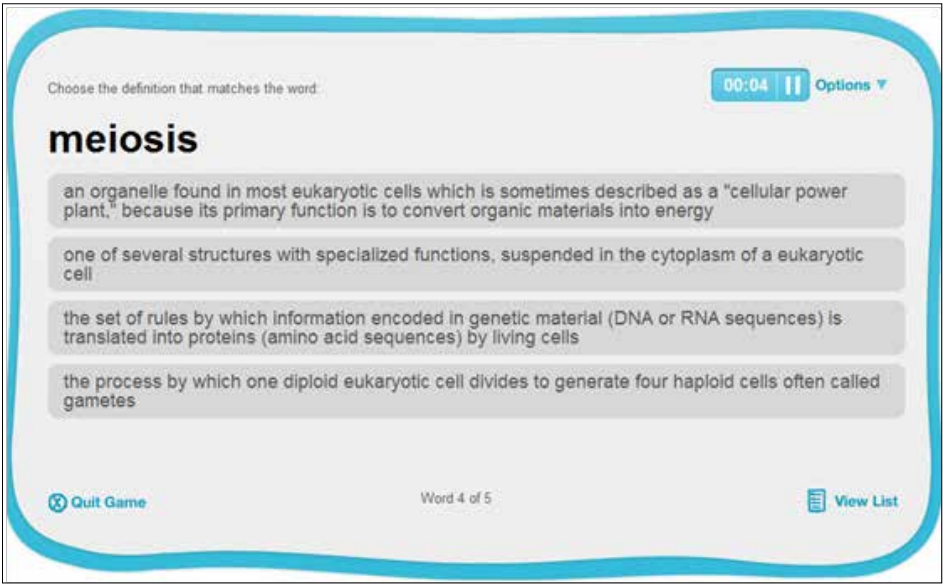


Figure 114. A Word Dynamo-generated multiple choice activity

In Word Dynamo-generated course-tailored activities, each targeted discipline-specific term is also provided in audio format thanks to text-to-speech software. Students are thus introduced to the pronunciation of subject-specific words prior to listening to the lecture, which makes it easier for learners to recognize the content-specific lexical items when the lecturer pronounces them. The automaticity of word processing is thus further enhanced, which is pivotal in listening tasks: «Part of mastering lexical items means that processing becomes more automatized. This is evidenced in faster recognition/comprehension speeds when listening» (Schmitt 2010, p. 106).

After answering each question, feedback is provided to learners (Fig. 115).

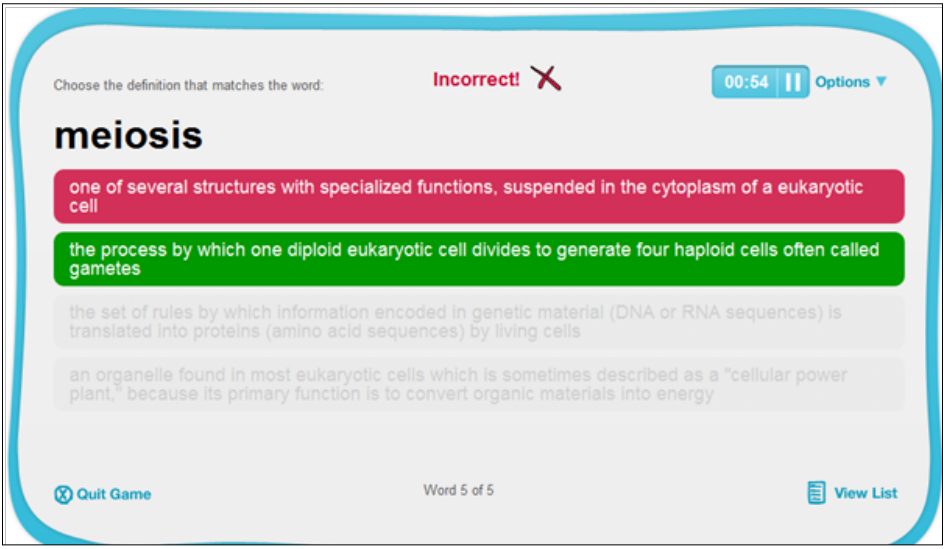


Figure 115. Feedback provided by Word Dynamo-generated activities

Upon completion of multiple choice activities, students can use customized flashcards to review the newly introduced concepts (Fig. 116). They can choose whether to first see only the key content-specific word and therefore provide their own definition, or to first see the definition and so provide the corresponding term. Students are expected to carry out the activities autonomously before watching the lecture. Working on key content-specific concepts is also instrumental in promoting lexical priming.

Figure 116. Word Dynamo-generated flashcard

Pre-viewing activities aimed to introduce subject-specific collocations are also created (Fig. 117).

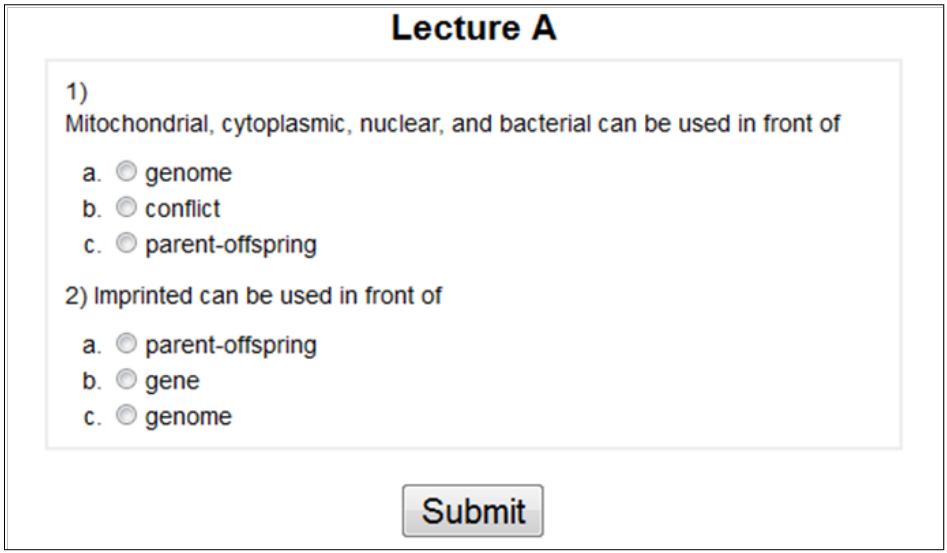


Figure 117. Learnclick-generated activity

Chunking the input decreases learners' cognitive load and enhances students' sense of self-efficacy. Thus, to enhance top-down processing and lexical priming, learners are expected to carry out customized activities prior to watching each chapter of the video. To this purpose, key discipline-specific terms and collocations are introduced through Word Dynamo-generated matching activities. The content-specific vocabulary items and collocations presented are chosen from the keyword lists and concordance outputs generated for each chapter of the video lecture. Some of the keyword lists retrieved for each chapter are provided below (Figs. 118 and 119).

evolutionary	2.0	5747.1	1021.0	9.1	569.0
faculty	1.0	2873.6	534.0	4.8	499.0
disgust	1.0	2873.6	613.0	5.5	444.7
reliable	3.0	8620.7	2144.0	19.1	428.7
conflicts	2.0	5747.1	1396.0	12.4	427.6

Figure 118. Keywords of chapter 1 of the lecture

Word list

Corpus: **Genome 2**

Reference corpus: **British National Corpus**

Page 1 Go Next >

	Genome 2		*British National Corpus*		
word	Freq	Freq/mill	Freq	Freq/mill	Score
genome	11.0	5620.8	182.0	1.6	2143.8
organelles	5.0	2554.9	25.0	0.2	2090.1
Wolbachia	4.0	2043.9	5.0	0.0	1957.7
plasmids	6.0	3065.9	125.0	1.1	1450.6
mitochondrial	5.0	2554.9	103.0	0.9	1332.5
petite	5.0	2554.9	120.0	1.1	1234.9
cytoplasmic	6.0	3065.9	183.0	1.6	1165.6
bacterium	5.0	2554.9	135.0	1.2	1160.0
meiotic	3.0	1533.0	38.0	0.3	1145.8
genomes	3.0	1533.0	40.0	0.4	1130.8
Basho	2.0	1022.0	2.0	0.0	1005.1
asexual	3.0	1533.0	75.0	0.7	919.3
mitochondria	3.0	1533.0	83.0	0.7	881.6
haiku	2.0	1022.0	20.0	0.2	868.2
two-level	2.0	1022.0	21.0	0.2	861.7
replication	4.0	2043.9	260.0	2.3	616.4
mutation	6.0	3065.9	446.0	4.0	616.4
yeast	5.0	2554.9	363.0	3.2	603.4
antidote	3.0	1533.0	182.0	1.6	584.9
bacterial	6.0	3065.9	520.0	4.6	544.2

Figure 119. Keywords of chapter 2 of the lecture

Furthermore, Text-Lex Compare was used to identify the content-specific words used for the first time in each chapter of the lecture (Tab. 2).

Table 2. Content-specific words introduced in each chapter retrieved with Text-Lex Compare

New content-specific words chapter one	New content-specific words in chapter two	New content-specific words in chapter three	New content-specific words in chapter four	New content-specific words in chapter five	New content-specific words in chapter six
Introduction	Hierarchical Selection and Conflicts	Segregation Distortion	Reproductive Conflicts	Reproductive Conflict and Mental Disorder	Evolutionary Principles of Conflict Resolution
conflicts 2	conflict 13	gene 5	placenta 7 fetus 6	syndrome 5	pathogens 1
	genome 11	homozygotes 3	embryo 5 imprinted 5	idiopathic 4	
	cell 10 cells 7	biochemical 1	fetal 4 germ 4	hyperlexic 1 hypogonadism 1	
	genetic 8	distorter 1 distorters 1	preeclampsia 2	lateralized 1	
	bacterial 6 cytoplasmic 6	gamete 1 genotype 1	vaccine 2	neurogenomic 1	
	plasmids 6	heterozygous 1	eclampsia 1		
	bacterium 5	meiosis 1	epigenetic 1		
	mitochondrial 5	mitochondrion 1	hormones 1		
	genes 5	nucleus 1	methylated 1 methylating 1		
	organelles 5	organelle 1	monogamous 1 monogamy 1		
	genomes 3 genomic 3	pleiotropy 1			
	meiotic 3 mitochondria 3				
	segregation 3				
	offspring 2				
	antibiotic 1 antibiotics 1				
	chloroplasts 1 chromosomal 1 chromosome 1 chromosomes 1				
	genetics 1				
	gynodioecious 1				
	mitotic 1				
	parasites 1				

The data generated for each chapter with Text-Lex Compare were compared with the content-specific lexical items featured in the Sketch Engine-generated keyword list of the 'genome corpus' mentioned above. The results emerging from the comparison of the two lists were used to select the content-specific vocabulary items and collocations to be introduced prior to watching each chapter of the video and focused on while-viewing activities.

2.12.1 While-viewing activities

As they are viewing the lecture on 'genomic conflict', students are encouraged to carry out online activities created with ProProfs (`http://www.proprofs.com`), an online tool. ProProfs-generated activities – such as multiple-choice, true/false and matching exercises (Figs. 120 and 121) – are embedded in the CLIL website; their URL address is also provided. Furthermore, paper-based versions of the activities are available together with keys.

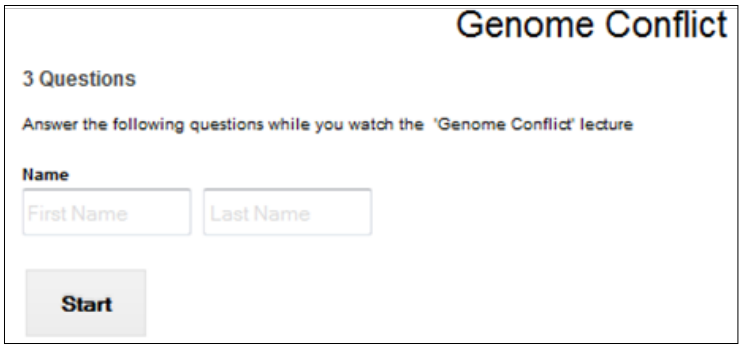


Figure 120. Students' login page

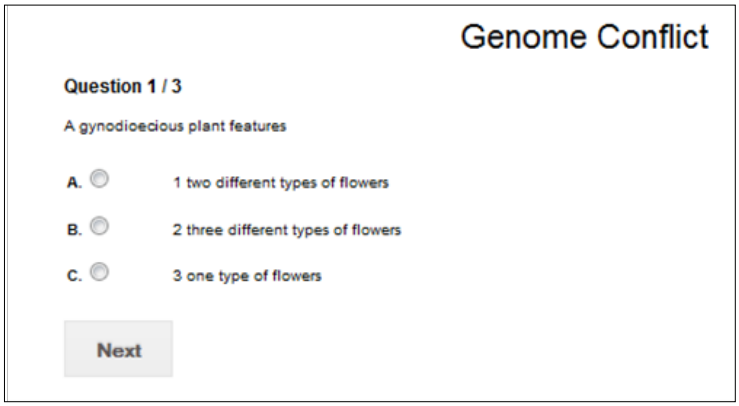


Figure 121. While-viewing activities created with ProProfs

After answering each question, learners are immediately provided with feedback (Fig. 122). As a result, students' listening process is consistently scaffolded; learners' monitoring skills are also enhanced.

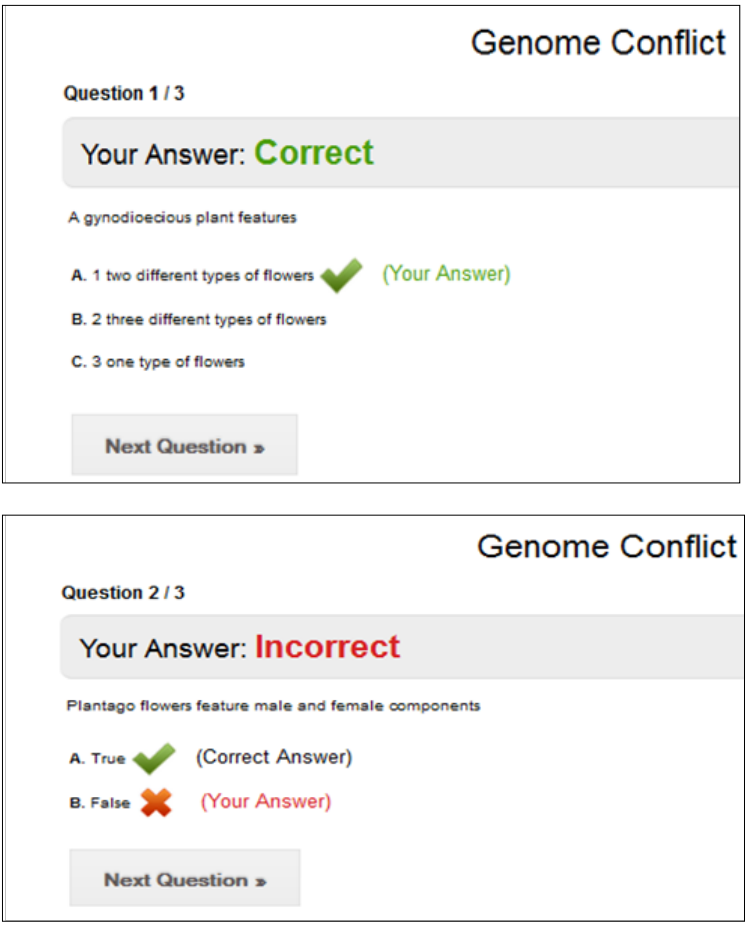
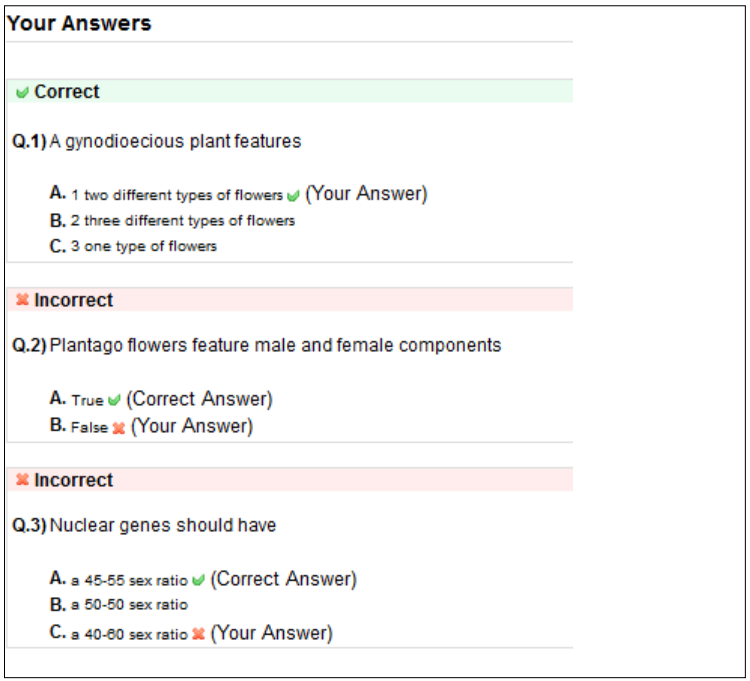


Figure 122. Feedback provided by ProProfs-generated activities

Upon completion of the while-viewing activities, students can view a complete report of their answers (Fig. 123). The report can be printed out.


Your Answers

Correct

Q.1) A gynodioecious plant features

A. 1 two different types of flowers (Your Answer)
B. 2 three different types of flowers
C. 3 one type of flowers

Incorrect

Q.2) Plantago flowers feature male and female components

A. True (Correct Answer)
B. False (Your Answer)

Incorrect

Q.3) Nuclear genes should have

A. a 45-55 sex ratio (Correct Answer)
B. a 50-50 sex ratio
C. a 40-60 sex ratio (Your Answer)

Figure 123. ProProfs-generated report

Reports are also available to course instructors, CLIL and language experts.

2.12.2 Note-taking skills in a CLIL setting

Note taking in a foreign language is a skill which needs to be acquired and consistently enhanced, especially in an academic setting. To this purpose, CLIL learners are trained to take notes while listening to professors deliver their lectures in English at a normal speech rate: «the average rate of speech is 150 words per minute, with peak rates of about 300 words per minute» (Schmitt 2010, p. 107). To this purpose, free online videos featuring instructors delivering content-specific lectures in English using Power Point presentations were used.

At the beginning, CLIL students are provided with guidelines, such as questions, on what to take notes on. Activities are created with Videonot.es (`http://www.videonot.es`), a free tool which allows users to take notes while watching videos and track the footage where the notes are inserted; the notes added are synchronized with the video (Fig. 124). Learners can watch the videos as many times as they want and type in the information requested when

they feel comfortable doing so. Clicking on the questions provided, students activate the corresponding footage, which makes note taking easier. Learners can share their notes with their peers and compare them. Students can also decide to let their peers edit their work if they so desire.

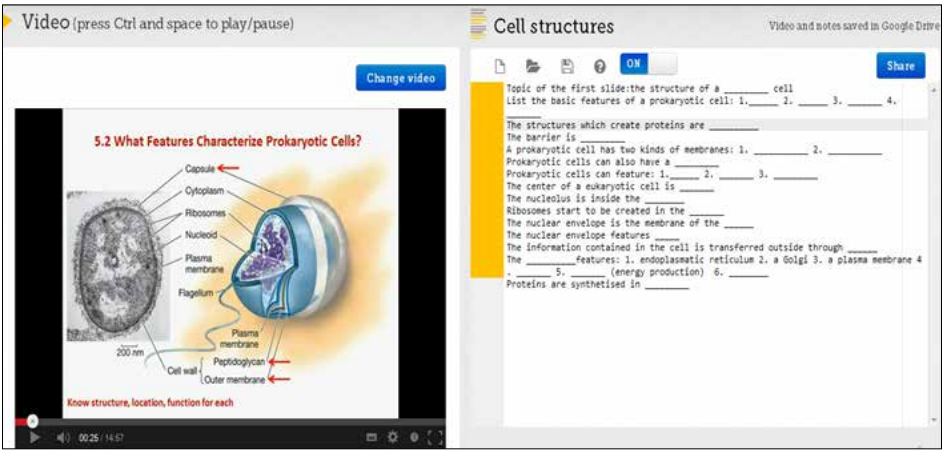


Figure 124. While-viewing activities created with Vidoenot.es

In order for learners to gradually develop note taking skills in English, the scaffolding provided – namely the questions devised by the CLIL expert together with content instructors – is gradually decreased. Students are thus trained step by step to take notes effectively during content-specific lectures delivered in English.

They can discuss their notes with their peers using Towtruck (`https://towtruck.mozillalabs.com`), a free text and voice chat collaborative tool (Fig. 125), embedded in the CLIL website, which enables students to collaborate in real time.

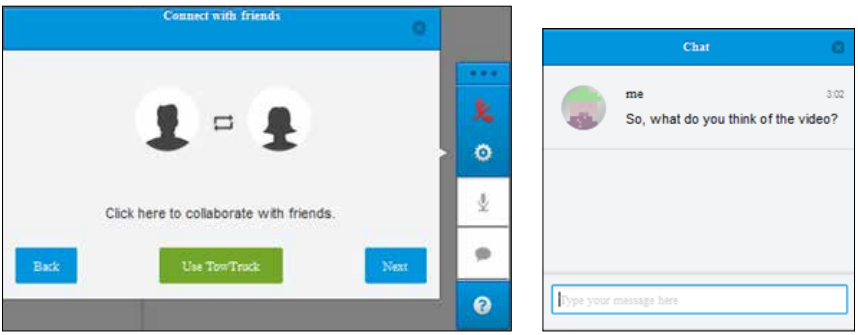


Figure 125. TowTruck collaborative space

After comparing their notes, learners are encouraged to collaboratively write a short summary of the targeted lecture using Draftin (`https://draftin.com`), a free online collaborative tool. Students can write their summaries co-editing each other's work. With this tool, learners get a high degree of control over the co-writing process, as the changes suggested by peers need to be approved (Fig. 126).

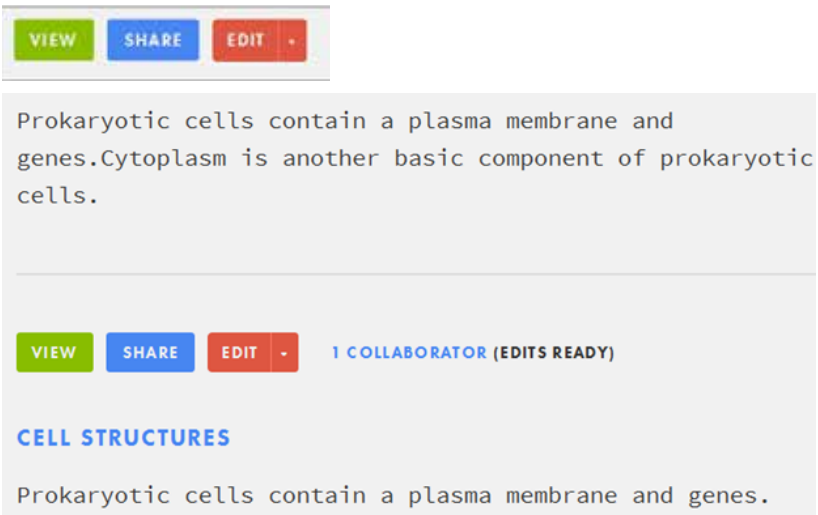


Figure 126. Co-written summary created with Draftin

Furthermore, in Draftin, students can access and compare the different drafts of their work and if they so desire, instructors can see the different changes made by each learner and to what extent each student collaborated in the writing process.

3 Online methodological support: the CLIL website

Summary 3.1. The 'Courses' pages: students' engagement in CLIL learning. – 3.2. Students and the CLIL Learning Center. – 3.3. Feedback on technical problems and suggestions. – 3.4. The online learning environment: activity profiling. – 3.5. The CLIL website design model.

In order to support instructors, the CLIL expert has devised a blended scaffolding system as previously mentioned. The face-to-face consultation mode has thus been integrated with online support delivered through a customized website.

Web-based services and spaces have been designed to enhance the adoption of new technologies in higher-education practices in keeping with the trends highlighted in the *2010 Horizon Report*, which identifies four trends as key drivers of technology adoption in higher education for the period 2010-2015:

- the abundance of online resources and relationships inviting a rethink of the educators' role in sense-making, coaching and credentialing;
- an increased emphasis on, and expectation of, ubiquitous, just-in-time, augmented, personalised and informal learning;
- the increased use of cloud computing challenges existing institutional IT infrastructures, leading to notions of IT support becoming more decentralised;
- the work of students being seen as more collaborative in nature and therefore there is potential for more intra- and inter- institutional collaboration (Conole, Alevizou 2010, p. 10).

The CLIL website (`http://clilteaching.weebly.com`), created and managed by the CLIL specialist, aims to provide methodological support to instructors throughout the course and to cater to students' needs before and after class. To ensure security and privacy for instructors and students, the site is password-protected. The visual and interaction design of the CLIL website is geared to effectively implement its multiple functions.

The site has been created with a free website builder, Weebly (`http://www.weebly.com`), which allows adding, uploading, and managing content elements such as text and audio files, photos, images, and videos. A mobile-

optimized version of the ads-free site is also available, which is a pivotal benefit for digital-age users. Additionally, through the detailed traffic stats tool, the number of visits and visitors to the site can be tracked.

Content elements, learning objects, add-ons, and plug-ins added to the site are created using free web-based tools or software. The CLIL expert made this decision to show instructors how they can successfully integrate CLIL face-to-face and out-of-class instruction with free technology-enhanced scaffolding at no further costs.

In this study, the following definition of a learning object has been adopted: «A digital self-contained and reusable entity, with a clear educational purpose, with at least three internal and editable components: content, learning activities and elements of context» (Chiappe et al. 2007). At a granularity level, the taxonomy elaborated by (Conole 2010b, p. 8) has been adopted to create the website:

> Littlejohn et al. (2008) identify four levels of granularity: i) digital assets – a single file, raw media asset, ii) information objects – structured aggregation of digital assets, iii) learning activities – tasks involving interactions with information to attend a specific learning outcome, iv) learning design – structure sequences of information and activities.

Firstly, the site displays the services offered by the Learning Center and presents the CLIL project (Fig. 127). Secondly, it provides instructors with methodological guidelines and free web-based tools. Thirdly, it makes customized, course-tailored materials and activities available to instructors and students. Fourthly, it connects instructors and students involved in the project, increasing interaction time in the target language between learners and faculty and setting up a community of practice. Fifthly, it introduces students to the methodological and language services provided by the Learning Center and connects learners with CLIL and English native-speaker experts. The systematic embedding of technologies in CLIL learning environments and cooperation through networking is thus enhanced.

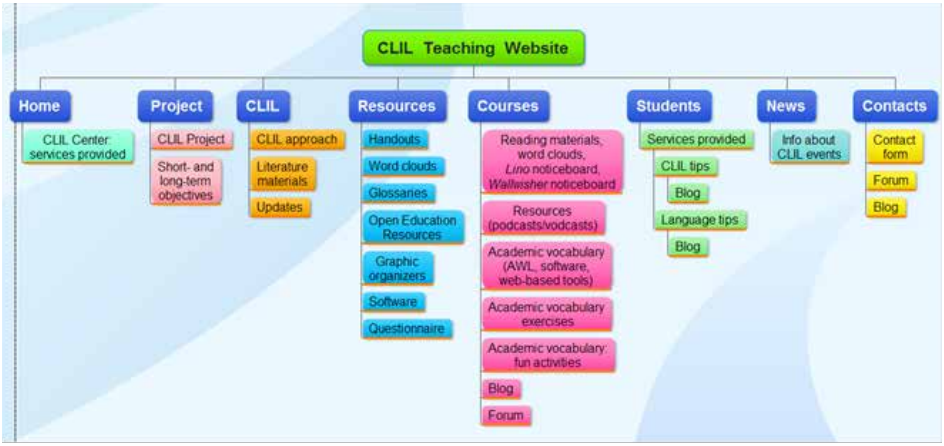


Figure 127. The CLIL website design

The website homepage contains a downloadable handout displaying the services provided by the CLIL Learning Center in detail. Visitors are introduced to the homepage content with a Voki (`http://www.voki.com`), a customized English-speaking avatar created with a free web-based tool that allows users to add voice with text-to-speech software. To make the CLIL website an English input-rich learning environment, English-speaking avatars are consistently used. Avatars introduce and describe objectives, learning objects, and services. Visitors can thus access spoken input in various English accents. Avatars speaking American, Australian, or British English introduce users to the different webpage contents.

The 'CLIL project' pages describe the theoretical underpinning of CLIL in higher education as well as its objectives and intended outcomes in the short and long run.

The 'CLIL' page features readings as well as audio-visual materials instructors can browse to get an understanding of the CLIL approach and its European dimension. Professors can advance their knowledge of CLIL core concepts through a consistently updated literature review archive displayed though BagTheWeb (`http://bagtheweb.com`), a free web-based tool that allows the CLIL expert to collect and share web resources. Updates on CLIL and new technologies are also provided through a customized Twitter account and a free online Paperli–based (`http://paper.li`) magazine, also accessible from mobile devices and managed by the CLIL expert.

The 'Resources' pages provide instructors with a wide range of web-based tools they can use as scaffolding for CLIL courses: «As has been noted by a range of researchers, (Puntambekar, Hübscher, Pea), scaffolding is used to describe a wide range of support provided by both humans and technology» (Luckin 2010, p. 41). An attempt has been made to encour-

age instructors to integrate new technologies into CLIL higher-education pedagogical practices:

> Although studies in OECD countries show that teachers may indeed be amongst the most skilled technology users, it appears that they are unable to take advantage of their competence and apply it to the way they teach (OECD: see also Blin and Munro; Zang). According to the OECD, three reasons emerge as the most salient for explaining this paradox:
> – the absence of appropriate incentives to use technology in the classroom and, more generally, getting involved in any innovation regarding teaching;
> – the dominant culture in the teaching profession is one of applied practice, which does not rely very much on research-based evidence to identify good teaching methodologies and strategies;
> – the observation that academic teachers lack the vision and the personal experience of what a technology-enhanced teaching could look like. (Conole, Alevizou 2010, pp. 21-22)

With regard to students, studies carried out show that learners are willing to interact with web-based technologies in formal education as long as they are effectively integrated with the course objectives:

> One of the main reasons cited by students for using technologies in their courses is convenience. Technologies are seen as adding value to courses, not as mechanisms for radical transformation. […] From the student perspective, technology is not necessarily a substitute, but a tool for added convenience and control. (Conole, Alevizou 2010, pp. 19-20)

Customized paper-based recommendations for instructors and students are also made available in the 'Resources' subsections. The tips for instructors and students are based on the CLIL university-based research project and carried out at the Brussels Faculty of Engineering previously mentioned. An avatar invites instructors to give 'recommendations for students' a few days before the first CLIL class to prepare them to autonomously scaffold their own learning experience in the new educational setting. Professors can thus enhance students' metacognition within a self-directed learning perspective.

Afterwards, Wordle – the free word cloud generating software previously mentioned – is introduced and a few examples of course-tailored word clouds are provided. Instructors can then browse various glossary URL addresses and links to educational resource materials. Links are also provided to Open Education Resources (OER) such as MIT (Massachusetts Institute of Technology) Open courseware, Open Yale courses, and open courses from Berkeley, Princeton, Stanford, and Harvard University.

An avatar describes graphic organizers and their effectiveness in a CLIL environment; links to free downloadable graphic organizers are also inserted. Instructors are then introduced to Cacoo (http://www.cacoo.com), a free web-based presentation tool that allows users to create diagrams, use stencils, make drawings, take screenshots, upload images, and write texts collaboratively. Exploratree (http://www.exploratree.org.uk), free cloud-based software, is presented next. With Exploratree, instructors can easily create customized graphic organizers or interactive thinking guides that «support the thinking or working through of an issue, topic or question and help to shape, define and focus an idea and support the planning needed to research it further» (http://www.exploratree.org.uk). Thinking guides can be printed, sent by mail, or shared through a URL posted on the teacher's homepage or on the CLIL website. With Exploratree, instructors can create interactive thinking guides to scaffold the comprehension of course reading materials and lectures. Thinking guides can also be used to scaffold tasks aimed to activate higher-order thinking skills.

Various web-based tools are introduced. Instructors can use Pearltrees (http://www.pearltrees.com), a free social bookmarking tool, to collect and organize content-specific material found on the web through a user-friendly visual interface. Instructors can embed their Pearltrees on the web-based notice boards provided in the 'Courses' section of the website. Learners can manipulate the Pearltrees embedded by instructors, aggregating further materials, or they can create their own domain-specific Pearltrees to be shared with peers and instructors. This social bookmarking tool is suitable to promote student-generated content and sharing. Pinterest (http://www.pinterest.com), a pinboard-style content sharing service, is also presented; instructors can use the tool to provide image-enriched reading lists and to promote students' creation and sharing of collaborative infographic-based artifacts. Pinterest-based content can be tweeted, emailed, pinned, repinned, and embedded in the CLIL website. Next, NaturalReader, which was previously mentioned, is presented along with SoundGecko.

Further, MS Word downloadable evaluation questionnaires are available. Instructors can administer the questionnaires to students at the end of the CLIL experience. Students' evaluations can help instructors improve their teaching and the Learning Center staff to address students' needs and expectations.

3.1 The 'Courses' pages: students' engagement in CLIL learning

The 'Courses' pages feature learning objects and digital spaces customized for specific courses by the CLIL methodology expert. Online learning objectives, materials, and learner engagement are aligned to course

structures to enable students to actively achieve the expected learning outcomes: «There is a growing body of research literature that clearly demonstrates that active engagement in online learning environments is effective and highly desired by students (Creanor et al.; Weaver et al.)» (Tynan, Barnes 2011, p. 371). In course-tailored pages, software-powered Voki English-speaking avatars give learners clearly stated instructions on how and when to use the tools, materials, and spaces provided. Students can thus experience an online English input-rich learning environment. Avatars are effective tools with regard to autonomous learning and are instrumental in promoting self-directed learning: «Learner control has a huge role to play and is consistent with the intentions behind Web 2.0 technologies (Dron)» (Tynan, Barnes 2011, p. 371).

In the 'Courses' section, digital spaces and activities have been devised in keeping with De Freitas and Conole's research findings (cited in Conole, Alevizou 2010, p. 9):

> De Freitas and Conole suggest five broad technological trends that are likely to have a significant impact on education:
> – a shift towards ubiquitous and networked technologies;
> – the emergence of context and location aware devices;
> – the increasingly rich and diverse different forms of representations and stimulatory environments possible;
> – a trend towards more mobile and adaptive and devices;
> – a technological infrastructure which is global, distributed and interoperable.

On the introductory page of each course section, a notice board provides instructors with a space to post news, assignments, questions, comments, and polls in the target language using multimodal sticky notes. With Lino, a free web-based application, teachers can post editable stickies on a shared canvas while also creating and sending them via email; sticky notes can feature text as well as embedded videos, images, and links. Instructors can embed Voki avatars, attach files, and email stickies to students to keep them posted on course activities, while teachers can send students reminders of assignment due dates. Deadlines can also be made visible in the sticky notes. Furthermore, learners can download a Windows widget to see stickies and due dates on their desktop; alarm settings can be set by students within a selected time frame before the deadline. Dashboard widgets for Mac users are additionally available. Sticky notes can be peeled off when they are no longer useful and instructors can embed polls in stickies. Flisti (`http://flisti.com`), a free web-based poll-authoring tool, enables instructors to get feedback from learners and enables them, through polling, to monitor their class. Instructors and learners can access Lino canvases and post notes by iPhone and iPad and professors are

notified whenever someone posts a sticky. Overall, an attempt is made to encourage instructors to use digital tools to interact with students.

Instructors can use Dipity (`http://www.dipity.com`) or Timetoast (`http://www.timetoast.com`), free web-based timeline builder tools, to create digital timelines about the topics presented in class. Events can be either text- or audio/video-based. Students can thus keep track of the main topics tackled in class, shown chronologically, and are provided with topic-connected multimodal resources.

Customized word clouds, which can be printed, are saved in the public Wordle gallery and are embedded in each instructor's course-tailored pages. Clicking on them, users are redirected to the source page where embed codes and links can be retrieved. Instructors can provide students with word clouds before class to help them zero in on key subject-specific vocabulary items or in class to trigger students' hypotheses on the lesson topic. Word clouds can also be used at the end of class to summarize the key concepts introduced. Students and instructors can cooperatively create new word clouds based on learners' perceptions of the importance of the specialized key concepts introduced. These word clouds can be compared with those previously created using Sketch Engine statistically driven data; a discussion based on the differences and likenesses detected may follow. Word clouds represent flexible teaching materials that instructors can use in various ways in CLIL teaching.

Students can access print-based reading materials on their course-tailored pages. Avatars inform students that they can listen to the reading assignments being read in the foreign language selected using the floating bar of NaturalReader, the free text-to-speech software mentioned earlier. Students can synchronize the text-to-speech software on the sentence and word level and listen to the text while seeing the synchronized highlighting of both the sentence and word being read. Also, texts converted into downloadable MP3 or WAV audio files can be listened to on the move. Learners can, as a result, access input in multimodal formats. The use of SoundGecko, ProfessorWord and Visuwords is also encouraged through avatars.

Either before or after class, students can write questions in the target language about the assigned study materials or the topics introduced in class on a customized Padlet-generated (`http://padlet.com`) notice board. Padlet is a free web-based application that allows users to express their ideas using multimedia sticky notes. Students can post and share their questions and comments asynchronously. Sticky notes can feature text, hyperlinks, videos, audios, and images. Learners can also ask questions using Voki speaking avatars that they can embed in the notice board provided; students who do not feel comfortable asking questions in English in class can use sticky notes or avatars to become active participants. Learners can use text-to-speech software or record their voice with a Voki audio-casting tool, and can decide whether to stay anonymous or provide their names in sticky notes and avatars.

Learners can also ask instructors questions in the target language using a voicemail. To leave messages, learners can use the voicemail widget embedded in each 'Courses' section. To activate the service, students have only to press the button appearing on the left-hand side in the middle-right of the screen. Instructors are notified by mail of incoming messages. If students write their email addresses while leaving a voice message, instructors can record their answers in the target language and send them as voice messages to students; learners will be notified of the incoming voicemail. This service was implemented using SpeakPipe (http://www.speakpipe.com). Professors are instructed on how to create an account, which is necessary to set up the service. The CLIL expert is available to embed the SpeakPipe widgets in instructors' course pages.

The 'Resources' page is a repository of technology-enhanced materials in the target language learners can use to: (a) be introduced to key content-specific vocabulary items in a self-directed way; (b) carry out reading comprehension, while-viewing, and note-taking activities autonomously; (c) and engage in Data-Driven Learning independently. In the 'Resources' page, students can also find materials to review: for instance, key content concepts through rich-media presentations. To provide learners with further support and guidance, instructors are invited to publish podcasts in the foreign language on a regular basis, that is, with a daily or weekly frequency, on a Lino notice board. Podcasts can thus be accessed 24/7, which makes learning time and space independent. In our CLIL course design, podcasts are conceived in keeping with:

> the broader definition of podcast suggested by Salmon et al., who described a podcast as a digital media file that plays sound or video, is made available from a website and can be downloaded and played on a computer or a portable player. (Nie et al. 2010, p. 106)

In CLIL course planning, podcasts are not used as stand-alone but, as suggested by the IMPALA (Informal Mobile Podcasting and Learning Application) podcast development model, they are integrated with other media in a VLE, Virtual Learning Environment (http://www.atimod.com/podcasting/PDModel.html) and aligned with the course objectives. Perceiving podcasts as an essential part of instruction is likely to promote learners' motivation of engagement. Through podcasting in the target language, flexibility can be added to the course, enhancing learners' autonomous learning. Personalization is also fostered, since subject-specific content is delivered in different formats which learners can access at their own convenience.

Based on the framework of pedagogical podcasting devised by the IMPALA project (http://www.impala.ac.uk/outputs/Imala%20Podcasting%20Model_files/slide0002.html; http://www.impala.ac.uk/outputs/model.html), instructors involved in the CLIL project can use podcasts as pre-

paratory materials, summary lectures, key concept reviews, extensions to lectures, guidance of discipline-specific software/tool applications, and suggestions about further readings; additionally, native-speaker experts in the field can be asked to create podcasts that students can access online. Professors are encouraged to create podcasts that students can use autonomously during the preparation process. Preparatory podcast materials can feature a brief overview of the main issues that will be dealt with in lectures, explanations of key concepts, activity instructions and research questions students should try to answer while reading the material assigned and/or while listening to face-to-face lectures. With this purpose, podcasts may be largely reusable. On the other hand, if podcasts are used to provide either daily or weekly updates, they may end up being once-off materials. Also, as the podcast model elaborated by IMPALA indicates, podcasting should be done using an informal delivery style, which is more likely to engage students who may perceive podcasts as a pleasant communication between speakers and listeners: «Informal style of podcasts appealed to students. A friendly tone invites students to learn and helps to build intimacy with the speaker» (`http://www.impala.ac.uk/outputs/A%20model%20of%20podcasting.pdf`).

One of the main challenges students face in CLIL lessons is listening to lectures delivered in a foreign language, which can sometimes cause anxiety, since learners are afraid of missing important content knowledge. Therefore, podcasting can be highly valued by CLIL learners who can rely consistently on out-of-class audio/video scaffolding to be used at their own pace. The effectiveness of podcasting to scaffold content learning is supported by various studies as suggested by Nie et al. (2010, p. 106):

> More recently, Salmon and Edirisingha described a wide range of pedagogical applications of podcasting to support diverse disciplines and subjects. They highlighted the potential of podcasting to bring to learners immediacy, engagement and stimulation (Fothergill), convey more personalised and individualised content (France, Ribchester) and enable students to experience better quality contact with their lecturers (Cox et al.). Podcasts providing input, assessment guidelines and feedback were found effective in supporting students' assessed work (Sutton-Brady et al.). Although most of these case studies were based on podcasting to campus-based learners, the same approaches are transferable to distance learning.

Instructors can prepare customized online lectures after each face-to-face class or after weekly meetings. As Tynan and Barnes (in Lee et al. 2011, p. 371) suggest, «There is good evidence that students benefit from such podcast lectures (Tynan, Colbran; Lee, Chan)». Podcasts can be created using Audacity (`http://audacity.sourceforge.net`) and Soundcloud (`http://soundcloud.com`), free audio podcasting authoring tools, or MailVu (`http://mailvu.com`), a free video-

podcasting authoring tool. Screencast-O-Matic (http://www.screencast-o-matic.com), a free online screen recorder, can be used to create voice-enriched video presentations of the content discussed in class. Instructors can also use MyBrainshark (http://www.brainshark.com/mybrainsharkb), a free cloud-based slidecast authoring tool, to create user-friendly voice-enriched mobile video presentations; professors can generate slidecasts recording comments synchronized with Power Point slide sequencing. MyBrainshark audio-textual synchronizing process can be easily managed. Besides promoting personalization, multimodal representation of input further heightens learners' deeper understanding of discipline-specific content concepts.

Instructors can also decide to record discussions carried out with students in class or in other informal contexts and make them available as podcasts. The effectiveness of this kind of podcasting has emerged in studies carried out in academic settings:

> For example, to support his course in General Psychology at the University of Connecticut, Professor David B. Miller hosts weekly informal discussions with students following each week's lectures. During these discussions, students are able to seek clarification on the course material and talk about it in greater depth, as well as to actively explore and discuss issues not covered during the lecture that are of interest and relevance to the group (participation). The discussions are recorded and made available to other members of the class as a series of podcasts for individual listening at a convenient time and place (personalization). The process of creating and participating in the discussions becomes a form of student-generated content (productivity). All students in the cohort are welcome to submit questions in advance of the discussion via email; these questions, as well as those asked by students who attend in person, are answered during the discussion. (McLoughlin, Lee 2008, p. 18)

Instructors are invited to provide audio-podcasts ten minutes long at the most since this length seems to better suit users who can listen to the material on mobile devices and on the move: «The maximum recommended length of a podcast is 10 minutes (Edirisingha, Salmon and Nie)» (Nie et al. 2010, p. 108). Likewise, it is advisable to provide learners with video-podcasts that do not exceed fifteen/twenty minutes. Converting materials to the online environment may be time-consuming for instructors especially at the beginning but cost-effective in the long run.

Course-related podcasts are also likely to work effectively to help students revise before exams and therefore perform better during final examinations. The effectiveness of podcasting in content instruction, including assessment, has emerged in scholarly findings:

> Additional pedagogical benefits attributed to podcasting, as summarized by Lazzari, include significant increases in final project grades and students' receptiveness to material being presented in that format for learning and revision purposes. (Nie et al. 2010, pp. 106-107)

Instructors are also encouraged to get students involved in podcasting development. Learners can generate podcasts in the target language to present and share research findings, activity results, and experiments.

'Courses' sections are provided with a blog that instructors can use to promote learners' reflective skills and ideas are thus shared in the foreign language with peers. Moreover, using a public space can be motivating for students; a forum is made available to promote discussions about instructor- and/or student-generated content-specific issues. Among the student-generated questions posted on the Padlet notice board or in the course blog, through polling, peers can pick those most interesting to discuss in the forum while the others can be answered by the instructor in class or through podcasting. Forum discussions can greatly enhance written output in the target language. If requested by the instructor, more forums can be assigned to each course to promote group discussions. Native-speaker experts in the field can also participate in the discussions, making the interactions between participants even more authentic. These asynchronous exchanges are pivotal to enhancing intercultural awareness in a CLIL setting:

- CLIL actively seeks to promote intercultural understanding by planning and providing rich opportunities to investigate and reflect on different cultures, traditions, values and behaviour;
- this approach not only involves learning content through another language but also often involves learning content through another cultural lens. This helps learners to redefine the familiar, offering multiple perspectives and developing knowledge of and understanding about issues of shared global relevance;
- themes with relevance across the curriculum provide an appropriate arena to develop citizenship, addressing challenging ideas and fostering a human rights perspective on individual roles and responsibilities. (Coyle et al. 2009, p. 14)

Professors can also ask students to interact in the target language on Twitter with remote peers and experts in the field. Learners can thus start establishing PLNs (Professional Learning Networks), exchanging ideas, sharing resources and overall being active in ever-growing, highly-interactive, international communities of practice. Becoming active participants in Twitter subject-content networking can play a significant role in learners' motivation. Interacting in the target language with professionals, students can perceive the added value of CLIL. Noticeably, tweeting with peers

and experts in a foreign language can help learners, as Perry suggests (cited in Conole, Alevizou 2010, p. 34), to «[familiarize] with both disciplinary and professional discourses», a key objective of the CLIL approach. In keeping with the CLIL inquiry-based pedagogical approach, students can be provided with tasks requiring them to use Twitter for «Crowdsourcing of news and evidence from the ground» (Conole, Alevizou 2010, pp. 31-32). Furthermore, through social digital spaces, learners can start developing their online social identity in a foreign language. Dunlap and Lowenthal (cited in Conole, Alevizou 2010, p. 32) hold that «the use of Twitter can enhance students' perception of a sense of 'social presence', an important quality that helps promote student involvement, commitment and retention». Using Twitter would further foster learners' intercultural awareness. As Coyle et al. (2009, p. 13) suggest in relation to intercultural understanding within the 4Cs Framework:

- CLIL is particularly relevant in classrooms where learners bring diverse language and cultural experiences;
- CLIL is an appropriate vehicle for exploring the links between language and cultural identity, examining behaviours, attitudes and values;
- CLIL involves contexts and content which enrich the learners' understanding of their own culture and those of others;
- CLIL strengthens intercultural understanding and promotes global citizenship.

Instructors can use Tweetdeck (`http://www.tweetdeck.com`), a free tool that allows users to customize Twitterstreams, to monitor activities learners are required to carry out through Twitter. They can get access to more columns of Twitterstreams concurrently, such as replies to their Twitter username, students' own tweets, and the tweets featuring the class hashtags. Learners' tweet activity in the target language with peers and experts in the field can thus be monitored and evaluated.

On the basis of the practices presented so far, Web 2.0 tools are widely interlaced in online CLIL learning:

> What, then, are the implications of Web 2.0 for education? As Web 2.0 is participatory and collaborative, reflecting the way youth engage with technologies and connect with multiple social worlds, there is an increasing gap between the formalized interactions that occur in educational establishments and the modes of learning, socialization, and communication taking place in the everyday world. Siemens (2007b) states, [«]Our institutions need to change because of the increasing complexity of society and globalization. Schools and universities play a dual role: accommodating learner's method and mode of learning and

> transforming learners and preparing them to function in the world that is unfolding[»]. (McLoughlin, Lee 2008, p. 11)

In keeping with Pedagogy 2.0, instructors are invited to create Web 2.0-scaffolded tasks students can carry out in the target language out of class, either autonomously or in small groups, and present them online or during face-to-face instruction. As McLoughlin and Lee (2008, p. 15) suggest, «learning tasks: Should be authentic, personalized, experiential, and learner driven and designed, and enable the creation of content and innovative ideas by learners». Tasks requiring learners to accomplish projects, such as videos and presentations in the target language, entail students' critical engagement with the course materials and consequently the promotion of critical thinking and deep learning, which is pivotal in CLIL learning environments. To carry out these tasks, EduGlogster (`http://edu.glogster.com`), a free collaborative online platform where videos, sounds, graphics, images, text, and walls can be inserted, can be used by learners to create interactive multimedia-rich posters. Furthermore, students can publish a free online Paperli-based magazine to showcase and share their end-user generated knowledge gathered partly through crowdsourcing. Meaningful contextualized output in a foreign language is thus enhanced.

Through web-based technologies and digital spaces provided in the CLIL networked online environment, students can communicate and interact extensively in the target language with course contents, peers, instructors, and experts. Moreover, more control is shifted to learners, as advocated not only by CLIL tenets as previously mentioned, but also by Pedagogy 2.0:

> we are witnessing a re-definition of the roles of both teachers and learners, with the latter assuming more active roles as contributors of course content and ideas while also demonstrating learning outcomes through performance and production of ideas. (McLoughlin, Lee 2008, p. 18)

In the online CLIL learning environment, communication is enhanced within the following framework: «Students should be offered multiple opportunities for open, social, peer-to-peer, and multi-faceted forms of visual, verbal, and auditory communication, using multiple media types to achieve relevance, immediacy, and clarity» (McLoughlin, Lee 2008, p. 15). It is also noteworthy to mention that student-generated media-rich sticky notes can trigger discussions about issues relevant to learners themselves; emerging technologies can therefore enable students to negotiate the syllabus to a certain extent. Furthermore, through students' multimedia contributions, instructors can learn in advance what concepts they need to go over or explain again at the beginning of the next class and/or what new issues/ideas to tackle. As McLoughlin and Lee (2008, p. 15) suggest,

«Curriculum: Should not be fixed but dynamic, open to negotiation and learner input, consisting of 'bite-sized' modules, interdisciplinary in focus, and blending formal and informal learning». Networked-based interaction also aims to heighten students' engagement in their own learning process, which «[s]hould be situated, contextualized, reflective, integrated with thinking processes, iterative, dynamic, performance, and inquiry-based» (McLoughlin, Lee 2008, p. 15).

Online scaffolding provided through the CLIL website aims to foster inquiry-based content-specific pushed output in the target language, which is pivotal to enhancing effective content and foreign language acquisition. Likewise, productivity, a key asset of Pedagogy 2.0, is promoted: «Students are capable of creating and generating ideas, concepts, and knowledge, and it is arguable that the ultimate goal of learning in the knowledge age is to enable this form of creativity and productivity» (McLoughlin, Lee 2008, p. 17).

3.1.1 Autonomous academic language development

In CLIL face-to-face classes, instructors are most likely to focus on content subject learning, and to a certain extent, on subject-specific vocabulary. Academic language development in terms of guided practice may therefore be neglected to a certain extent during class hours. As a result, awareness-raising tools and customized language-focused activities with which learners can autonomously work to gradually build academic vocabulary competence are made available online.[1] With regard to academic language, proactive form-focused instruction is provided in the CLIL networked learning environment:

> Proactive form-focused instruction involves pre-planned instruction designed to enable students to notice and to use target language features that might otherwise not be used or even noticed in classroom discourse. (Lyster 2007, pp. 44-45)

The CLIL counterbalanced instructional framework advocated by Lyster (2007), which integrates content-based and form-focused instruction in both its reactive and proactive forms, is implemented to promote learners' intake in terms of academic language in particular.

It is noteworthy to mention that learners, through evaluation questionnaires, explicitly requested out-of-class academic language support to cope with CLIL lessons better. Therefore, the decision to scaffold CLIL

1 For a brief introduction to vocabulary building see Carloni (2012).

instruction by means of academic language-focused, web-based activities addresses students' needs and expectations.

To enhance self-directed learning of academic vocabulary, an avatar first introduces learners to Avril Coxhead's (2000) AWL (Academic Word List) in the 'Academic vocabulary' section; here, all AWL sub-lists can be accessed. With the support of screenshots, an avatar also introduces students to the Word and Phrase – Academic service previously mentioned. This tool enables learners to monitor and manage their academic language development autonomously while also catering to subject-specific language. To analyze the academic texts they have to study, students are also encouraged to use WordSift and Vocabulary Profiler, web-based tools previously introduced. Interactive visual thesauri, such as Visual Thesaurus and Graph Words (`http://graphwords.com`), are presented; learners are invited to use them on their own initiative.

An avatar introduces students to Concord Writer, free software previously mentioned, which they can use to search for collocations while writing in English in the box provided. Learners are also encouraged to use Just the Word, a web-based tool already introduced, to retrieve collocations in a user-friendly way as well as through Wordle-generated word clouds.

To enhance academic language development, guided practice geared to facilitate automaticity is also operationalized online. Automaticity is pivotal to language fluency:

> It is clear that a high degree of automaticity, that is fast, accurate and spontaneous effortless use of knowledge, however hard it may be to achieve, is the ultimate goal for most learners. That is both because of the impact on the quality of linguistic output and because how it frees up resources for processing message content instead of language (Segalowitz). (Meyer 2010, p. 16)

Academic language learning through guided practice is necessary to heighten intake, output, and negotiation of meaning in CLIL settings:

> Systematic language work is of paramount importance when teaching thinking. Students need to be shown how to express their thoughts in an increasingly complex manner: «Every learning involves language learning or is language learning at the same time and (that) communication, therefore is of overriding importance also in subject learning» (Vollmer). (Meyer 2010, p. 21)

To promote academic language building, gap-fill, academic language-focused exercises are created with free software (`http://www.nottingham.ac.uk/~alzsh3/acvocab/index.htm`) made available by Nottingham University. To create cloze passages, course-specific English reading materials are

firstly pasted into the box provided by AWL Gapmaker software (http://www.nottingham.ac.uk/~alzsh3/acvocab/awlgapmaker.htm). Secondly, the frequency level of academic words[2] that best suits students' competence in English is selected. Thirdly, a list of the extracted words is requested and made available at the bottom of the exercise. Finally, the submit button is pressed. In the HTML gap-fill exercise created, all the academic words featured in the pasted text are replaced by gaps. The HTML gap-fill exercise is then copied to an MS Word file and uploaded in the 'Academic vocabulary exercises' subsection of each course. In addition, AWL Highlighter software (http://www.nottingham.ac.uk/~alzsh3/acvocab/awlhighlighter.htm) is used to identify and boldface all the academic words featured in the reading materials. To this purpose, the CLIL expert pastes the targeted texts in the box provided, selects the suitable frequency level of academic words, and presses submit. The HTML-created text features all the core academic words in bold. The MS Word version of the text thereby created is made available to learners providing a key to the gap-fill exercises mentioned above. Students can thus do the exercises and self-evaluate their answers offline. Interactive gap-fill academic language-focused exercises are also created with Learnclick, and made available on the website. To this purpose, the HTML AWL Highlighter-generated texts featuring all the targeted academic words in bold are pasted into the box provided. The academic words identified with AWL Highlighter software are then manually selected; afterwards, drop-down or drag-and-drop answer choices are generated. The HTML gap-fill exercises thereby created are embedded in the website. Learners can check their answers and decide to have the solutions displayed. Course-tailored, interactive Learnclick-generated, gap-fill content, language-focused exercises are embedded in the website. The subject-specific vocabulary items retrieved through Sketch Engine and the other corpus-based tools mentioned so far are used to create the exercises. Paper-based and interactive gap-fill activities are created for each course. On demand, shorter cloze passages can be customized for use in class.

In addition, to make academic vocabulary learning more fun, links to academic language crosswords and matching activities (http://www2.elc.polyu.edu.hk/CILL/eap/2004/u6/pg137ex3&4vocab.htm; http://www2.elc.polyu.edu.hk/CILL/eap/2004/pg15verbsxword.htm; http://www2.elc.polyu.edu.hk/CILL/eap/2004/pg15nounsxword.htm) are provided in the 'Academic vocabulary: fun activities' subsection. Here, students can familiarize themselves with academic language functions through the materials provided; for example with the UEFA, Using English for Academic Purposes (http://www.uefap.com/index.htm) website. In particular, browsing the

2 «Useful academic vocabulary is contained in the Academic Word List, which is divided into 10 sublists in order of frequency, i.e. Sublist 1 has the most frequent academic words» (http://www.nottingham.ac.uk/~alzsh3/acvocab/wordlists.htm).

'Function' pages (http://www.uefap.com/speaking/spkfram.htm) of the UEFA 'Speaking' section (http://www.uefap.com/speaking/spkfram.htm), learners can access academic (1) rhetorical functions (such as describing objects, reporting and narrating, defining, giving instructions, describing processes, classifying, giving examples, comparing and contrasting, generalizing, expressing degrees of certainty, expressing reasons and explanations, expressing cause and effect, arguing and discussing, giving introductions, and drawing conclusions) and (2) language structures. Learners can also work on academic language functions through activities based on the MICASE, Michigan Corpus of Academic Spoken English (http://micase.elicorpora.info). On this website, students can engage with functions, such as clarifying and confirming (http://micase.elicorpora.info/esl/clarifying/Intro.htm) through various stages: introduction, process, exercises targeted to identify steps and strategies and a summary. Learners can study spoken academic English formulas (http://micase.elicorpora.info/esl/formulaic-expressions/Definition.htm) starting with an introduction and then moving on to written and listening activities. Students can carry out activities regarding complaining, giving instructions, giving and asking for advice, learning how to hedge claims, talking about other texts and discourses, standing aside from what one is saying, listening for conversational patterns, introducing a speaker, idioms and politeness strategies (http://micase.elicorpora.info/esl-eap-teaching-materials). Learners can familiarize themselves with vague category markers (http://micase.elicorpora.info/micase-kibbitzers/15-vague-language-in-academia) widely used in academic settings: «vagueness […] enables speakers to refer to categories of people and things in an open-ended way which calls on shared cultural and real-world knowledge to fill in the category members referred to only obliquely» (O'Keeffe et al. 2007, p. 74). Self-repair strategies can also be browsed (http://micase.elicorpora.info/micase-kibbitzers/8-announcements-of-self-repair).

3.1.2 The 'Courses' pages: a snapshot

Multimodal self-paced online learning implemented in the 'Courses' sections of the CLIL website allows students to experience an input-rich, space- and time-independent learning environment. The resources provided are in keeping with Pedagogy 2.0: «Resources: Should include multiple informal and formal sources that are media rich, interdisciplinary, and global in reach; from a network of peers, teachers, experts, and communities» (McLoughlin, Lee 2008, p. 15). Moreover, the tools available allow instructors «to 'extend the class' beyond normal instructional hours, 'address varying skill levels' and 'individualize lessons'» (Warschauer, Liaw 2011, p. 108). A graphic representation of the interrelationship between

face-to-face and online learning in the devised CLIL blended model is provided below (Fig. 128).

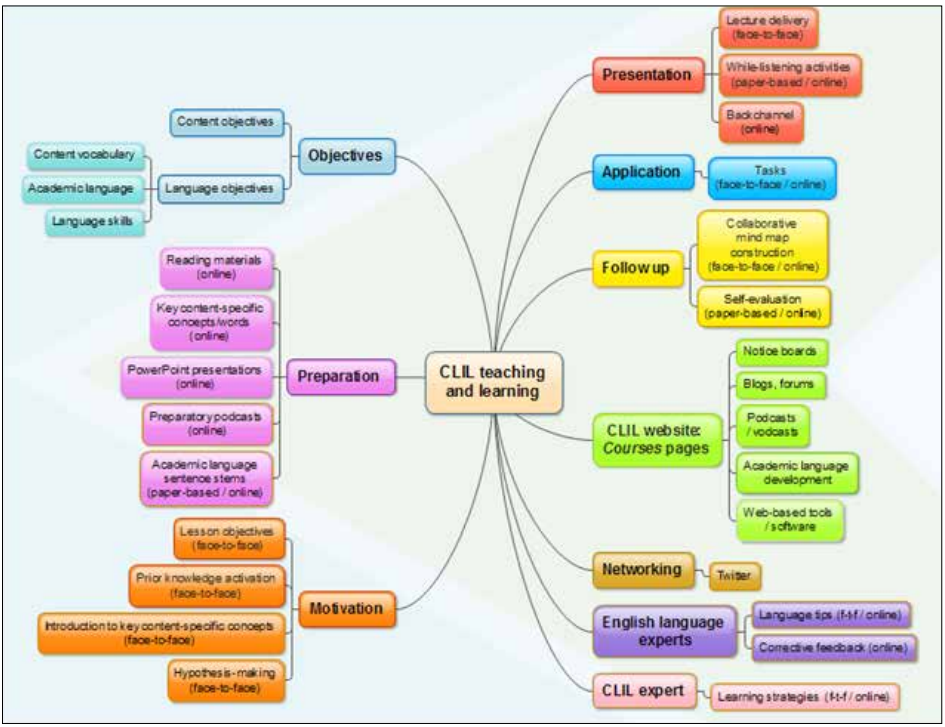


Figure 128. CLIL blended model

Within a sound pedagogical framework, the 'Courses' pages enhance students' knowledge building process in the target language through a consistent educational use of technology targeted towards promoting active learning in a variety of ways:

> If employed in conjunction with appropriate strategies, learning technologies are capable of supporting and encouraging informal conversation, dialogue, collaborative content generation and the sharing of knowledge, thereby opening up access to a vast array of representations and ideas. (McLoughlin, Lee 2010, p. 29)

Scaffolding is fundamental for online autonomous learning to be effective. To this purpose, customized resources and tools are made available to CLIL learners: «in order for self-regulated learning to come to fruition, students need not only to be able to choose and personalise what tools and content are available, but also to have access to the necessary

scaffolding to support their learning» (McLoughlin, Lee 2010, p. 28). In particular, the web-based tools deployed in the online CLIL learning environment are geared to «shift control to the learner, through promoting learner agency, autonomy and engagement in social networks that straddle multiple real and virtual learning spaces» (McLoughlin, Lee 2010, p. 28). In the networked CLIL learning environment, learners have the opportunity of experiencing greater freedom to take control of their learning process:

> [I]n instructed online environments […] [s]tudents play […] a greater role in managing the discourse, for example, they fe[el] greater freedom to suggest a new topic, follow up on someone else's ideas, or request more information. In general, they t[ake] the initiative more than they do in the normal classroom, due to the more decentralized role of the instructor. (Chun in Sieloff Magnan 2008, p. 28)

Providing learners with the possibility to manage their self-directed study through media-rich web technologies is also likely to successfully address students' ever-shifting cognitive and learning needs:

> The UK-based Committee of Inquiry into the Changing Learner Experience (CLEX, 2009) concludes that
> Web 2.0, the Social Web, has had a profound effect on behaviours, particularly those of young people whose medium and metier it is. They inhabit it with ease and it has led them to a strong sense of communities of interest linked in their own web spaces and to a disposition to share and participate. (p. 9)
> This indicates that digital-age students want an active learning experience that is social, participatory and supported by rich media. (McLoughlin, Lee 2010, p. 28)

Scaffolding learning processes through new technologies is important to foster learners' self-direction and therefore self-regulated learning in content and language instruction:

> Self regulated learning (Biggs; Zimmerman, Schunk; Simons) refers to the ability of a learner to prepare for his/her own learning, take the necessary steps to learn, manage and evaluate the learning and provide self feedback and judgment, while simultaneously maintaining a high level of motivation. A self regulated learner is able to execute learning activities that lead to knowledge creation, comprehension and higher order learning (Stubbé, Theunissen) by using processes such as monitoring, reflection, testing, questioning and self evaluation. The quest for learning to be 'student centred', self directed and self regulated has long

> been a pursuit of educators, and recent reports from various countries including the UK (see Owen, Grant, Sayers, Facer; Bryant; Minocha; CLEX), USA (see New Media Consortium; Salaway, Caruso, Nelson) and Australia (see Fitzgerald, Steele) indicate that the integration of social software into learning design can make a qualitative difference to giving students a sense of ownership and control over their own learning and career planning. (McLoughlin, Lee 2010, p. 29)

On the CLIL website, online self-directed study is especially promoted in terms of classroom preparation and academic language development. Learner autonomy is as such enhanced in CLIL teaching as suggested by Marsh et al. (2011, p. 34) who advocate «Learner autonomy and agency – Deciding on and managing one's own learning».

Online self-directed learning can also be instrumental in enhancing personalization, a core issue of CLIL teaching. Within this theoretical framework, the online tools and activities provided allow CLIL learners to experience personalized content learning in the target language, keeping with the guidelines provided below:

> Green, Facer, Rudd, Dillon and Humphreys summarize four key areas pivotal to enabling personalised learning through digital technologies. According to them, pedagogy must:
> - ensure that learners are capable of making informed educational decisions;
> - diversify and recognise different forms of skills and knowledge;
> - create diverse learning environments; and
> - include learner focused forms of feedback and assessment. (McLoughlin, Lee 2010, p. 30)

Students can set up PLEs (Personalized Learning Environments). Net learner-centric PLEs encompass «the tools, communities and services that make up the individual educational platforms learners use to direct their own learning and pursue educational goals» (`http://net.educause.edu/ir/library/pdf/ELI7049.pdf`). PLEs can be set up through a mash up of tools distributed over the cloud, which «is a state-of-the-art internet-based technology that provides access to services, storage space and resources on demand without the worry of downloading or installing anything on your computer» (Kop, Carroll 2011, p. 1). Within a self-regulated personalized learning framework, cloud computing can play a pivotal role:

> the Cloud might have the potential to give the control of learning to learners or to personalize the learning experience, by providing flexibility in adapting to the specific user's educational requirements and her or his conditions of use (Kop, Fournier,). (Kop, Carroll 2011, p. 2)

The CLIL website design envisages PLEs and learner-directed learning within a socially distributed space:

> PLEs stand in stark contrast to institutionally controlled, content-centric CMSs and VLEs as they provide learners with contextually-appropriate tool sets by enabling them to adjust, select, integrate and use various software, services and options based on their needs and circumstances. The result is, ideally, a model where learner needs, not technologies, drive the learning process (Attwell). (McLoughlin, Lee 2010, p. 31)

In a student-centered learning environment, the online scaffolding designed for CLIL courses also aims to meet digital-age learners' expectations: «In the Web 2.0 era, the need to close these gaps to achieve truly student centred learning is paramount, as learners, more so than ever before, desire and demand high degrees of autonomy, connectivity and socio-experiential learning (McLoughlin, Lee)» (McLoughlin, Lee 2010, p. 37). Moreover, through the online learning environment, digital-age learners' needs for informal learning and expectations in terms of learning models are catered to:

> Of crucial importance to attaining the longstanding goal of student centred learning is the need to acknowledge the importance of including informal modes of learning in the learning experience, to realise that learner needs and preferences cannot be addressed as static constructs during the design and development phases of instructional design, and to provide suitable scaffolds to support the learning outcomes to be attained. Educators need to prepare to accept and face the reality that learners' needs, preferences, perceptions and mental models will contribute significantly to the dynamic process that is learning design. (McLoughlin, Lee 2010, pp. 29-30)

Technology-enhanced learning has been integrated into CLIL teaching so that instructors can benefit from some of the affordances of Web 2.0:

> These indicate that these tools can result in pedagogical innovation in a number of ways. Firstly, by providing new ways of collaborative creation and exchange of learning content. Secondly, by providing new forms of communication amongst learners and teachers. Thirdly, by providing more personalised and learner-centred environments. Fourthly, these are resulting in new forms of blended learning contexts emerging. Fifthly, they are motivational in terms of providing active, discovery-based learning approaches and a sense of learner ownership. (Conole

2012, p. 52)

In particular, CLIL learning objects and digital spaces have been designed to «seek to achieve balance between self regulated and personalized learning and scaffolding support, while integrating Web 2.0 tools as well as the production, sharing and use of student-generated content» (McLoughlin, Lee 2010, p. 33). The interrelationship between scaffolding, self-regulated learning, personalization enhanced by the activities, digital communication spaces, and web-based technologies featured in the 'Courses' pages is mapped below (Tab. 3).

Table 3. Scaffolding, self-regulated learning, and personalization in the CLIL website

Context	**Scaffolding**	**Self-regulated learning**	**Personalization**
University courses taught either entirely or in part in a foreign language	Scaffolding is provided through: • resources devised by the CLIL expert • resources provided and/or created by the instructor • daily/weekly updates about course topics and assignments posted by the instructor • daily/weekly podcasts • instructor's and peers' answers and comments • mind maps of key content-specific concepts • web-based collaboratively constructed graphic organizers	Students can learn independently and in a self-directed perspective in the following ways: • by choosing the activities they consider suitable to their content and language needs • by selecting the podcasts/vodcasts they consider liable to scaffold their learning process • by choosing among a variety of tools/ activities • by providing their own contributions as well as commenting on the others' contributions • by mixing instructor-directed and learner-directed suggestions, activities, and tools	Students can: • voice their opinions • ask questions • add comments • collaboratively construct content • use peer–to-peer content sharing • use tools of their choice • decide when, where, and how to use the online environment provided (in terms of activities, communication, etc.)

Sound CLIL pedagogical theory implemented within a Pedagogy 2.0 framework has underpinned the use of Web 2.0 and other web-based tools in university-based CLIL scaffolding. As McLoughlin and Lee (2008, p. 16) hold with regard to these practices:

> They represent principles that are congruent with the philosophy of the relatively new concept of Web 2.0, but, nevertheless, they are well supported by established and accepted learning concepts and theories including motivation and self-regulation (Pintrich; Pintrich, Schunk), information processing theory (Miller), multimedia learning theory (Mayer), sociocultural learning theory (Brown, Collins, Duguid; Lave, Wenger; Vygotsky) and experiential learning theory (Kolb; Kolb, Fry).

Hockly (2011, pp. 175-176) has devised a continuum to map mobile learning, that is, learning carried out through any mobile device, which can be summarized as follows:

in the classroom ——————————————————— on the move
class sets ——————————————————————— own devices
rich content ————————————————————— discrete content
push content ————————————————————— pull content
strategic use ———————————————————————— discrete use

On the grounds of this continuum, the activities provided on the CLIL website are scheduled to take place mainly on the move. In general, learners are expected to engage with the learning objects on a voluntary basis and in a time- and space-independent way. Since no class sets are available, students need to use their own devices. In class, mobile devices are likely to mainly be used during instructors' 20-minute-long lectures and soon afterwards by means of web-based tools, such as Mentimeter, TodaysMeet and Twitter. Mobile devices can also be used in class to communicate with English native speakers. On the CLIL website, learners mostly access rich content, that is, multimedia resources as well as social spaces. On the other hand, during lectures, students primarily engage with discrete online content, such as quizzes and polls. Overall, learners engage with online pushed content, that is, materials and activities provided by instructors, even though students can be asked to pull content themselves. Mobile learning is mainly geared towards strategic use; it is aligned with the course objectives and includes a wide array of activities. However, since activities scheduled to be accomplished on the move can be carried out on a voluntary basis to a great extent, learning can end up being largely discrete.

In brief, through various web-based tools and participatory social media, an attempt has been made to increase learners' autonomy and the possibility for them to personalize their learning environments. Communicating,

sharing and networking disciplinary content in the target language has thus been promoted, while user-generated content production and social mediation has been fostered. User-generated content-production practices have been informed by sound pedagogical tenets:

> Besides, in fostering learning processes that encourage learner-generated content there is still a need for accountability and recognition of authoritative sources of information; however, the review, editing, and quality assurance of content can be done collaboratively and in partnership with learners, while simultaneously drawing on input from the wider community (i.e., «wisdom of crowds»). (McLoughlin, Lee 2008, p. 19)

The CLIL project is just at the beginning and a further shift to more user-generated content production and sharing is scheduled for the future.

3.2 Students and the CLIL Learning Center

In the 'Students' pages, learners can learn about the support services provided by the CLIL Learning Center and how to contact the CLIL and English native-speaker experts. In particular, students can email the CLIL specialist to arrange a meeting or they can post a question on the blog. Likewise, students can contact the language specialists by mail or write questions on the blog provided. Students need to feel that support is available to help them effectively cope with CLIL classes. Open online office hours to answer students' questions can be scheduled monthly on Skype.

3.3 Feedback on technical problems and suggestions

Instructors and students can report technical problems using Wishbox (`http://www.jotform.com/wishbox`), a visual-enriched tool embedded in the 'Feedback' page. The free software allows users to visually map the problems arisen with a variety of forms (arrows, circles, lines and colours), insert text-based comments and take screenshots of the webpage targeted. This feedback system allows the webmaster to identify and therefore fix the problem effectively as well as email back the senders.

Instructors and students are also encouraged to use Wishbox to ask for the implementation of further services. In particular, instructors can describe the kinds of activities they want to provide online and learners can express their specific needs to communicate, create, share and effectively and engagingly learn online.

3.4 The online learning environment: activity profiling

To effectively implement new technologies in educational practices, a learning design approach has been elaborated in recent years. As Conole (2010b, p. 10) suggests:

> The learning design research work has developed in response to a perceived gap between the potential of technologies in terms of their use to support learning and their actual use in practice (Conole; Herrington et al.; Bennett et al.).

Conole (2007, p. 84) has devised a taxonomy to map the parameters – namely the contexts, pedagogical underpinnings and components – foregrounding learning activities:

> A taxonomy has been developed that defines the components involved in a learning activity (Conole, Fill; Bailey et al.). The taxonomy attempts to consider all aspects and factors involved in developing a learning activity, from the pedagogical context in which the activity occurs through to the nature and types of tasks undertaken by the learner. At the heart of the taxonomy is the assertion that learning activities are achieved through completion of a series of tasks in order to achieve intended learning outcomes.

The context, which is the first component of the framework introduced, is defined as «the context within which the activity occurs; this includes the subject, level of difficulty, the intended learning outcomes and the environment within which the activity takes place» (Conole 2007, p. 84). In the 'Courses' pages, the context refers to the content-specific subject being taught, the degree of difficulties involved in carrying out online content activities in a foreign language, the extent to which the online tasks and the digital discussion-based spaces help learners to reach the intended learning outcomes, the online environment in which activities are carried out and the degree of scaffolding provided by online activities and digital social spaces to CLIL face-to-face instruction.

The pedagogical theories mapped by Mayes and de Freitas constitute the second component of Conole's taxonomy (cited in Conole 2010a, p. 5):

> Mayes and de Freitas grouped learning theories into three categories:
> – *Associative* (learning as activity through structured tasks)
> – *Cognitive* (learning through understanding)
> – *Situative* (learning as social practice).

The pedagogical categories featured in Mayes and de Freitas's classification include various approaches. Within this theoretical framework, in the 'Courses' sections, the structured activities focusing on academic and content-specific language development can be classified, for example, as associative since they promote, as Conole suggests, the «acquisition of skills through sequences of concepts/tasks and feedback» (http://cloudworks.ac.uk/cloud/view/2473), and conceive «learning as activity through structured tasks» (http://cloudworks.ac.uk/cloud/view/2982). On the other hand, cognitive pedagogy includes constructivism as well as reflective and dialogic learning:

> The *cognitive* perspective views learning as transformations in internal cognitive structures. Pedagogically, it is characterised by processing and transmitting information through communication, explanation, recombination, contrast, inference and problem solving. It gives rise to *constructivist* and *experiential/reflective* positions. (Conole 2010a, p. 12)

In the 'Courses' sections, this pedagogy has been implemented through activities carried out by means of asynchronous communication tools, such as notice boards, discussion boards and forums, targeted to promote dialogic co-construction of knowledge. Moreover, as Conole (2010a, p. 12) holds with regard to reflection:

> One mechanism for promoting a constructive environment that has been widely adopted in the creation of e-learning environments is cognitive scaffolding, where the activities that the learner engages with are supported by a series of guidelines to support them and help them to reflect on their actions.

In the 'Courses' sections, reflective learning has been implemented through reflective practices operationalized by means of self-directed learning objects – including DDL activities – as well as asynchronous communication tools, such as blogs and notice boards.

Situative pedagogy envisions learning as social and taking place in a community of learning:

> The *situative* perspective views learning as social participation and emphasises interpersonal relationships involving imitation, modelling, and the joint construction of knowledge. It views the ultimate objective of learning as to enable us to experience the world as meaningful. (Conole 2010a, p. 17)

In the 'Courses' sections, social learning, communication, collaboration and the creation of communities of practice are fostered, for example,

through notice boards, collaborative mind-mapping building, forums, and Twitter-based activities. In addition, instructors and native-speaker experts in the field represent access to expertise.

Connectivism envisaging knowledge creation as the product of networking collaborative learning is included in the situative category:

> Siemens has developed connectivism as an approach that emphasises the connected and networked nature within which modern learning occurs. This includes a learning ecology model that considers the elements involved in the learning process and how they can be facilitated within a networked ecology. It emphasises the networking affordances of technologies. In particular it addresses the question: How does learning change when knowledge growth is overwhelming and technology replaces many basic tasks we have previously performed? (Conole 2010a, p. 19)

Activities requiring learners to act as nodes of a networking system are featured in the 'Courses' section, especially in relation to end-user generated knowledge, even though these tasks need to be further promoted.

Thus, in terms of pedagogy, most online content-specific activities can be classified as pertaining to the cognitive and situative category: «cognitive (construction of meaning based on prior experience and context) and situative (learning in social and/or authentic settings)» (`http://cloudworks.ac.uk/cloud/view/2473`). On the other hand, as previously mentioned, the structured activities targeted to enhance academic and content-specific language development can be defined as associative.

Laurillard has classified the resources and tools used by educators to create online activities as follows:

> Tasks also involve different resources and, when appropriate, tools to access or manipulate them. Taxonomies for resources and tools were developed based on Laurillard's five principal media forms. These, together with the type of learning they support are:
> – Narrative ~attending, apprehending.
> – Interactive~investigating, exploring.
> – Communicative ~discussing, debating.
> – Adaptive~experimenting, practicing.
> – Productive~articulating, expressing.
>
> Narrative media tell or show the learner something (e.g. text, image). Interactive media respond in a limited way to what the learner does (e.g. search engines, multiple choice tests, simple models). Communicative media facilitate exchanges between people (e.g. email, discussion forum). Adaptive media are changed by what the learner does (e.g. some simulations, virtual worlds). Productive media allow the learner to produce something (e.g. word processor, spreadsheet). (Fill, Conole, Bailey 2008)

The third component of Conole's (2007, pp. 84-85) taxonomy is informed by Laurillard's work and includes six kinds of tasks learners can engage with:

> The taxonomy is similar to that developed by Laurillard and classifies task types into six areas:
> – assimilative tasks (essentially passive in nature such as reading, viewing or listening);
> – information handling (such as getting students to gather and classify resources from the Web or manipulate data in a spreadsheet);
> – adaptive (where students are engaged in using modelling or simulation software);
> – communicative (in terms of engaging in a range of dialogic activities, such as pair dialogue group-based discussions);
> – productive (where the students actively construct an artefact such as a written essay, production of a new chemical compound or creation of a sculpture);
> – experiential (such as practising skills in a particular context or undertaking an investigation).

On the grounds of Conole's (2007) learning activity taxonomy, some of the online activities provided to learners in the customized 'Courses' sections can be classified as follows:

- assimilative: reading assignments, Power Point presentations, glossaries, videos, podcasts/vodcasts, slidecasts;
- information handling: visual organizers used to classify content-specific words; visual organizers used to classify key discipline-specific concepts; researching and classifying key online resources through free social bookmarking tools;
- adaptive: videos showing how to use software;
- communicative: Padlet notice boards used by learners to ask questions and provide new ideas; discussion-based activities provided by instructors; forums used to discuss instructor- and student-generated content-specific issues; blogs used to reflect on specific topics; Twitter-based activities;
- productive: collaborative construction of key content-specific mind maps; production of podcasts and slidecasts; creation of artifacts as end-user generated knowledge;
- experiential: enquiry-based tasks and simulations.

In keeping with the learning design approach, a graphic representation of the educational tenets underpinning the CLIL online scaffolding can be provided: «Inherent in the approach was the assumption that educational practice can be represented in a design description, i.e. that underlying design ideas and principles can be captured in an explicit representation» (Conole

2010b, p. 10). Within this theoretical framework, the classification of some of the CLIL activities carried out above through Conole's (2007) taxonomy is graphically represented below by means of the pedagogy profiler devised by the scholar (Fig. 129, `http://cloudworks.ac.uk/cloud/view/2459`).

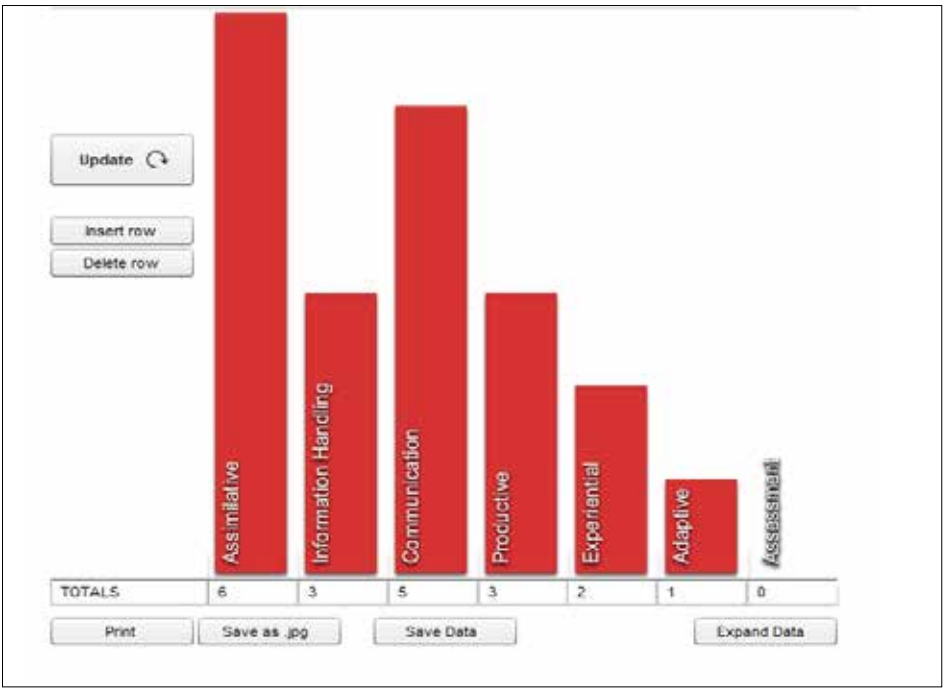


Figure 129. The Pedagogy Profile of some of the 'Courses' section website

A course map view is useful to visually track course architectures:

> The course map view provides an 'at a glance' view of a course defined around the four main categories that a course is made up of:
> - Guidance and Support
> - Content and Activities
> - Communication and Collaboration
> - Reflection and Demonstration. (Conole 2010a, p. 22)

In keeping with Conole's (2010a) guidelines in terms of visual representation of a course structure, the course map view of the online CLIL scaffolding provided can be seen below:

Table 4. Online CLIL course map view

Guidance and Support		Content and Experience	
Tools and Resources	*Roles and Relationships*	*Tools and Resources*	*Roles and Relationships*
- Guided learning pathways - *Lino* notice boards featuring instructors' posts (such as course information, assignments, advice, etc.) - Avatars providing instructions - Activity guidelines - Social bookmarking tools - CLIL methodological support (CLIL expert) - Foreign language support service (English language specialists)	- Students are expected to already be able to use undergraduate level study skills - Students are expected to have basic ICT skills - Instructors encourage students to be active learners and consistently use the target language	- Module materials including course reading materials (such as articles, reports, readings, and related activities), PowerPoint presentations, videos and related activities, podcasts, screencasts, mind maps, graphic organizers, word clouds, glossaries - *Padlet* notice boards for students' posts (such as questions, ideas) - Learner-generated contents - Academic language development (AWL, subject-specific academic language activities, academic language fun activities) - Content-specific language development - DDL activities - Web-based tools (such as *Vocabulary Profiler*, *Word and Phrase – Academic*, *WordSift*, *Just the Word*, etc.) - Discussion forum content - Blog content	- Students study assigned materials and carry out activities prior to class (preparation phase - self-directed study) - Students post comments and/or questions on content-specific issues either before or after class - Students carry out academic language exercises in a self-directed mode - Students carry out activities and tasks assigned by instructors after class - Students carry out various activities (such as listening/viewing, reading, discussing, creating end-user generated knowledge)

Reflection and Demonstration		Communication and Collaboration	
Tools and Resources	*Roles and Relationships*	*Tools and Resources*	*Roles and Relationships*
- *Padlet* notice boards - Recording reflection (podcasting and avatars) - Reflective blogs - Forums	- Use of *Padlet* notice boards is encouraged to reflect on content-specific issues - Use of podcasts and avatars is encouraged to reflect on content-specific issues - Use of a reflective blog is encouraged to promote reflection about student-generated content-related topics - Use of a forum is encouraged to promote reflection about instructor- and/or student-generated content-related topics	- Asynchronous online Padlet notice and discussion boards - Asynchronous forums - Asynchronous use of Twitter - Asynchronous use of TodaysMeet - Live online discussions via *Twitter* or *TodaysMeet* - Collaborative mind-mapping construction with Popplet - Podcasts - Peer communication	- Emphasis on instructor-student communication - Emphasis on peer communication and collaboration - Emphasis on learning from one another's textual and multimodal reflections and ideas - Emphasis on being an active member of a community of practice - Emphasis on interacting with experts in the field - Emphasis on the consistent use of the target language to co-construct meaning and artifacts

Some tools and activities –including the kind of enhanced learning featured in the 'Courses' section– have also been mapped using the pedagogy framework devised by Conole (2008):

> We argued that any particular instance of learning lies somewhere along a combination of [...] three dimensions [...]. The first is the dimension of learning on one's own, individually through to learning socially. The second dimension is learning through information versus learning through experience. The third dimension is learning passively versus actively.

For example, on the grounds of the framework featured above, reading information provided by instructors on the Lino notice boards is individual, passive and information-based, while listening to avatars' instructions is individual, passive and information-based. Reading and/or listening to reading assignments/podcasts is individual, active and experience-based. Posting questions related to the reading material assigned on the Padlet notice boards is individual, active and experience-based. Posting answers and/or ideas on the Padlet boards is social, active and information-based. Posting ideas in the course blog is individual, active and information-based. Discussing ideas in the course forum is social, active and experience-based. Tweeting is social, active and experience-based. Co-constructing concept

maps is active, social and in between information- and experience-based. Doing academic language-focused cloze passages and multiple choice/matching activities is individual, active and experience-based; and asking the CLIL and native-speaker experts for support using the blog or email is social, active and experience-based.

3.5 The CLIL website design model

While designing the CLIL website, Luckin's (2010, p. 28) educational technology Zone of Collaboration model has been largely implemented. Courseware was thus:

> informed [by Luckin's] […] interpretation of the ZPD [Vygotsky's Zone of Proximal Development], which explores the relationship between the identification of a learner's collaborative capability and the specification of the assistance that needs to be offered to the learner in order for them to succeed at a particular task. (Luckin 2010, p. 28)

The relationship between learners' needs, materials, activities, web-based technologies and instructor's and CLIL expert's roles, were analyzed. The findings were used to create effective learner-tailored scaffolding working as a More Able Partner:

> Effective scaffolding […] involves simplification of the learner's role and interactions in which learners and their More Able Partners (MAPs) work together to achieve success, but the contribution from each vary according to the [learner's] level of ability (Wood). (Luckin 2010, p. 26)

In this way, technology can act as a More Able Partner: «Scaffolding through technology has been used to support both collaborative learning and individual progress where the technology plays the role of the learner's More Able Partners (MAPs)» (Luckin 2010, p. 40). In keeping with the Zone of Collaboration model, the materials in the 'Resources', 'Courses' and 'Students' sections make up the Zone of Available Assistance as defined by Luckin (2010, p. 28):

> I introduce two additional constructs called:
> 1. The Zone of Available Assistance (ZAA); and
> 2. The Zone of Proximal Adjustment (ZPA).
>
> The ZAA describes the variety of resources within a learner's world that could provide different qualities and quantities of assistance and that may be available to a learner at a particular point in time.

Within the ZPD framework, a key role is played by the people or technologies that provide the scaffolding. In the online CLIL learning environment, the CLIL expert and course instructors act as MAPs, which first identify learners' needs and then provide students with suitable online web-based customized learning materials, activities, services as well as learner-instructor and peer-to-peer interaction: «Part of the aim of person, people, or technology acting in the role of the learners' MAP […], is to identify and possibly introduce a variety of types of assistance from the available resources» (Luckin 2010, p. 28). Furthermore, MAPs help learners to identify their ZPA, which «represents a subset of resources from the ZAA that are appropriate for a learner's needs» (Luckin 2010, p. 28). For MAPs to provide suitable scaffolding, learners need to voice their needs. To this purpose, students can communicate their content-related needs through media-rich sticky notes on Padlet notice boards or by sending voicemail messages. Further, learners can contact the CLIL and language experts, also acting as MAPs, through the 'Students' pages; the experts will be able to help them get personalized scaffolding. The ZPA is thus created through the interaction of learners with MAPs operationalized via web-based tools:

> One of the goals of the learner is to express their current understanding and learning needs through their interactions with the elements of the ZAA, including interactions with the More Able Partner. The ZPA is constructed through a negotiation between a learner and More Able Partners. An impoverished ZAA will limit the possibilities for the construction of the ZPA. However, even if the ZAA offers a rich and versatile range of resources, the success with which a ZPA meets the needs of the learner will depend upon the negotiation between learners and More Able Partners. (Luckin 2010, p. 28)

The Ecology of Resources Model, devised by Luckin (2010), envisions learners interacting with three different sets of resources. The 'Knowledge and Skills' category, featuring content-specific concepts and skills, represents the first set (Luckin 2010). The second set is the 'Tool and People' category which «includes books, pens and paper, technology and other people who know more about the knowledge or skills to be learnt than the learner does» (Luckin 2010, p. 91). The third set is the 'Environment': «This category includes the location and surrounding environment with which the learner interacts: for example, a [university] classroom […] or a place of work» (Luckin 2010, p. 91). An element affecting how learners interface with the resources is the filter: «for example, the subject matter to be learnt is often filtered through some kind of organization, such as a curriculum, that has been the subject of a process of validation by other members of the learner's society» (Luckin 2010, p. 93). Also, tools and people can be filtered to a certain extent: «for example, a teacher taking

a French conversation evening class is available also during that time, or perhaps at some other times via email» (Luckin 2010, p. 93). Moreover, «a learner's access to the Environment is mediated by that Environment's organization [...] [such as] timetables and regulations» (Luckin 2010, p. 93). The following grid[3] shows how the three categories of resources are organized in the CLIL courses in general, and in the website in particular:

Table 5. Organization of resources in the CLIL classes and website

Resource element	Filter element
Knowledge Content-specific concepts and values	**Filter: Curriculum** Course requirements - course contents - course objectives - course evaluation and assessment
Tools and people *Human resources:* instructors, CLIL expert, English language experts *Physical resources:* reading materials, materials used in class, computers *Digital resources: CLIL teaching* website	**Filter: Administration** *Types of administration of human resources include:* the type of contracts of employment (working hours and contact time) *Types of administration of physical resources include:* classroom *Types of administration of digital resources include:* the web-based tools embedded in the website, the learning objects made available on the website, the contents selected by the instructor and CLIL expert
Environment Classroom *CLIL teaching* website	**Filter: Organization** Classroom - timetable - lecture styles - seating arrangement Website - the materials and links made available - availability of computers on campus - availability of Wi-Fi on campus - added value perceived by instructors

3 The grid format has been elaborated by Luckin (2010, p. 127).

4 Conclusion

In the CLIL scenario set up to enhance CLIL courses in higher education, technology has not been conceived as an add-on, but has been fully integrated and interlaced with face-to-face CLIL teaching/learning experiences. Thus, within a CLIL pedagogic framework, face-to-face and online consultation services have been provided to instructors to effectively support them in lesson planning, material design and lesson delivery.

Scaffolding of learners' content knowledge acquisition and foreign language development has been extensively and consistently provided through the CLIL website. The online environment has been designed to allow for a further move from teacher-directed to more learner-directed learning, a core tenet of the CLIL approach. Online scaffolding has also been designed to allow learners to exponentially increase their access to input-rich personalized learning experiences in the target language. Increasing learners' engagement and exposure to the foreign language is one of the main assets of CLIL pedagogy. In this environment, learners can master multiple modes of representation. Learning objects and social digital spaces have been particularly geared to scaffold students' self-directed learning as well as communication and cooperation; pushed output in the target language has as such been enhanced. Overall, customized technology-enhanced learning has been targeted to help digital-age learners manage their CLIL experience. Learners have been enabled to experience educational benefits from innovative practices performed through pedagogically grounded scaffolding carried out in a mash up of cloud-based learning environments. Overall, the CLIL website design has tried to enhance content and foreign language acquisition through learners' proactive and critical engagement with (1) discipline-specific content, (2) collaboration between peers, (3) more time on task, (4) learners' autonomous learning and (5) creation of a community of practice.

In CLIL teaching, instructors need to promote active, strategic and self-directed learning: «The challenge for educators is to enable self-direction, knowledge building and learner control by providing options and choice while still supplying the necessary structure and scaffolding» (McLoughlin, Lee 2008, p. 17). The CLIL blended model implemented has

therefore been designed to support instructors and learners accomplish this goal as supported by research:

> Recent research attests to a growing appreciation of the importance of the learner's self-direction and control over the whole learning process (Fazey; Narciss, Proske, Koerndle). Evidence suggests that we can improve learning effectiveness by giving the learner control over, and responsibility for their own learning (Dron; Nesbit, Winne). This is the foundation for such approaches as problem-based and inquiry-based learning (Desharnais, Limson; Edelson, Gordin, Pea) and is central to the grand vision of Pedagogy 2.0, where learners have the freedom to decide how to engage in personally meaningful learning. (McLoughlin, Lee 2008, p. 17)

References

Ala-Mutka, K. (2012). *Mapping digital competence: towards a conceptual understanding*. Seville: Institute for Prospective Technological Studies.

Baker, C. (2006). *Foundations of bilingual education and bilingualism*. Clevedon, (England); Buffalo (NY): Multilingual Matters.

Ball, P.; Lindsay, D. (2013). «Language demands and support for English-medium instruction in tertiary education: Learning from a specific context». In: Doiz, A.; Lasagabaster, D.; Sierra, J.M. (eds.), *English-medium instruction at universities: Global Challenges*. Bristol; Buffalo; Toronto: Multilingual Matters, pp. 44-61.

Barnhardt, S. (1997). «Self-efficacy and second language learning». *The Learning Strategies Instructional Resource*, vol. 1, 5, p. 26. NCLRC (National Capital Language Resource Center) Georgetown University; George Washington University: The Center for Applied Linguistics. Disponibile all'indirizzo: `http://nclrc.org/teaching_materials/assessment/assessment_tools/ls_instructional_resource.pdf` (2012-04-17).

Bartik, K.; Maerten, C.; Tudor, I.; Valcke, J. (2009). *A discussion brief of content and language integrated learning (CLIL) at the faculty of applied sciences*. Disponibile all'indirizzo: `http://difusion.academiewb.be/vufind/Record/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/84137/Details` (2012-02-13).

Baoutina, A.; Alexander, I.E.; Rasko, J.E.J.; Emslie, K.R. (2008). «Developing strategies for detection of gene doping». *The Journal of Gene Medicine*, 10 (1), pp. 3-20.

Biber, D.; Conrad, S.; Leech, G. (2002). *Student grammar of spoken and written English*. Harlow: Longman.

Biber, D. (2006). *University Language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Carloni, G. (2012). «Online CLIL scaffolding at university level: building learners' academic language and content-specific vocabulary across disciplines through online learning». In: Bradley, L.; Thouësny, S. (eds.), *CALL: Using, learning, knowing, EUROCALL conference*, Gothenburg, Sweden, 22-25 August 2012, Proceedings, vol. 1, pp. 37-42. Dublin: Research-publishing.net.

Carloni, G. (2013a). «Content and Language Integrated Learning in Higher Education: A Technology-Enhanced Model». In: Aitken, J.E. (ed.), *Cases on communication technology for second language acquisition and learning*.

Hershey (PA): IGI Global, pp. 484-508.
Carloni, G. (2013b). «English corpus linguistics: a case study at university level». In: *Proceedings of 6Th international conference of education, research and innovation* (ICERI2013), Seville (Spain), pp. 326-332.
Carloni, G. (2013c). «English corpus linguistics and Content and Language Integrated Learning in higher education». In: *Proceedings of* INNODOCT 2013*, the international conference on innovation, documentation and teaching technologies*, 6-7 May 2013, Polytechnic University of Valencia, Valencia (Spain).
Carloni, G. (2013d). «Integrating English corpus linguistics and online learning environments at university level». *International Journal of Information Technology & Computer Science* (IJITCS), vol. 12 (1), pp. 68-76.
Carloni, G. (2014). «Content and Language Integrated Learning: a Blended Model in Higher Education». *The International Journal of Technology, Knowledge and Society*, vol. 9 (4), pp. 61-71.
Chamot, A.; Barnhardt, S.; El-Dinary, P.B.; Robbins, J. (1999). *The learning strategies handbook*. New York: Longman.
Cheng, W. (2009). «Income/interest/net: using internal criteria to determine aboutness of a text». In: Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam: John Benjamins, pp. 157-177.
Chiappe, R.; Segovia Cifuentes, Y.; Rincon Rodriguez, H.Y. (2007). «Toward an instructional design model based on learning objects». *Educational Technology Research and Development*, 55 (6), pp. 671-681.
Chun, D.M. (2008). «Computer-mediated discourse in instructed environments». In: Sieloff Magnan, S. (ed.), *Mediating discourse online*. Amsterdam: John Benjamins, pp. 15-45.
Commission of the European Communities (2005). *A new framework strategy for multilingualism*, COM (2005), 596 final. Disponibile all'indirizzo: http://ec.europa.eu/education/policies/lang/doc/com596_en.pdf (2012-01-02).
Commission of the European Communities (2010). *Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of regions. A Digital Agenda for Europe*. Disponibile all'indirizzo: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245&from=EN (2014-09-22).
Conole, G. (2007) «Describing learning activities: tools and resources to guide practice». In: Beetham, H.; Sharpe, R. (eds.), *Rethinking pedagogy for a digital age*. London: Routledge, pp. 81-91.
Conole, G. (2008). «New schemas for mapping pedagogies and technologies». *Ariadne Magazine* 56, July. Disponibile all'indirizzo: http://www.ariadne.ac.uk/issue56/conole/ (2012-03-21).
Conole, G. (2010a). *Review of pedagogical models and their use in e-learning*. Disponibile all'indirizzo: http://www.slideshare.net/grainne/pedagogical-models-and-their-use-in-elearning-20100304 (2012-01-10).

Conole, G. (2010b). «Learning design – making practice explicit». *ConnectEd Design Conference*, 28 June – 2 July 2010, Sydney, Australia. Disponibile all'indirizzo: http://oro.open.ac.uk/22890/2/Conole.pdf (2012-02-09).

Conole, G. (2012). *Designing for learning in an open world*. Heidelberg: Springer.

Conole, G. (s.d.) *Pedagogy profiler*. Disponibile all'indirizzo: http://cloudworks.ac.uk/cloud/view/2459 (2012-02-18).

Conole, G. (s.d.) *Design tips: what is a learning activity?* Disponibile all'indirizzo: http://cloudworks.ac.uk/cloud/view/2473 (2012-02-18).

Conole, G. (s.d.) *Paper: review of pedagogical models and frameworks*. Disponibile all'indirizzo: http://cloudworks.ac.uk/cloud/view/2982 (2012-18-02).

Conole, G.; Alevizou, P. (2010). *A literature review of the use of web 2.0 tools in higher education*. Disponibile all'indirizzo: http://www.jisctechdis.ac.uk/assets/EvidenceNet/Conole_Alevizou_2010.pdf (2012-04-27).

Coonan, C. (2008). «The foreign language in the CLIL lesson. Problems and implications». In: Coonan, C. (a cura di), *CLIL e l'apprendimento delle lingue: Le sfide del nuovo ambiente di apprendimento*. Venezia: Cafoscarina, pp. 13-35.

Coxhead, A. (2000). «A new academic word list». *TESOL Quarterly*, 34, pp. 213-238.

Coyle, D.; Holmes, B.; King, L. (2009). *Towards an integrated curriculum – CLIL national statement and guidelines*. Disponibile all'indirizzo: http://www.rachelhawkes.com/PandT/CLIL/CLILnationalstatementandguidelines.pdf (2012-01-25).

Coyle, D.; Hood, P.; Marsh, D. (2010). *CLIL content and language integrated learning*. Cambridge: Cambridge University Press.

Crawford Camiciottoli, B. (2010). «A corpus-informed approach to teaching lecture comprehension skills in english for business studies». In: Campoy-Cubillo M.C.; Bellés-Fortunño, B.; Gea-Valor, L. (eds.), *Corpus-based Approaches to English Language Teaching*. London: Continuum, pp. 95-106.

Cummins, J. (1979). «Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters». *Working Papers on Bilingualism*, 19, pp. 121-129.

Delia, T. (2012). «Cell Structure and Function: Part 2: Cell Structures and Endomembrane System». Disponibile all'indirizzo: https://www.youtube.com/watch?v=V33p4pGa79s (2014-07-08).

Ellis, R. (1999). *The study of second language acquisition*. Cambridge: Cambridge University Press.

European Objectives (2011). *Progress towards the common European objectives in education and training indicators and benchmarks 2010/2011. Commission of the European Communities*. Disponibile all'indirizzo:

http://www.eqavet.eu/Libraries/Policy_Documents_2011/European_Commission_s_report_on_progress_towards_common_European_objectives_in_education_and_training.sflb.ashx?download=true (2014-09-22).

European Union (2009). «Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ('ET 2020')». *Official Journal of the European Union*, 119/2, 2009/C. Disponibile all'indirizzo: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:119:0002:0010:EN:PDF (2013-06-09).

Fill, K.; Conole, G.; Baily, C. (2008). «A toolkit to guide the design of effective learning activities». In: Rees, P.;Mackay, L; Durham, H.; Martin, D. (eds.), *E-learning for geographers: Online materials, resources, and repositories*. Hershey (PA): IGI Global, pp. 156-169.

Flowerdew, L. (2012). *Corpora and language education*. Basingstoke: Palgrave McMillan.

Graesser, A.; D'Mello, S.K. (2011). «Theoretical Perspectives on Affect and Deep Learning». In: Calvo, R.A.; D'Mello, S.K. (eds.), *New Perspectives on Affect and Learning Technologies*. New York: Springer, pp. 11-21.

Granath, S. (2009). «Who benefits from learning how to use corpora?». In: Aijmer, K (ed.), *Corpora and language teaching*. Amsterdam: John Benjamins, pp. 47-65.

Greaves, C. (2009). *ConcGram 1.0. a phraseological search engine*. Amsterdam: John Benjamins.

Hales, K.G. (2010). «Mitochondrial Fusion and Division». *Nature Education*, 3 (9). Disponibile all'indirizzo: http://www.nature.com/scitable/topicpage/mitochondrial-fusion-and-division-14264007 (2014-07-08).

Halliday, M.; Teubert, W.; Yallop, C.; Čermáková, A. (2004). *Lexicology and corpus linguistics. An introduction*. London: Continuum.

Hancioglu, N.; Eldridge J. (2007). «Texts and frequency lists: some implications for practicing teachers». *ELT Journal*, 61 (4), pp. 330-340.

Harasim, L. (2012). *Learning theory and online technology*. Abingdon, Oxon: Routledge.

Hoffman, S.; Evert, S.; Smith, N.; Lee, D.; Berglund Prytz, Y. (2008). *Corpus linguistics with BNCweb – a practical guide*. Frankfurt: Peter Lang.

Hockly, N. (2011). «Teaching the mobile generation». *IATEFL 2010: Harrogate Conference Selections*, Canterbury: IATEFL, pp. 175-176.

Hoey, M. (2005). *Lexical priming. A new theory of words and language*. Abingdon: Routledge.

Johansson, S. (2009). «Some thoughts on corpora and second-language acquisition». In: Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam: John Benjamins, pp. 33-44.

Johnson, L.; Levine, A.; Smith, R.; Stone, S. (2010). *The 2010 Horizon Report*. Austin (TX): The New Media Consortium. Disponibile all'indirizzo: http://www.nmc.org/pdf/2010-Horizon-Report.pdf (2014-07-13).

Kalyuga, S. (2011). «Cognitive load in adaptive multimedia learning». In: Calvo, R.A.; D'Mello, S.K. (eds.), *New perspectives on affect and learning technologies*. New York: Springer, pp. 203-215.
Kilgarriff, A.; Rychly, P.; Smrz, P.; Tugwell, D. (2004). «The Sketch Engine». Disponibile all'indirizzo: ftp://ftp.itri.bton.ac.uk/reports/ITRI-04-08.pdf (2013-06-13).
Kop, R.; Carroll, F. (2011). «Cloud computing and creativity: learning on a massive open online course». *European Journal of Open, Distance and E-learning*. Disponibile all'indirizzo: http://www.eurodl.org/?article=457 (2012-03-03).
Kost, C.R., (2008). «Use of communication strategies in a synchronous CMC environment». In: Sieloff Magnan, S. (ed.), *Mediating discourse online*. Amsterdam: John Benjamins, pp. 153-189.
Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.
Larsen-Freeman, D.; Long, M. H. (1991). *An introduction to second language acquisition research*. New York: Addison Wesley Longman.
Lightbown, P.; Spada, N. (1997). *How languages are learned*. Oxford: Oxford University Press.
Long, M. (1996). «The role of the linguistic environment in second language acquisition». In: Ritchie, W.; Bhatia, T. (eds.), *Handbook of language acquisition*, *vol. 2: Second Language Acquisition*. New York: Academic Press, pp. 413-468.
Luckin, R. (2010). *Re-designing learning contexts*. London: Routledge.
Lyster, R. (2007). *Learning and teaching languages through content. A counterbalanced approach*. Amsterdam: John Benjamins.
Marsh, D.; Mehisto, P.; Wolff, D.; Frigols Martín, M.J. (2011). *European framework for CLIL teacher education. A framework for the professional development of CLIL teachers.* Disponibile all'indirizzo: http://www.ecml.at/tabid/277/PublicationID/62/Default.aspx (2013-09-10).
Marzano, R. J. (2004). *Building background knowledge for academic achievement*. Alexandria (VA): Association for Supervision and Curriculum Development.
Matthews, J.R.; Bowen, J.M.; Matthews, R.W. (2000). *Successful scientific writing*, Cambridge: Cambridge University Press.
McCarthy, M. (2004). *Touchstone. From corpus to course book*. Cambridge: Cambridge University Press.
McCarthy, M. (2006). *Explorations in corpus linguistics*. Cambridge: Cambridge University Press. Disponibile all'indirizzo: http://www.cambridge.org/other_files/downloads/esl/booklets/McCarthy-Corpus-Linguistics.pdf (2012-03-29).
McCarthy, M.; O'Keefe, A.; Walsh, S. (2010). *Vocabulary matrix*. Andover: Heinle Cengage Learning.
McDonough, K.; Trofimovich, P. (2008). *Using priming methods in second*

language research. New York: Routledge.
McEnery, T.; Xiao, R.; Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
McLoughlin, C.; Lee, M. J. W. (2008). «The three P's of pedagogy for the networked society: personalization, participation and productivity». *International Journal of Teaching and Learning in Higher Education*, 20 (1), pp. 10-27. Disponibile all'indirizzo: http://www.isetl.org/ijtlhe (2012-02-28).
McLoughlin, C.; Lee, M.J.W. (2010). «Personalised and self regulated learning in the web 2.0 era: international exemplars of innovative pedagogy using social software». *Australasian Journal of Educational Technology*, 26/1, pp. 28-43.
Meara, P. (s.d.) «Plausible Non-Words (PNWs) as devised by Paul Meara & colleagues for use in Yes-No Tests». Disponibile all'indirizzo: http://www.lextutor.ca/freq/lists_download/pnwords.html (2012-02-28).
Meyer, O. (2010). «Towards quality-CLIL: successful planning and teaching strategies». *Pulso: revista de educación*, 33, pp. 11-29. Disponibile all'indirizzo: http://dialnet.unirioja.es/servlet/articulo?codigo=3311569 (2012-04-02).
Nie, M.; Armellini, A.; Harrington, S.; Barklamb, K.; Randall, R. (2010). «The role of podcasting in effective curriculum renewal». *ALT-J, Research in Learning Technology*, 18/2, pp. 105-118.
OECD (2007). *Giving knowledge for free – The emergence of open educational resources: OECD*. Disponibile all'indirizzo: http://www.oecd.org/dataoecd/35/7/38654317.pdf (2014-09-24).
O'Keeffe, A.; McCarthy, M.; Carter, R. A. (2007). *From corpus to classroom*. Cambridge: Cambridge University Press.
Oksefjell Ebeling, S. (2009). «Oslo interactive English: corpus-driven exercises on the web». In: Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam: John Benjamins, pp. 67-82.
Prabhu, N. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
Richards, J.C.; Rodgers, T. (2001). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.
Römer, U. (2010). «Using general and specialized corpora in English language teaching: past, present and future». In: Campoy-Cubillo, M.C.; Bellés-Fortunño, B.; Gea-Valor, L. (eds.), *Corpus-based approaches to English language teaching*. London: Continuum, pp. 18-35.
Saville-Troike, M. (2006). *Introducing second language acquisition*. Cambridge: Cambridge University Press.
Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
Short, D.J.; Vogt, M.; Echevarria, J. (2011). *The SIOP model for teaching history-social studies to English learners*. Boston: Pearson.
Simpson-Vlach, R.; Ellis, N. (2009). «An academic formulas list: new

methods in phraseology research». Prepublication draft. Disponibile all'indirizzo: http://www-personal.umich.edu/~ncellis/NickEllis/Publications_files/AFL_paper_AppLinxPrepub.pdf (2013-06-09).
Simpson-Vlach, R.; Ellis, N. (2010). «An academic formulas list: new methods in phraseology research». *Applied Linguistic*s, 31/4, pp. 487-512.
Teubert, W. (2004). «Language and corpus linguistics». In: Halliday, M.; Teubert, W.; Yallop, C.; Čermáková, A. (eds.), *Lexicology and corpus linguistics: An introduction*. London: Continuum, pp. 73-111.
Tynan, B.; Barnes, C. (2011). «Web 2.0 and professional development of academic staff». In: Lee, M.J.W.; McLoughlin, C. (eds.), *Web 2.0-based e-learning: applying social informatics for tertiary teaching*. New York: Information Science Reference, pp. 365-379.
VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood; NJ: Ablex Publishing Corporation.
Vargas Parada, L. (2010). «The Apicoplast: An Organelle with a Green Past». *Nature Education*, 3 (9). Disponibile all'indirizzo: http://www.nature.com/scitable/topicpage/the-apicoplast-an-organelle-with-a-green-14231555 (2014-03-03).
Warschauer, M.; Liaw, M. (2011). «Emerging technologies for autonomous language learning». *Studies in Self-Access Learning Journal*, 2/3, pp. 107-118. Disponibile all'indirizzo: https://docs.google.com/file/d/0B1kU_KPg-p3-YWNiMjdjYWMtODU0OS00NTY3LThmODktNDdhYjAwN2JiYmY1/edit?hl=en_GB&pli=1 (2012-03-09).
Williams, M.; Burden, R. (1997). *Psychology for language teachers*. Cambridge: Cambridge University Press.
Wolff, D. (2011). «CLIL and learner autonomy: relating two educational concepts». *Education et Sociétés Plurilingue*s, 30/June, pp. 69-80. Disponibile all'indirizzo: http://www.cebip.com/datapage.asp?l=1&id=43 (2013-06-09).

Websites

Academic Vocabulary
http://www.nottingham.ac.uk/~alzsh3/acvocab/index.htm.
Academic Word List
http://www.victoria.ac.nz/lals/resources/academicwordlist.
Academic language crosswords and matching activities
http://www2.elc.polyu.edu.hk/CILL/eap/2004/u6/pg137ex3&4vocab.htm.
http://www2.elc.polyu.edu.hk/CILL/eap/2004/pg15verbsxword.htm.
http://www2.elc.polyu.edu.hk/CILL/eap/2004/pg15nounsxword.htm.
Audacity
http://audacity.sourceforge.net.
AWL Gapmaker
http://www.nottingham.ac.uk/~alzsh3/acvocab/awlgapmaker.htm.
AWL Highlighter
http://www.nottingham.ac.uk/~alzsh3/acvocab/awlhighlighter.htm.
Bagtheweb
http://bagtheweb.com.
myBrainshark
http://www.brainshark.com/mybrainshark.
Cacoo
http://www.cacoo.com.
CLIL Teaching website
http://clilteaching.weebly.com.
Compleat Lexical Tutor
http://www.lextutor.ca.
Compleat Reaction Time
http://www.lextutor.ca/rt.
Concord Writer
http://conc.lextutor.ca/concord_writer/index.pl?lingo=English.
Conole's cloudworks
http://cloudworks.ac.uk/cloud/view/2473.
http://cloudworks.ac.uk/cloud/view/2982.
http://cloudworks.ac.uk/cloud/view/2459.
Dipity
http://www.dipity.com.
Draftin
https://draftin.com.
EAP activities
http://www2.elc.polyu.edu.hk/CILL/eap.
EduGlogster
http://edu.glogster.com.
Exploratree
http://www.exploratree.org.uk.

Familizer
http://www.lextutor.ca/familizer.
Flisti
http://flisti.com.
Graph Words
http://graphwords.com.
IMPALA podcast development model
http://www.impala.ac.uk/outputs/model.html.
http://www.atimod.com/podcasting/PDModel.html.
Just the Word
http://www.just-the-word.com.
Learnclick
http://www.learnclick.com.
Lino
http://en.linoit.com.
MailVu
http://mailvu.com.
Mentimeter
http://mentimeter.com.
MICASE
http://micase.elicorpora.info.
Model of Learning Objectives
http://www.celt.iastate.edu/pdfs-docs/teaching/RevisedBloomsHandout.pdf.
Natural Reader
http://www.naturalreaders.com/index.htm.
Open Yale Courses
http://oyc.yale.edu.
Genomic conflict (lecture)
http://oyc.yale.edu/ecology-and-evolutionary-biology/eeb-122/lecture-10.
Genomic conflict (transcript)
http://oyc.yale.edu/transcript/185/eeb-122.
Padlet
http://www.padlet.com.
Pearltrees
http://www.pearltrees.com.
Popplet
http://www.popplet.com.
ProfessorWord
http://www.professorword.com.
ProProfs
http://www.proprofs.com.

Range
http://www.victoria.ac.nz/lals/resources/range.aspx.
Screencat-o-Matic
http://www.screencast-o-matic.com.
Sketch Engine
http://sketchengine.co.uk.
SoundCloud
http://soundcloud.com.
SoundGecko
http://soundgecko.com.
SpeakPipe
http://www.speakpipe.com.
SurveyMonkey
http://www.surveymonkey.com.
Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives
http://www.celt.iastate.edu/teaching/RevisedBlooms1.html.
Texcompactor
http://textcompactor.com.
Text-Lex Compare
http://www.lextutor.ca/text_lex_compare.
Timetoast
www.timetoast.com.
TodaysMeet
http://todaysmeet.com.
Towtruck
https://towtruck.mozillalabs.com.
TweetChat
http://tweetchat.com.
TweetDeck
http://www.tweetdeck.com.
TweetShow
http://tweet-show.com/index.html.
Twitter
http://www.twitter.com.
UEFAP
http://www.uefap.com/index.htm.
Videonot.es
http://www.videonot.es.
Visual Thesaurus
http://www.visualthesaurus.com.
Visuwords
http://www.visuwords.com.

VocabProfile
http://conc.lextutor.ca/vp/.
Vocabulary Profiler
http://www.lextutor.ca/vp/eng.
Voki
http://www.voki.com.
Web VP
http://www.er.uqam.ca/nobel/r21270/cgi-bin/webfreqs/web_vp.html.
Wishbox
http://www.jotform.com/wishbox.
Word and Phrase – Academic
http://www.wordandphrase.info/academic.
Word Dynamo
http://dynamo.dictionary.com.
Wordle
http://www.wordle.net.
WordNet
http://wordnetweb.princeton.edu/perl/webwn.
WordSift
http://www.wordsift.com.

This volume presents blended methodological and language support services provided by CLIL (Content and Language Integrated Learning) experts to professors and students involved in CLIL teaching/learning across university departments. The theoretical framework underpinning the design of technology-enhanced learning environments and CLIL corpus-informed content-specific teaching materials aimed to support CLIL learning effectively are analyzed. The blended model devised to scaffold professors' CLIL teaching and students' learning of content-specific knowledge and foreign language in CLIL classes are thus thoroughly examined.

Università
Ca'Foscari
Venezia