DDL&CLIL Integration
Learning Activities and Resources Based on the Use of Corpora for CLIL Geography in a Cambridge International IGCSE® High School in Italy

Emma Abbate

Liceo Statale Alessandro Manzoni, Caserta, Italia

Abstract   This paper presents a series of e-CLIL (Content and Language Integrated Learning combined with ICT) activities based on the use of corpora, taken from experiments during the 2017-2018 school term in four Italian upper secondary classes. In the four classes considered in the present study, Geography is taught in English in the CLIL approach for Cambridge IGCSE (International General Certificate of Secondary Education). The Task-based activities of reading and listening comprehension, grammar learning and vocabulary building, have been focused on Cooperative Learning's strategy and submitted on a teacher monitored e-learning platform through constant application of BYOD (Bring Your Own Device) in the classroom.

Keywords   Content and Language Integrated Learning (CLIL). Corpus- based CLIL. Data Driven Learning (DDL). Technology-enhanced teaching. High School.

Summary   1 CLIL: Content Based Learning through Target Language. – 2 The Importance of Disciplinary Literacy and Language Awareness (LA) in CLIL Settings. – 3 Lexical Approach Applied to CLIL: A Lesson Plan Framework. – 4 Examples of FLAX Activities in a CLIL Geography Lesson. – 5 Vocabulary.com: A New, Original, Online Dictionary. – 6 Unconventional and Unaligned Corpora for DDL&CLIL Informal Approach: Youglish, Dollar Street and BrainyQuote. – 7 Conclusions.

> Language is not an abstract construction of the
> learned, or of dictionary-makers, but is something
> arising out of the work, needs, ties, joys, affections,
> tastes, of long generations of humanity, and has its
> bases broad and low, close to the ground.
>
> (Walt Whitman, *Slang in America*, 1885)

1 CLIL: Content Based Learning through Target Language

CLIL (Content and Language Integrated Learning) implies teaching subject-specific content through a target language (Marsh 1994): exercises, activities and classroom oral communication are all in the vehicular language. In CLIL settings, interaction and negotiation of meaning are fostered by cooperative learning strategies applied using the foreign language as a medium of instruction. Thus, one of the core aims of CLIL methodology is to encourage the acquisition of content-obligatory language (Coyle, Hood, Marsh 2010, 35).

In CLIL environments the teacher is required to identify learners' language demands and to support their linguistic needs while considering content vocabulary, functional language and language skills, according to the following three sets of questions suggested by Chadwick (2012, 4):
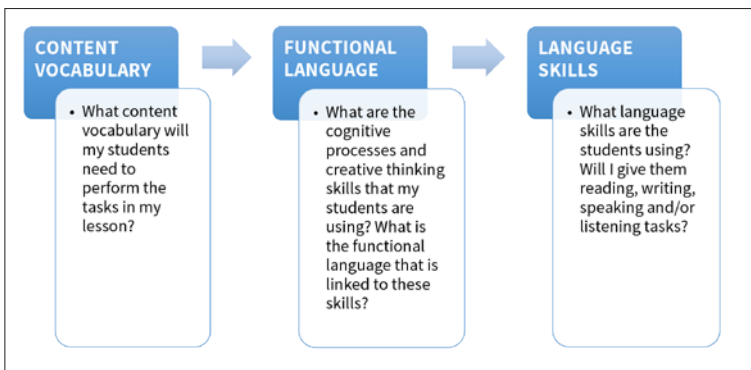

CONTENT VOCABULARY
- What content vocabulary will my students need to perform the tasks in my lesson?

FUNCTIONAL LANGUAGE
- What are the cognitive processes and creative thinking skills that my students are using? What is the functional language that is linked to these skills?

LANGUAGE SKILLS
- What language skills are the students using? Will I give them reading, writing, speaking and/or listening tasks?

Figure 1 Sets of questions to plan language support in content lessons. Adapted from Chadwick's framework (2012)

Similarly, Ball, Kelly and Clegg (2015, 75) identify three different layers of language in CLIL settings:
  i.   One related specifically to the subject area (subject-specific terminology);
  ii.   One that could be defined as cross-curricular and referred to as general academic language and the language that builds the classroom's conversation;
  iii. The interactional language of communication between people during the lesson; the so-called 'peripheral language' of instruction.

An awareness of these 'languages' is important for the CLIL teacher working with learners in an additional language.

The distinction operated by the aforementioned authors is pivotal to guide the CLIL teacher in detecting which language is needed for any given topic in the curriculum.

In a series of articles published on OneStopEnglish website, Keith Kelly looks at examples of process language in geography, covering common verb and noun phrases, structures and sequencing phrases.[1] He also gives examples of the language of comparison from the area of geography, covering grammatical patterns and structures, useful verbal phrases and specific prefixes,[2] and provides a comprehensive lesson plan with both a language and content focus accompanied by an alphabetical list of root words used to form words commonly used in geography.[3] Kelly's articles and his book on Geography language (published for Macmillan Vocabulary Practice series – Kelly 2009), are good Geographical resources that allow students to work in new terminologies.

In CLIL classrooms the learner has to be regarded as an active language producer and grammarian, a type of "self-contained language processor and grammar builder" (Dalton-Puffer, Nikula, Smit 2010, 7). This should be triggered by the teacher's rich content input, who should be able to predict all of their language demands, planning accurately the language given for content input and the language required to process the content output.

---

1 http://www.onestopenglish.com/clil/clil-teacher-magazine/your-clil/process/process-geography/551674.article.

2 http://www.onestopenglish.com/clil/clil-teacher-magazine/your-clil/comparisons/comparisons-geography/551281.article.

3 http://www.onestopenglish.com/clil/clil-teacher-magazine/your-clil/roots/roots-geography/551787.article.

2 The Importance of Disciplinary Literacy and Language Awareness (LA) in CLIL Settings

Latest research in CLIL has focussed on the role of subject-specific or disciplinary literacies (Meyer et al. 2015) in order to help learners become better meaning-makers, able to draw on content knowledge to communicate successfully across languages, disciplines and cultures.

The Council of European Union's recommendation of 22 May 2018[4] indicates literacy competence among the eight key competences for lifelong learning: it implies the ability to communicate and connect effectively with others, in an appropriate and creative way.

Literacy competence is conceptually connected to Language Awareness (LA). For David Marsh (2012, 58) LA it is essentially about "moving learners from viewing language learning as an object of study, towards an explicit understanding of how language is used in a variety of contexts".

A recent Recommendation from the European Commission (2018) underlines the role of LA as a channel for a full comprehension of how languages work and how people learn and use them: improving language learning implies encouraging inclusive education and the European dimension in teaching by promoting mobility, exchange and intercultural understanding. The document suggests reconsidering language teaching, breaking it down to the silos of language learning, and assisting teaching staff for a comprehensive language approach and linguistic-sensitive teaching: CLIL, enhanced with digital tools, is recommended as an efficient and innovative pedagogy to enrich language competence considered as a transversal element across the curricula.

Disciplinary Literacy and Language Awareness are both mediated from text, one of the main teaching tools used by teachers. However, school texts, also in L1, are often too difficult to be assimilated by students because of their verbosity and grammatical density. A major obstacle for learners of English-medium content is the cognitive load of texts that needs appropriate 'scaffolds': the study material should be adapted to learners' language proficiency (Mehisto 2012).

Scaffolding is a CLIL's pillar: it underpins language's building by supplying sentences, content-specific terms and collocations necessary to accomplish tasks. It supports pupils to express their ideas compatible with the subject and topic way (Bentley 2010, 69-70). A good scaffolding will enforce learners' cognitive Academic Language Proficiency (CALP), linking the gap between BICS (Basic Interper-

---

4 Council Recommendation of 22 May 2018 on key competences for lifelong learning. Official Journal of the European Union, 04-06-2018 (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0604(01)&rid=7).

sonal Communication Skills, the everyday language) and formal, abstract, subject-specific language, allowing the full Language Awareness (LA) acquisition and a deeper knowledge of contents.

The application of Academic Word List (AWL), developed by Averil Coxhead,[5] as a scaffolding tool to improve as well as differentiate learner awareness of CLIL-related vocabulary, has been underlined by the authors of *CLIL Activities* (Dale, Tanner 2012, 141-3).

3　Lexical Approach Applied to CLIL: A Lesson Plan Framework

The use of AWL in CLIL settings (as mentioned above) to enhance academic language proficiency represents a good and practical example of integration between CLIL and Data-Driven Learning (DDL): the two approaches promote authentic study materials and they are both aimed at the development of metalinguistic knowledge and at the increase of learner's autonomy. Thus the combination of those methodologies offers interesting tips and resources to CLIL teachers: the use of web corpora to implement CLIL learning through the creation of suitable and instructional materials ready-to-use. This is a scaffolding technique in itself.

As already emphasized, the integration of content and language in CLIL entails accurate consideration of the linguistic features of subject-specific proficiency and of the role played from written and oral communication within the learning process.

In a comprehensive article published in 2011, looking ahead in CLIL's investigation field, Christiane Dalton-Puffer indicated in systemic functional linguistics the future task of the research community in order to "build the necessary bridges to general learning theories based on ideas of discursiveness and performativity" (Dalton-Puffer 2011, 196).

Gathering linguistic data through corpora in CLIL environments allows students to acquire the topic's micro-language and the subject-specific grammatical features: teachers may use them to plan and design corpus-informed activities for a more student-centered approach to learning, according to the blended model provided by Giovanna Carloni (2014; 2015).

CLIL teachers can provide learners with a number of techniques from corpus linguistics and natural language processing for the investigation of subject-specific language patterns and words in large-scale data sets: it is a top-down approach, relying on frequency counts and statistical analyses.

5　https://www.victoria.ac.nz/lals/resources/academicwordlist.

In this paper we suggest a five-stage procedure of a typical lesson based on CLIL&DDL activities, summed up in the figure below:
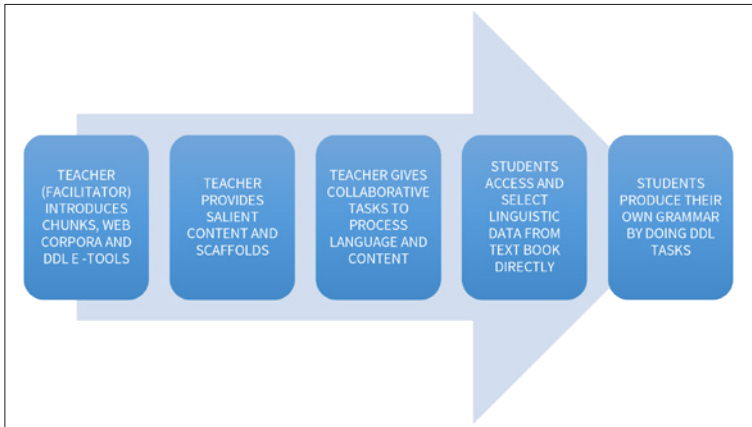
According to this procedure, the teacher, acting as a facilitator, introduces the concepts of linguistic chunks and e-corpora and shows the correct use of web tools for the lexical approach to salient content and cognitively engaging study material. Working on a basic micro-linguistic lexicon and on a corpus of LSP (Language for Specific Purposes) terms, identifying lexical chunks and structure of subject-text, reflecting on collocations and frequencies, are all activities mediated by the teacher in the first and second step of the procedure. In the last phases, thanks to training on tasks, students are able to explore and interpret the CLIL text on their own, accessing and selecting linguistic data directly.

This model helps teachers and students to recognise common and frequently recurring features of language in their subject. It is an aid for teachers to focus on what students need to understand and to identify potential language demands, and offers a principled process for constructing activities which integrate language and content within a lexical based approach.

In a DDL and CLIL integrated setting, students learn autonomously subject-specific collocations and colligations, improving the discipline's knowledge: the exposition to authentic language data and the engagement in active analysis of data to uncover language usage patterns, converting them from passive learners to learner-researchers. DDL is based on the cognitive process of inference that is likely to enhance the input of content-specific vocabulary.

Combined with CLIL methodology, DDL develops lexical links, working on semantic fields determined by the subject's specificity.

Through corpora, lexicon is presented in a contextualized frame and disciplinary knowledge is not acquired passively, but is used to complete tasks and group activities. Subject micro-language terms are often monosemic and should be presented in their own context so as to be learnt easily.

Moving the focus from how to say to what to say is the basis for a lexical enrichment that favours both the development of language skills and subject's mastery. In CLIL settings particular attention to the lexicon and the lexical repertoire are required: the teacher must be able to anticipate words that students should know but don't.

Some of the keywords may have been encountered in reading or heard in class, but not completely understood. Others might be 'kind of' understandable, but students are not confident enough to use them in speaking or writing: I call them *grasp words*. In order to prepare pupils to handle web corpora and to get acquainted with practicing essential DDL skills, teachers should gradually introduce chunks and subject's language patterns and structures training students with grasp words contained into topic-specific chunks.

This sort of 'chunks workout', a series of lexicon's expansion and reinforcement activities, warms 'learners' muscles' before 'lifting' the heavy e-corpora collection. It is a way to develop the language awareness and ability to chunk as part of a text assault strategy.

Reflecting on language is more important than developing lexicon quantitatively.

Chunks are two or more words which are often used together (e.g. 'in the context of', 'a large number of'). In order to read, speak and write articulately, students need to possess general and topic-specific chunks. Phrases of a language can be learned in chunks – which breaks the learning process into easy stages. The progress from a single disconnected term to chunks will boost conceptual comprehension and production in both BICS and CALP.

Working with partners, students can get used to recognizing lexical sets and chunks, making 'pragmatic generalizations' about the patterns that are typical of the subject's academic text that they are studying. Induction-type tasks introduce students to detect recurring patterns and start an inferring learning process that encourages reflection and increases study motivation.

Textalyser[6] is a free software which allows teachers to find the most frequent phrases and frequencies of words in a text's page that can be copied and pasted in the box to calculate the lexical density as well.

---

6 http://textalyser.net.

The frequencies and chunks highlighted from the program can be used for targeted exercises of content-specific vocabulary building: students are asked to detect general and topic-specific chunks in order to activate semantic association as in the following activities designed for a Geography CLIL lesson (topic: plate tectonics):

Activities:

1. Chunks: *Look at the text with a partner. One of you find and writes the general chunks and the other the topic specific chunks.*

General chunks	Topic specific chunks
such as	earth's crust
located at	molten magma
affected for	world's climate
a number of	plate boundaries
ejected at	earth's interior
top of	earth's surface
	volcanic eruption

2. Lexical sets: In pairs or alone, read the text. Make lexical sets of words/phrases for:

Types of volcanoes	Features of volcanoes

Figure 3 Examples of chunks activities on a Geography text

Through this form of *enquire-based learning*, the student gradually becomes an inquirer and discoverer of new keywords, a sort of 'linguistic Hercule Poirot' scanning and identifying discipline-specific language features.

To familiarize with this study technique, as homework, learners can be given DDL tasks that require them to create word webs using terms found in the CLIL text. A word web is a picture that helps to connect word together increasing the target language's vocabulary. Online word web generators, such as Visuwords,[7] can help students to design their own word webs.

After the teacher has introduced students to the concept of corpus as a good resource by justifying its use and presenting the various data that can be obtained from it, pupils can start a monitored computer-based practice to use online corpus.

---

[7] https://visuwords.com.

Corpora allow learners to investigate language rules for themselves and, in doing so, build up an understanding of how language works.

4 Examples of FLAX Activities in a CLIL Geography Lesson

When students and teachers have acquired enough technical skills they can be exposed to authentic data in the form of concordance lines: at this point they are ready to approach FLAX, Flexible Language Acquisition.[8] FLAX is an automated language system to extract salient linguistic features from academic texts: it displays them through an interface developed for ESP (English for Specific Purposes) students who are learning academic writing and reading.

It is an inductive tool that stimulates higher cognitive process and learners' autonomy and offers them the chance to investigate text by experiencing repeated examples of lexicon-grammatical features' use through Language Awareness activities. In the section named "FLAX Resource Collections", users can find collections that draw on large reference corpora like the British National Corpus (BNC) and even larger datasets from Google and Wikipedia. They present a lot of examples of language in context for some of the most demanding areas of English language learning – collocations and phrases – where there are hundreds of thousands of possibilities for combining words.

In the table below there are some example of CLIL Geography questions on population dynamics and volcanoes (both topics are part of Cambridge ICGSE Geography syllabus[9]) based on a FLAX research of the relevant topic's terms:

Table 1 Examples of CLIL investigation tasks FLAX-based

Ares of investigation	Sample questions
Collocation/phraseology	What common adjectives are generally used in front of the word 'volcano'? What adjectives can we use to describe a volcano?
Grammatical questions	Is 'population' a countable or uncountable noun?
Connotation/semantic	What are the differences between 'extinguish' and 'extinct'?

The use of such investigative tasks enables learners to move beyond the words to what is really at the heart of Geography learning: the

---

8 http://flax.nzdl.org.

9 https://www.cambridgeinternational.org/programmes-and-qualifications/cambridge-igcse-geography-0460/.

words that convey core information of the texts can be used to plan practice exercises to use the vocabulary at word, sentence and text level like the ones shown in order to get students to create their investigative tasks on their own, using the models provided from their teacher. Analyzing new vocabulary in web corpora sentences ensures comprehension and correct formation in a greater language context.

In Geography texts, for example, the future events expressed by the clause "be likely + to-infinitive", the cause and effect connection with "due to" or the use of phrasal verb "depend on", are very common: learners can get used to Geography text's language patterns through a series of DDL tasks based on FLAX's sections named "Learning Collocations", "BAWE (British Academic Written English)", "Book Phrase" as the ones following:

- A FLAX exercise on the use of "likely":

Likely is an adjective. How can we use to express future tense (give the pattern)? Does it express something will probably happen or something that won't happen for sure? Can we use it before a noun? Which verbs and adverbs are used with "likely"? Give example. We can follow *likely* by a *that*-clause with *will*. Does this pattern is more common than *likely/unlikely* + *to*-infinitive? Check the most frequent occurrence in the learning collocation section in Flax.

- A FLAX Learning Collocation's activity on the verb "depend on":

1. In the learning collocation, using the BAWE (British Academic Written English) corpus, look up the verb "depend". What's the most common collocation?
2. Open the window with the example sentences. What word and grammatical pattern tend to mostly commonly come after the collocation?
3. What verbs tend to came before it? What often comes before the verb?
4. Take a minute to further investigate the "depend" collocation in either the *Book Phrase* or the *British Academic Written English Collections* section in Flax. Then try to create your own sentence using the collocation, following the normal grammatical pattern (colligation) identified above.

- A FLAX task on the connective "Due to":

What's the meaning of "due to"? Check it on Flax collocations' section then answer: do you know similar expression in English to express the same meaning?

Students can also work on FLAX-generated concordance lines printed on paper rather than having them query for and explore concordance lines on computers.

5   Vocabulary.com: A New, Original, Online Dictionary

Another useful tool to strengthen CLIL lexical learning and to help teachers to prepare authentic class material from digital texts creating words lists and quiz documents is Vocabulary.com:[10] an online dictionary that gives friendly, funny but rigorous explanations of words, easy to understand and remember, like this one:
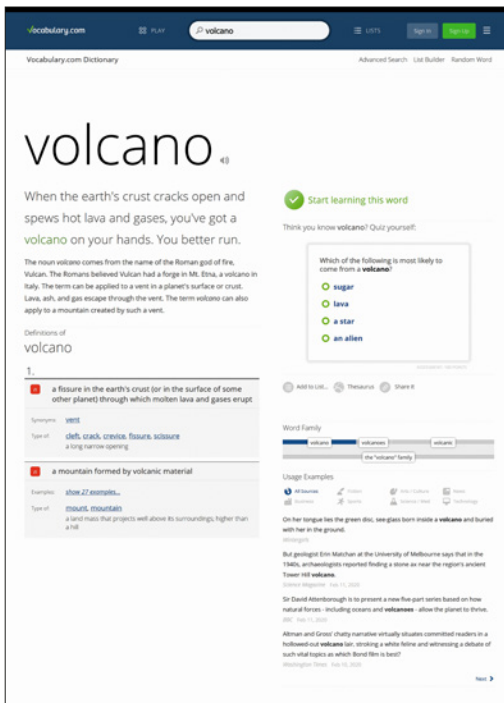
Figure 4 *Volcano*'s definition on Vocabulary.com

10 https://www.vocabulary.com.

Vocabulary.com provides more conversational and student-friendly explanations of words in addition to typical definitions, together with audio pronunciations, "word families" of inflectionally and derivationally related forms (indicating relative corpus frequency for each form), and corpus example sentences (from current newspapers, magazines, and literature) arranged by genre and topic.

Another added value of Vocabulary.com is the gamified environment: CLIL teacher can create a student's customized topic-specific terms list (also simply by copying and pasting an online text) and, on this list, the system will generate some activities like pictures quiz, vocabulary jam and spelling challenge as incentives to keep learners intensely engaged. In the examples below two activities on the key terms *ridge* and *magmas* (topic: Plate Tectonics) are shown:
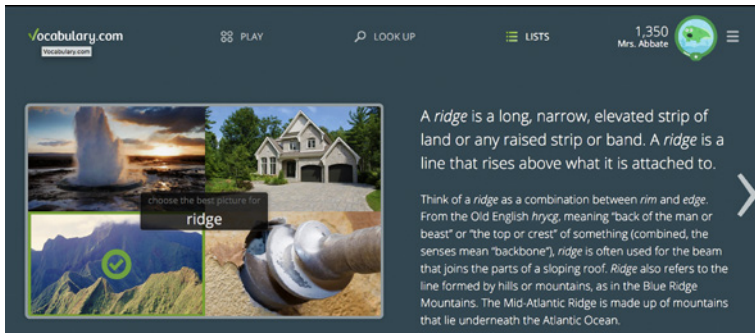
Figure 5   A visual quiz on Vocabulary.com
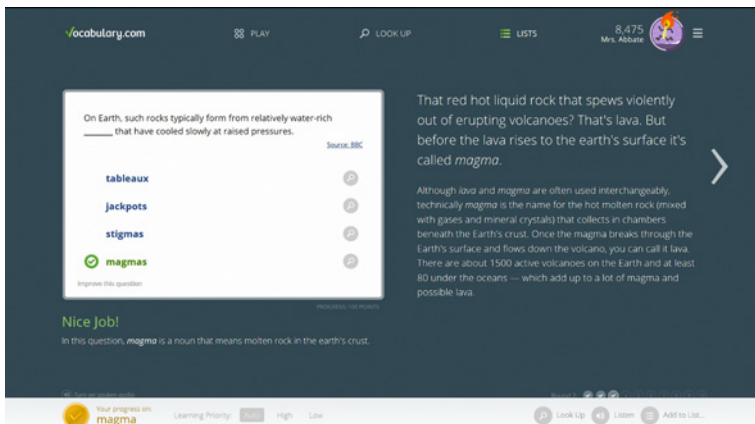
Figure 6   A multiple choice exercise on Vocabulary.com

6 Unconventional and Unaligned Corpora for DDL&CLIL Informal Approach: YouGlish, Dollar Street and BrainyQuote

Apart from traditional collections, the Internet allows one to find many unconventional corpora for atypical consultation activities that might inspire original CLIL tasks:[11] for example, video watching activities. YouGlish[12] is a great tool that allows students to hear the correct pronunciation of words in English, using YouTube videos in the native language and in a meaningful context.[13] The learner simply enters the word(s) they would like to hear and study, then chooses between English, American or Australian pronunciation – or all together. Each video has subtitles in order to facilitate listening and understanding: the entered word is highlighted and shown many times in various sample sentences, which makes this resource very effective in studying new subject content terms in the genuine sentence's context, while improving pronunciation.

---

11 For an original and new perspective on corpora check the articles by Jamie Keddie on *One Stop English* (http://www.onestopenglish.com/skills/vocabulary/corpora/).

12 http://youglish.com/.

13 To find out the potentiality of Youglish as teaching tool read the e-article "Youglish: using authentic English videos for pronunciation and presentation practice" by Lara Wallace and Cassidi Hunker (http://newsmanager.commpartners.com/tesol-callis/issues/2017–09–12/7.html).

Figure 7 Search of the word *volcano* on YouGlish

The author of YouGlish, Dan Barhen, has designed another tool: *Fraze It*.[14] This allows one to investigate a corpus of phrases to detect how words are used in the sentence's context. Naturally, online dictionaries also offer usage models, but the added value of *Fraze* is that it gives the chance to look for defined types of phraseology such as interrogatives, negatives, present perfect sentences, etc. For any entered word, the dictionary definition is given together with synonyms, translations, pronunciation (using YouGlish), and even pictures. Signing up is free and enables one to save the searches' results.

Another atypical corpus could be considered the pictures 'web collections: in particular, for CLIL Geography teachers. I suggest the interesting photo-based project *Dollar Street* invented by Anna Rosling Rönnlund:[15] it is a website where images from homes from all over the World have been collected as data to show differences and similarities in income.
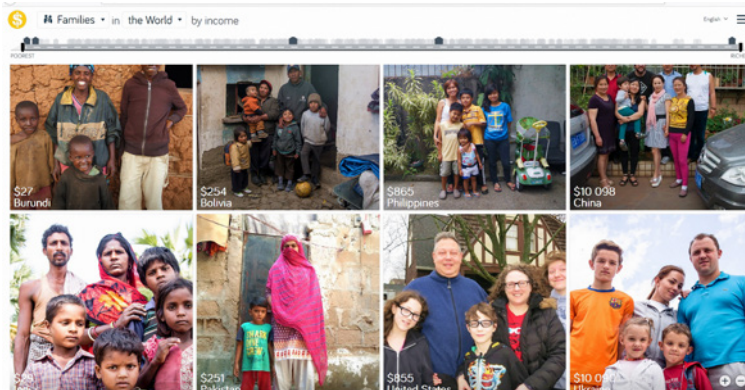

Figure 8 *Dollar Street*'s homepage

Quotations' corpora, such as the ones you can find entering a keyword in *Brainyquotes*, are also a good source for learning new words in a unusual way.[16]

---

14 https://fraze.it/.

15 https://www.gapminder.org/dollar-street/matrix.

16 https://www.brainyquote.com/.

7 Conclusions

Lexical approach to text-handling has had, in my experience, beneficial effects on learners: working with web corpora in the classroom (including unconventional ones), allowed my students to contribute actively to the lesson despite being apprentices within the subject community. It offered them the topic terminology which may result equally 'extraneous' and difficult in L1.

In the DDL&CLIL integrated approach illustrated in this brief, language has not been considered as a mere presentational tool but plays a pivotal role for meaning-making, thinking and the knowledge construction medium. The use of language features such as collocations, subject-specific vocabulary, and academic vocabulary, has been improved through corpora based tasks, both receptive and productive, appropriate to the students' learning pace.

The activities on web corpora have notably promoted co-operation among pupils. The active learning tasks have embedded an extensive view of phraseology into routine teaching by joining more significance to word determination: this has massively improved my students' fluency in L2.

Considering chunks, being able to recognize and use them in the productive phase, makes input comprehensible and allows learners to organize more easily the correct output, increasing at the same time their academic skills and STT (Student Talking Time).

Language is not an abstract construction: as in Walt Whitman's quote that I cited when introducing this article, language is "close to the ground". CLIL's socio-linguistic dimension, focussed on the pragmatic use of language in a communicative context, reminds us of this closeness.

References

Ball, P.; Kelly, K.; Clegg, J. (2015). *Putting CLIL into Practice*. Oxford: Oxford University Press.

Bentley, K. (2010). *The TKT Course CLIL Module*. Cambridge: Cambridge University Press.

Cambers, G.; Sibley, S. (2018). *Cambridge IGCSE and O Level Geography. Coursebook*. Cambridge: Cambridge University Press.

Carloni, G. (2014). "Content and Language Integrated Learning: A Blended Model in Higher Education". *The International Journal of Technology, Knowledge, and Society*, 9(4), 61-71. `https://doi.org/10.18848/1832-3669/CGP/v09i04/56358`.

Carloni, G. (2015). *CLIL in Higher Education and the Role of Corpora. A Blended Model of Consultation Services and Learning Environments*. Venezia: Edizioni Ca'Foscari. `https://doi.org/10.14277/978-88-6969-038-9`.

Chadwick, T. (2012). *Language Awareness in Teaching*. Cambridge: Cambridge University Press.

Coyle, D.; Hood, P.; Marsh, D. (2010). *CLIL Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Dale, L.; Tanner, R. (2012). *CLIL Activities*. Cambridge: Cambridge University Press.

Dalton-Puffer, C. (2011). "Content-and-Language Integrated Learning: From Practice to Principles?". *Annual Review of Applied Linguistics*, 31, 182-204. `https://doi.org/10.1017/S0267190511000092`.

Dalton-Puffer, C.; Nikula, T.; Smit, U. (eds) (2010). *Language Use and Language Learning in CLIL*. Amsterdam: John Benjamins.

European Commission (2018). *Commission Staff Working Document Accompanying the Document Proposal for a Council Recommendation on a Comprehensive Approach to the Teaching and Learning of Languages*. Brussels. `https://eur-lex.europa.eu/resource.html?uri=cellar:36ae2f05-5dc7-11e8-ab9c-01aa75ed71a1.0001.02/DOC_1&format=PDF`.

Kelly, K. (2009). *Geography Macmillan Vocabulary Practice Series*. Macmillan Education.

Marsh, D. (1994). "Bilingual Education & Content and Language Integrated Learning". International Association for Cross-cultural Communication (eds), *Language Teaching in the Member States of the European Union*. Paris: University of Sorbonne.

Marsh, D. (2012). *Content and Language Integrated Learning (CLIL). A Development Trajectory*. Cordoba: Universidad de Córdoba. `https://helvia.uco.es/bitstream/handle/10396/8689/2013000000658.pdf?sequence`.

Mehisto, P. (2012). "Criteria for Producing CLIL Learning Material". *Encuentro*, 21, 15-33. `http://www.encuentrojournal.org/textos/3.%20Mehisto_rev.pdf`).

Meyer, O. et al. (2015). "A pluriliteracies approach to content and language integrated learning – mapping learner progressions in knowledge construction and meaning-making". *Language, Culture and Curriculum*, 28(1), 41-57. `https://doi.org/10.1080/07908318.2014.1000924`.

Websites

https://www.brainyquote.com
https://www.cambridgeinternational.org/programmes-and-quali-
   fications/cambridge-igcse-geography-0460/
http://flax.nzdl.org
https://fraze.it/
https://www.gapminder.org/dollar-street/matrix
http://www.onestopenglish.com
http://textalyser.net
https://www.victoria.ac.nz/lals/resources/academicwordlist
   (AWL by Averil Coxhead)
https://visuwords.com
https://www.vocabulary.com
http://youglish.com/