Linguæ &

Rivista di lingue e culture moderne

1
2019

Il ruolo e le sfide dei Centri Linguistici universitari – Parte prima
a cura di Enrica Rossi

Cesare Zanca

Università degli Studi di Siena

Language Centres, Online Authentic Materials and Learners' Needs: Improving Autonomy and Discovery in Language Learning

1. INTRODUCTION: ONLINE LEARNING, SELF-ACCESS CENTRES AND ONLINE RESOURCES

Most language centres in Europe are now implementing or at least considering the implementation of courses, activities or websites that imply or recommend an autonomous use by the learner of authentic online materials in the foreign language. Similarly, the use of online resources, Google and corpora for language learning has been supported by many scholars and is at the core of several approaches. Milton 2006, for instance, encourages the use of Google searches by learners to check and refine their writing. Choi 2017 describes how e-resources can be integrated in a traditional self-access centre and Choi *et al.* 2018 describe examples of Self-Regulated Learning (SRL) of vocabulary. Other similar studies explore a Google-Assisted Language Learning (GALL) approach (Chinnery 2008). They investigate the advantages and possible pitfalls of the use of Google in searching for foreign language related information. Similarly, large collections of texts – both online and on personal computers – have been considered by the Corpus Based Learning (McEnery 1990, 370) and the Data Driven Learning (DDL) approaches (Johns 1991; 1994) to be invaluable resources in order to explore the usage and the contextual features of language items: teachers and learners observe concordances and statistical information so as to explore authentic texts and analyse lexi-

cal features. Other studies consider the use of online resources and corpora together (Conroy 2010; Shei 2008a, 2008b; Zanca 2013, 2014, 2017, 2018).

One common feature of these studies – one that this study considers to be of paramount importance in the process involving an effective use of authentic materials by learners both in self-access centres and online courses – is that the learner is seen as an explorer, "a research worker whose learning needs to be driven by access to linguistic data" (McEnery *et al.* 1990, 370). Nevertheless, a direct, unfiltered and autonomous exploration and use of online materials might turn out to be to be extremely challenging, even for advanced learners. The acquisition of autonomy and the ability to explore require a certain amount of confidence and expertise in the tools of the trade. Dealing with corpora, for instance, involves the ability to choose the right corpus, formulate an effective query, read and interpret concordance lines, frequency lists, collocation lists and so on. As argued in previous studies (Zanca 2013; 2014), the 'technicalities' of corpus investigations are probably the main reason why DDL activities and corpus-based learning is still far from familiar to the vast majority of language teachers and learners in schools and universities. When using Google or online tools, the learner's task may well appear more informal and straightforward as we all use Google, online dictionaries, Google Translate or similar resources. However, when applied to language learning even this kind of resource might be employed in partial or inappropriate ways. Conroy, for instance, describes a research project in which corpora and online resources were employed with Australian EAL (English as an Additional Language) university students. He found that:

> students typically used crude and often inappropriate search strategies. Many students were, for example, unaware that they could use quotation marks to focus their search onto specific and uninterrupted word sequences (e.g. a verb and an associated preposition). Also, during training it was observed that several students were often unable to search with sufficient critical skills to detect non-native texts and the non-standard usage often contained therein. […] Thus, contrary to what is commonly believed, young university students, many from cultures with a strong uptake of computer technology, do not appear to be skilled in using the Internet to do anything more than content searches. (Conroy 2010, 879)

These remarks fully correspond to my personal experience with university students: they are increasingly sophisticated in their use of online resources, but when it comes to making the most of them for language learning purposes, they need some guidance and training. Conroy concludes by arguing that:

> Internet-based corpus techniques offer a relatively new and under-used method for EAL university students to independently enhance their academic literacy and English language proficiency. At the moment however, this type of training appears to be either limited or unavailable in many Australian universities, who appear to have largely ignored these technologies. (Conroy 2010, 880)

A similar lack of training and awareness is likely to be experienced in most language centres in Europe affecting both learners and teachers, tutors or self-access advisers. An 'intelligent' use of Google search strategies, of other online tools and of corpora can be, as clarified in the next sections, a very helpful resource for the language learning process, but it is often disregarded and requires some methodological consideration and training.

1.1. The specific role of corpora and GALL in language learning: reducing the priming gap

One of the most relevant theoretical background studies into the opportunity of introducing corpora in language teaching courses in classroom – and, by extension, in online or blended courses and self-access centres – is Michael Hoey's idea of Lexical Priming, defined as:

> the process whereby "[a]s a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and cotexts 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". (2005, 8)

On the basis of theoretical assumptions shared by most corpus linguistics scholars (e.g. Sinclair 1991; Partington 1998), Hoey argues that an important aspect of language learning relates to: (a) the tendency of words to link with other lexical items to create patterns and (b) to the distinctive use of lexical items in specific contexts, which might significantly modify their meaning and interpretation. Combinations of words like hard rain, die hard, squeezed hard, hard to believe, hard luck, or a whole clause like the time has come create patterns (Hoey 2009, 1). The words in a pattern might be grammatically and lexically separated, but when it comes to the use of language, they can be considered as a single unit. This is evident for idiomatic expressions and proverbs like "When the going gets tough" (Zanca 2017, 117) or in Italian "Hai voluto la bicicletta, adesso pedala" ("You made your bed, now lie in it", Zanca 2014): when we hear the first part of the pattern we can predict the

words to come. Another common example are multi word expressions and the tendency to create acronyms like IDK (I Don't Know) used informally (e.g. in online chats) but also in more formal text types[1]. As for context, native speakers of English subconsciously know that a phrase like Kieren Perkins of Australia (Zanettin 2001, 212) probably comes from the sports section of a newspaper, where of + nationality patterns are more frequent than in other textual contexts. According to Hoey, a native speaker's fluency derives precisely from a lifelong subconscious process of noticing grammar and lexical features but also contextual features and lexical patterns:

> Noticing all these things is what makes it possible for a speaker to use the right phrase in the right context at the right time […] This is how native speakers are able to be fluent and because the things they say are subconsciously influenced by what everyone has previously said to them, it also explains why they almost always sound natural. (Hoey 2009, 1)

Accordingly, many of the linguistic limitations of non-native speakers derive, primarily, from their limited exposure and insufficient priming in the uses of the foreign language in context and, secondly, from the interference of their first language priming (*ibid.*). This dimension of linguistic performance goes beyond the ability to convey a message and ensure mutual comprehension as emphasized by the traditional communicative approach to language learning. The stress is on acquiring fluency, phraseology and on cultural, pragmatic and discourse appropriacy. A great deal of what non-native speakers say or write is not grammatically, semantically or lexically wrong, but simply 'doesn't sound right'. For instance, Italian learners of English, prompted by their first language priming, might opt for Can you make me a favour? which is grammatically feasible and is perfectly understandable, but is not used and something 'you do not say' because a native speaker's priming suggests alternative patterns (such as could you do me a favour?). In order to enhance fluency and reduce the 'priming gap' Hoey suggests that teachers should help learners acquire the means to (1) explore authentic texts and (2) focus their attention on the linguistic context and patterns because

---

[1] It appears, for instance, in an article published in the Art and Design section of *The Guardian*: "So Frieze is like, definitely a Thing™ isn't it? Some ppl hate it, some ppl live for it (idk anyone that actually lives for it, but if it's still goin, they must be out there…)", https://www.theguardian.com/artanddesign/2018/oct/05/the-five-best-and-worst-things-at-frieze-2018-according-to-the-white-pube.

> Native speakers have acquired a large corpus of examples of the words of English in their typical contexts, and from this they learn how the words are used. […] learners need to be exposed to as much authentic material as possible and teachers need to find ways of accelerating the priming processes by drawing attention to the patterns of use that a text (or conversation) reveals. (*ibid.*)

Learners do not possess an extensive internal corpus of patterns in context, therefore they should be guided in the use of external resources and databases – such as online resources and corpora – to try and compensate for their 'priming gap'. How can this be done? The following section suggests some practical ideas.

2. DISCOVERY ACTIVITIES AND AUTONOMOUS 'INTELLIGENT' USE OF ONLINE RESOURCES

As mentioned above, Hoey and Conroy suggest that, in the definition of linguistic educational goals, university advisors and teachers should help learners use online resources and corpora, authentic materials and language discovery strategies to improve their foreign language skills. Unfortunately, most teachers are still not aware of such studies or trained in this field. The approaches outlined in section 1 of this study are not well known and school courses are usually based on a traditional, fixed syllabus that leaves little room for creativity. In this respect, language centres could play an important and pioneering role: online tutors and language/self-access centres are usually more flexible and could easily devise activities based on language discovery, suggest methodologies, offer lists of sites, training and information with an aim to foster language exploration and awareness.

A detailed guide to all possible procedures, activities and materials in this field is beyond the scope of this paper but the following sections outline a series of resources – both simple and more complex – I have experimented in my university courses, and of some practical examples derived from my students' writing and translation activities. The suggestion is that corpus linguistics tools and methodology and frequently used online resources can be integrated and included in language courses.

2.1. Google and search engines

We may feel that we do not need to tell our students how to use Google, but, as mentioned above, this is often not the case: many learners do not know how to use more advanced Google searches or how to select some areas from its endless database of materials. The following are some suggestions from my courses (see also Zanca 2018).

2.1.1. Search Engines and spelling or syntax mistakes

A common yet simple way to search Google for reliable and quick answers is to check the spelling of words, or the structure of some fixed phrases like phrasal verbs. If we type *mispelt** in Google rather than *misspelt*, we immediately get links to grammar pages indicating a possible mistake. A similar outcome follows a search for *rather then**, *pronounciation**, *look this car** and the like. If learners still have doubts about the reliability or quality of Google's suggestions, they should employ an online checking routine, as described later in this paper.

2.1.2. Google advanced searches

Better and more reliable results are obtained by using more advanced Google searches. Not all users know, for instance, that if we enclose our search string in quotation marks, Google tends (in my experience it does not always work properly) to display exact matches for the string. If we type *I took up it*, for instance, we get a series of results relating to the individual words, but if we type *"I took up it"* in quotation marks, the results will signal that it is a common mistake made by learners of English. The asterisk in a string in quotation marks stands for 'any word or phrase in that position'. So if we type *"I'm looking * to"* we'll find that the most frequent word used in that position is *forward*; if we find a headline like *Panetta's Mission: Can He Cut the Pentagon?* [2] and want to discover the meaning of the second clause we can type *"Can he cut the *"* in Google to understand that it is an idiomatic expression and that the following word is usually *mustard*. We can also use

---

[2] https://www.americanprogress.org/issues/ext/2011/04/28/9376/panettas-mission-can-he-cut-the-pentagon/.

other search operators like *OR* to search for X or Y or both ("can OR cannot cut the"), – or + to exclude or necessarily include the following word ("cut the + mustard" or "cut the – mustard") and round brackets to group multiple terms or search operators. A full list of operators and advanced search tips is available at https://www.google.com/advanced_search.

Another useful feature of search engines is that they usually work by considering the most likely relations among the searched for items in a query. In a sense, they consider the possibility that they might be part of a pattern in a particular context and show different results if we search for a word on its own or together with other words from different contexts. The example of *cut the* together with *mustard* illustrates this circumstance, but learners could also try searching for the word *badger* on its own and then for *badger him*, *badger vehicle* and *badger Wisconsin* to see how the relational system works. Finally, by adding the terms *definition*, *synonyms* or *traduzione* to a fairly obscure word such as *ashen*, learners can obtain more relevant information about its meaning, synonyms and translation in Italian and fewer results concerning a popular video game: adding contextual elements to our search is definitely useful to explore language in context.

2.1.3. Limiting the research field in Google

The advanced use of Google queries outlined above is discussed in more detail by many scholars (see for instance Chinnery 2008; Conroy 2010; Eu 2017; Geiller 2014; Sha 2010; Wu *et al.* 2009), but search engines can also be used in a more reliable and more corpus-like way by limiting the search field to some of its sections. This can sometimes be necessary, because one of the main differences between a linguistic corpus and Google's database is that we do not know anything about the latter's extension or the kind of texts it uses: this means that we cannot always trust the reliability of the results, as they may be written by non-native speakers or inaccurately. Our learners can obtain far more reliable information by confining their searches to smaller and more dependable sections, such as Google Scholar, Google News or Google Books [3]. If we search for the string "I'm looking * to" in one of these sections, the number of results will be far smaller, but much more accurate and we will find valid linguistic observations like, for instance, that the expression is com-

---

[3] https://scholar.google.com; https://news.google.com; https://books.google.com.

monly followed by a present participle, a noun or a pronoun. By exploring the idiom *made your bed "now * in it"* (Zanca 2014) in Google News, the learner can explore its variations in authentic texts and find that sometimes instead of *lie* we can find other options such as *sleep*, *you can lie* and *you have to lie*.

Sometimes it may be useful to limit the search to a defined period of time. Zanca 2014, for instance describes how a student who thought the expression *you made your bed, now lie in it* might be obsolete and outdated, was prompted to reconsider by searching the previous three months of Google News texts and finding hundreds of occurrences of the idiom or its variants.

A further useful way to control Google searches and explore more appropriate contexts can be obtained by employing the *site*: operator. For instance, learners might try and see the different results they get by typing *thongs site:.uk* or *thongs site:.au*.

2.1.4. Google images

Surprisingly, many of the students in my courses – including teachers in training courses – are not aware of the fact that Google images can be a very effective and often amusing way to explore language. The aforementioned search for thongs site:.uk or thongs site:.au is much more effective using Google Images. Some clearly culture-bound expressions, like playing conkers, Shepherd's Pie or Toad in the hole [4], are much easier to understand if learners can see what they look like and images are excellent ways to add context to the lexical priming process. Also the consequences of some syntactic and lexical choices can become 'visible': if our learners do not know whether they should write "phone card" or "card phone" or think that a "high window" and a "tall window" are the same thing. the difference will immediately become clear when consulting Google Images.

2.1.5. Wikipedia and other encyclopaedic or specific sites

Linguistic exploration may also require encyclopaedic information or the consultation of sites dedicated to the linguistic items we are dealing with. Asked

---

[4] These and other examples mentioned in the study are not invented, but emerged during classroom activities.

to translate into Italian the expression "Simon Stevens, the NHS England chief executive, told the Public Accounts Committee on Monday…" [5] one of my students realized he lacked some information about the British cultural context: in Hoey's terms, his priming was not as accurate as that of an average British native newspaper reader. He had to find information about the NHS (using Wikipedia and the NHS site), Simon Stevens (Wikipedia) and the Public Accounts Committee (Wikipedia and PAC site), in order to compensate for his priming gap.

2.1.6. Blogs, online forums and interactive apps

A number of students have informed me that they use interactive blogs and online forums to ask native speakers and translators about the interpretation and translation of textual elements, as in the case of the previously mentioned proverb "you made your bed, now lie in it" [6]. Some apps, like *hinative* [7], might also be worth considering for future consultation.

2.1.7. Dictionaries, glossaries, Google translate and online parallel translation corpora

As mentioned above, online dictionaries are fundamental resources for learners. Unlike traditional paper-based dictionaries, they allow for complex searches, including multi-word unit queries, links to translations, synonyms, encyclopaedic information, images, sounds and sometimes even concordances and collocations [8]. Online translators, such as Google translate are very popular, and can be considered as important and serious commercial projects which can prove to be very useful, when we have no idea about the meaning of a term or a text in a foreign language. However, if we try and insert *Toad in the hole*, *playing conkers* or *Faculty of Public Health Medicine* into Google translate, we get, at least in Italian, literal and highly unsatisfactory translations.

---

[5] https://www.theguardian.com/society/2017/jan/11/nhs-england-chief-executive-simon-stevens-funding-theresa-may-public-accounts-committee.
[6] https://forum.wordreference.com/threads/hai-voluto-la-bicicletta-e-allora-pedala.714846/?hl=it.
[7] https://hinative.com.
[8] See for instance http://nav4.stringnet.org/.

Reverso Context [9], Linguee [10] and other online resources based on parallel translation corpora allow for similar considerations: students often use them without realising that some of the translations, may be inaccurate and might lead to inappropriate choices. Automatic translators and web based translation corpora are not always reliable or offer solutions that are appropriate to the learners' needs and need to be thoroughly verified (see section 4).

2.1.8. *Other online resources*

Innovative online resources appear every day. Learners engaged in specific tasks may garner useful information from online Thesauri, like the *Visual thesaurus* [11], terminological data-banks [12] and glossaries [13], grammar websites [14], phraseological tools [15], text analysis tools like *The Compleat Lexical tutor* [16], sites that bring together many different tools such as *The Language Toolbox* [17] and many others. We cannot predict what kind of online resources will be available in the future, but all of them, if our goal is an 'intelligent' use of the internet, require a certain amount of training.

3. FROM GALL AND ONLINE RESOURCES TO CORPORA

Although specific training is needed, the tools, websites and resources outlined above are fairly simple to use and in many cases they allow for an effective exploration of linguistic items in context and reliable solutions to the learners' needs. As mentioned above, they can help reduce the priming gap by exploring and observing lexical patterns, images and other paralinguistic

---

[9] http://context.reverso.net.
[10] https://www.linguee.com/.
[11] https://www.visualthesaurus.com.
[12] E.g. http://iate.europa.eu.
[13] E.g. http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm.
[14] E.g. https://www.ucl.ac.uk/internet-grammar/home.htm.
[15] https://www.phrases.org.uk/meanings/; https://www.wordandphrase.info/academic/x.asp.
[16] https://www.lextutor.ca/.
[17] http://itools.com/language.

features, variation in time, places and culture, and provide valid explanations. Why, therefore, should language learners turn to more technical, unfamiliar and professional tools such as corpora?

Corpora, both online and stored on personal devices, are collections of searchable texts that have been selected according to defined criteria. This latter feature makes them usually more reliable in terms of the quality of the linguistic information they can offer as compared with other online resources [18]. They have been considered for many years a potentially extraordinary resource for language learning and teaching, but, as reported by many studies (e.g. Boulton 2010; Meunier 2011; Römer 2008), teachers and learners are still far from familiar with them, probably because some of their more technical features – such as frequency lists, keyness value, collocation lists and concordances – are considered too challenging for non-specialist users. Nevertheless, this paper outlines three main reasons why online corpora can and should be used.

The first is that many corpora are now easily available online and their use is, at least for basic queries, not significantly different from the tools outlined previously. In my teaching experience, learners are now less intimidated by the more technical aspects of corpora mainly because they are used to searching the web, but also because the idea of tags (e.g. for online images) and the relevance of statistics and quantities (e.g. because of the number of likes or followers in social networks) are more familiar than before (see also Zanca 2017).

The second pertains to the previously mentioned risk that many online resources might be linguistically inaccurate and misleading: linguistic features need to be evaluated and considered in a specific context and corpora can provide information about the kind of texts and contexts used (see the 'check online procedure' suggested next).

The third is that the investigation of corpora allows for queries and explorations that are impossible with other online tools. For instance, by using a specific software like *AntConc* (Anthony 2019), dedicated websites like Mark Davies' corpora at Brigham Young University [19] or the Sketch Engine for Language Learning [20] sites, learners can (a) use the asterisk operators to

---

[18] Such as search engines or unmonitored corpus based resources. See also Zanettin 2012, 13 for considerations in terms of reliability among different types of corpora.

[19] https://corpus.byu.edu/.

[20] http://skell.sketchengine.co.uk/run.cgi/skell.

replace initial or final characters of a word (e.g. type *divers** and get results including *diverse*, *diversity* and *diversities*); (b) restrict the query through part-of-speech tags (e.g. search any adjective followed by *eyes* or any adverb preceding *unsatisfactory* – as I did to choose, a few lines above, between *extremely* and *highly* unsatisfactory [21] –); (c) compare or search different sections or different corpora (e.g. compare English and American English); (d) read and sort the concordance lines to observe specific patterns (e.g. by sorting the expression *can't cut* to the right we can reveal that *can't cut it* can be used as an alternative to *can't cut the mustard*); (e) query words only when they co-occur in a span of X words (e.g. *cut* only when used with *mustard*), (f) generate collocation lists (e.g. the collocates of *dog's day*); (g) for more expert users generate wordlists or cluster lists of different corpora and compare them to reveal diverse linguistic patterns [22]. If some of these functions, and others not mentioned here, are probably still perceived as too technical and challenging, others, like the use of wildcards, the use of part-of-speech tags and collocations or concordance lists need little training and learners soon realize that by using corpora they can observe, explore and produce results that are often better tailored to their specific needs.

Summing up, corpora should be introduced in learning processes because they are now easily available online, in much the same way as other more familiar resources. They offer types of queries that are not available with other resources and their putative technical complexity is mitigated by online support web pages and the learners' increased familiarity with online searches. The use of corpora allows for more detailed exploration of linguistic patterns and contexts and this appears particularly relevant when discussing the need for autonomous access to authentic online materials by learners in blended/online courses and self-access language centres. In such circumstances, students are often working on their own and it is virtually impossible to foresee the contexts of use and language problems they are going to come across.

It must be said that in traditional DDL or corpus based learning studies, corpora are usually considered the only tool for exploration and other online resources are ignored or criticized. This may well be rather limiting as the dif-

---

[21] The query string *_r* unsatisfactory* was used in the NOW corpus site: https://corpus.byu.edu/now/.

[22] For a more detailed guide see: https://corpus.byu.edu/queries.asp and https://www.sketchengine.eu/quick-start-guide/.

ferent online tools can be used in combination to achieve different purposes: learners may well find instant, more accessible and useful information by using Google images, Wikipedia, Reverso Context or other online resources, as in the example presented in section 2 above and these findings can be profitably used as a starting point for corpus queries.

4. A GUIDE TO AN AUTONOMOUS EXPLORATION OF LANGUAGE IN CONTEXT: THE 'CHECK ONLINE ROUTINE'

As suggested in previous studies (Zanca 2014; 2017), meaningful linguistic explorations of words and patterns using online tools and corpora can be introduced by asking our students to engage in traditional activities, such as reading comprehension, writing or translation tasks. Generally speaking asynchronous activities of this kind are particularly appropriate because they are normally focused on accuracy rather than on more general communicative outcomes and allow time to go online, to explore contextual features and to consider alternative linguistic options. Nevertheless, as mentioned previously, some of the outcomes of these explorations using online resources are likely to be inaccurate and/or misleading, so 'explorers' should learn how to verify their accuracy and check if the context of use is appropriate or not. The following simple verification procedure could be adopted (see Zanca 2013, 328):

a) Define linguistic problem / *research question* (e.g. translate *agency work*[23] in Italian.

b) Use online resources to obtain possible solutions (e.g. *lavoro di agenzia* in google translate).

c) Check solutions in context by observing them in relevant corpora or reliable online resources (e.g. comparing *agency work* and *lavoro di agenzia* in English and Italian newspaper articles using Google News or online corpora[24]: the latter expression is almost never used in Italian newspapers).

---

[23] From a text used for a translation activity: "I've been a chef for eight years and have had to walk out of two jobs. My last workplace, before I switched to agency work, was a bar, restaurant and club", https://www.theguardian.com/lifeandstyle/2017/mar/18/manager-selling-coke-staff-restaurants-tips-pay-chefs.

[24] For instance using the *BYU NOW corpus* and the *La repubblica online corpus*, https://www.english-corpora.org/now/; https://corpora.dipintra.it/public/run.cgi/first_form.

d) Possible emergence of alternative solutions (e.g. *agency work* collocates with and is often preceded by *temporary*, so a possible translation could be *lavoro temporaneo* or *lavoro interinale*).

e) Check these new solutions going back to point b of the routine (in the Italian corpus and Google news *interinale* is frequently used in similar contexts).

This routine fosters the exploration of language in context and brings to the fore different possible usages of words. For example, by browsing through the list of collocates and concordances of *agency work* our learners can observe a different use of the expression when preceded by nouns such as *estate* or *environmental*.

A more detailed example of an exploration activity, the translation of the headline of a newspaper article, might clarify how this routine can be applied and lead to an effective, autonomous and unconstrained exploration of online resources. The headline is: "I'll happily take the credit for a dog day afternoon" [25]. The subhead outlines the main theme of the article: "A few weeks ago I took my dog to the dog beach. The public beaches in Southern California are beautiful, but they are also highly-regulated…", however the interpretation and translation of the second part of the headline still creates some problems. When I asked my students [26] to translate it, some initially opted for the solution 'un pomeriggio a spasso con il cane' (An afternoon spent walking my dog). But this did not match the first part of the headline, why should one "happily take the credit" for something as ordinary as walking one's dog? The first part of the sentence seems to imply a negative, contrasting ensuing statement. Of course they all knew the meaning of the individual words (I, happily, take, credit, dog, day, afternoon), the real meaning had to depend on their combination in a pattern which matched the priming of native speakers, but not that of learners. The course explicitly asks the students to use online resources to accomplish the writing or translation tasks and the first resources normally used are online dictionaries. Learners can easily find *prendersi il merito* as a credible equivalent for *take the credit*, but when they look for *dog day* the results are less straightforward. Most online dictionaries do not provide any solutions for the two words together, others show some results,

---

[25] https://www.thenational.ae/i-ll-happily-take-the-credit-for-a-dog-day-after-noon-1.241348.

[26] I used this text with first-year Italian Literature university students with a mixed level of English, from B1 to B2.

but they are not very clearly relevant to our context. The *Cambridge Online Dictionary* redirects users to the expression *every dog has its days*, which is not very useful in our context: "'every dog has its day': saying; said to emphasize that everyone is successful or happy at some time in their life".

The Dictionary.com site offers a more interesting solution, which to some extent corresponds to the Italian negative expression *da cani*:

> **dog days** plural noun
> - the sultry part of the summer, supposed to occur during the period that Sirius, the Dog Star, rises at the same time as the sun: now often reckoned from July 3 to August 11
> - a period marked by lethargy, inactivity, or indolence.

We get something similar from the *Online Merriam-Webster Dictionary*:

> **Definition of dog day**
> **1 a:** **dog days** *plural*: the period between early July and early September when the hot sultry weather of summer usually occurs
> **b:** a day in dog days: a hot sultry day
> **2 dog days** *plural*: a period marked by dull lack of progress // *the dog days* following any major upheaval

We infer that *dog days* are probably something negative that could be related to the Italian *giornate da cani*, but the British Online English dictionary does not seem to acknowledge this interpretation and the learners were still doubtful. And what about afternoon? In this case Google can be revealing: if we search for *a dog day*, amongst the first five results we get one definition from Wikipedia which bears a resemblance to the *Dictionary.com* one [27], a reference to a TV programme, a third one which advertises dog services and two references to *a dog day afternoon* (see *Fig. 1*).

We can even see an indication of the Italian title of the 1975 American movie: *Quel pomeriggio di un giorno da cani*: the cultural connection with the movie is presumably something native speakers are aware of and something the writer intentionally refers to. Our learners could gather even more evidence by searching for *dog day afternoon* in Google images (see *Fig. 2*) [28].

---

[27] Results of Google searches are different for different users and times. These results were obtained on May 2017 and are very different at the time of writing this paper: this usually has little impact on the efficacy of the autonomous discovery process, but should be considered in the case of previously prepared activities.

[28] The query of *"dog day"* in Google images offers very different results, all relating to National Dog Days.

Conversely, the query of just *dog day* offers very different results in Google images, showing that the expression might also bring to the mind of native speakers another context of use: National Dog Day events. It is possible that the author of the headline intentionally played with the priming of his readers and these two cultural suggestions in order to create a contrast between positive and negative references (see *Fig. 3*).

Finally, as the routine used in the course suggests the use of more reliable linguistic tools, the learners also searched the expression in online corpora. In the Corpus of Contemporary American English [29], for instance, *dog day* appears 77 times and it collocates 55 times with *afternoon*, 4 times with *care*, 3 times with *National* and 3 times with *August*. These results suggest that the expression creates a pattern linked with different priming dimensions and that the association with *"a hot sultry day"* found in dictionaries is by no means the predominant one. A brief analysis of the concordances confirms the strong association of *dog day* with *afternoon* and the 1975 movie, something the learners would not have easily uncovered using traditional reference tools (see *Fig. 4*). By means of just a few clicks, the learners were able to access and discover a whole new aspect of the original text suggesting that the writer played with the reader's cultural familiarity with the celebrated movie and the less frequent reference to National dog days and dog care activities. By adopting the suggested procedure and by using additional online resources in order to verify the initial interpretation offered by dictionaries, learners were led to new and unexpected associations. Different tools revealed different information and only their combined exploration allowed learners to share some of the multifaceted priming of native speakers. Their translation options were now very different: they could refer directly to the Italian title of the movie with 'Mi prendo volentieri il merito di un pomeriggio di un giorno da cani' or more creatively with something like: 'Un pomeriggio o un giorno da cani? Me ne assumo volentieri la responsabilità'. What is crucial here is not the final translation, but the exploration, the priming process the learners have gone through and the opportunity to experience a deeper awareness of the foreign language.

The *dog day* example of language exploration is just one brief illustration of what learners can and should be encouraged to do. Other similar activities are described in Zanca 2013, 2014 and 2017.
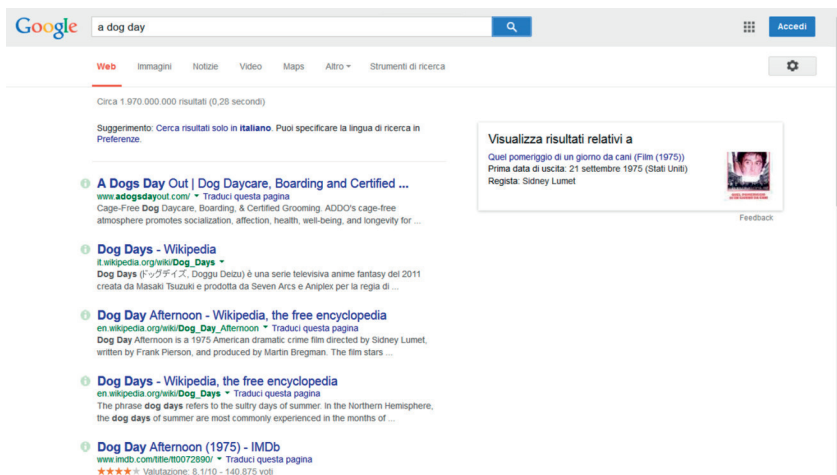
---

[29] https://corpus.byu.edu/coca/.

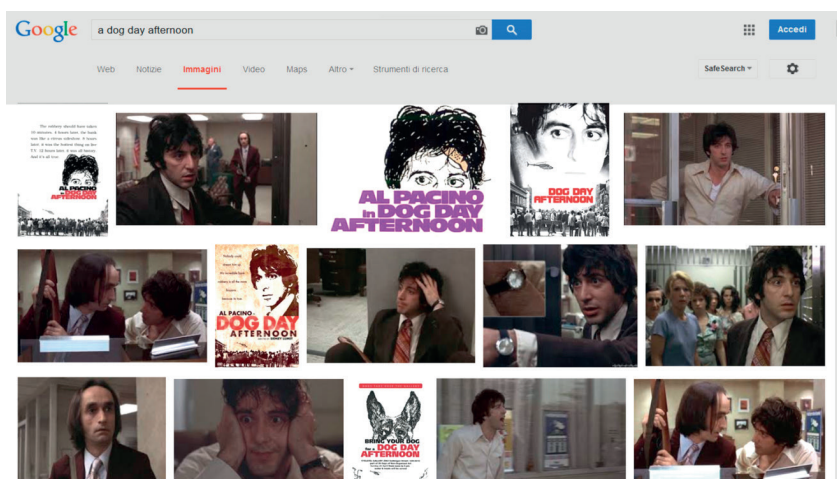*Figure 1. – Results for 'dog day' in Google.*

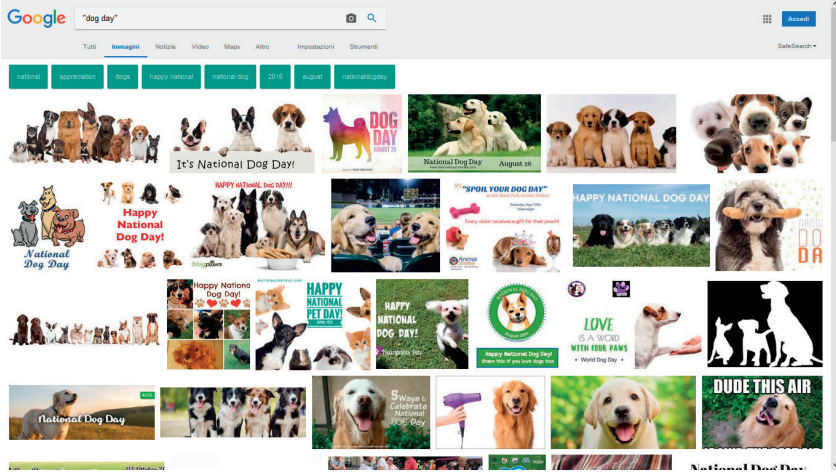*Figure 2. – Results for 'dog day afternoon' in Google Images.*

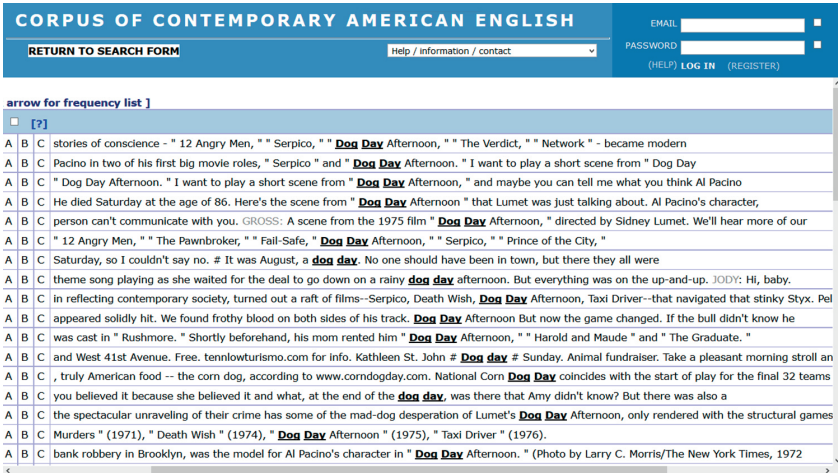*Figure 3. – Results for 'dog day' in Google images.*

*Figure 4. – Initial results for dog days in the Corpus
of Contemporary American English.*

5. Conclusions

Assuming that one of the aims of contemporary self-access centres and blended/online courses activities is to train learners to use the countless authentic materials and resources available online in a profitable manner, this paper argues that this should include the exploration of linguistic patterns and contextual features. The examples and experiences described suggest that our learners, who are today more familiar with online queries, can, in just a few minutes, autonomously access linguistic and contextual features that might compensate for their lower lexical priming and lack of cultural background in the foreign language. An advanced use of online resources combined with the use of corpora and training provided by a class teacher, or by an online or self-access tutor, can convert the learner into a language researcher. This study maintains that the opportunity to explore, carry out research and verify results against authentic, reliable data, enhances the learner's confidence in autonomously dealing with problematic elements in the foreign language and culture and grants her/him the 'tools of the trade' to independently accelerate her/his lexical priming. The introduction of a training process in language centres along similar methodological guidelines requires more research, but the practical indications, examples and suggestions mentioned in this study might become an integral part of our centres' activities or constitute the content matter of specific training courses.

References

Anthony, L. 2019. AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. http://www.laurenceanthony.net/software.

Boulton, Alex. 2010. "Empirical research in data-driven learning – a summary". https://nancy-universite.academia.edu/AlexBoulton.

Chinnery, George M. 2008. "You've got some GALL: Google assisted language learning". *Language Learning & Technology* 12 (1): 3-11.

Choi, Jessie. 2017. "The metamorphosis of a self-access centre in Hong Kong: From theory to practice (a case study)". *Studies in Self-Access Learning Journal* 8 (1): 23-33.

Choi, Yunjeong, Dongbo Zhang, Chin-Hsi Lin, and Yining Zhang. 2018. "Self-regulated learning of vocabulary in English as a Foreign Language". *The Asian EFL Journal Quarterly* 20 (1): 52-82.

Conroy, Mark A. 2010. "Internet tools for language learning: University students taking control of their writing". *Australasian Journal of Educational Technology* 26 (6): 861-82.

Eu, Jinseung. 2017. "Patterns of Google use in language reference and learning: A user survey". *Journal of Computers in Education* 4 (4): 419-39.

Geiller, Luc. 2014. "How EFL students can use Google to correct their 'untreatable' written errors". *European Association for Computer Assisted Language Learning* 22 (2): 26-45.

Granger, Sylviane. 2011. "From phraseology to pedagogy: Challenges and prospects". In *The Phraseological View of Language: A Tribute to John Sinclair*. Ed. by Tomas Herbst, Susan Schüller, and Peter Uhrig, 123-46. Berlin - New York: Mouton de Gruyter.

Higgins, John, and Tim Johns. 1984. *Computers in Language Learning*. Collin: London.

Hoey, Michael. 2005. *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hoey, Michael. 2009. "Lexical priming". http://www.macmillandictionaries.com/MED-Magazine/January2009/52-LA-LexicalPriming-Print.htm.

Johns, Tim. F. 1991. "Should you be persuaded: Two examples of data-driven learning". In *Classroom Concordancing*. Ed. by Tim F. Johns and P. King. Birmingham: ELR.

McEnery, Tony, and Richard Xiao. 2010. "What corpora can offer in language teaching and learning". In *Handbook of Research in Second Language Teaching and Learning*. Ed. by E. Hinkel, vol. 2, 364-80. London - New York: Routledge.

Meunier, Fanny. 2011. "Corpus linguistics and second/foreign language learning: exploring multiple paths". *Revista Brasileira de Linguística Aplicada* 11 (2): 459-77.

Milton, John. 2006. "Resource-rich Web-based feedback: Helping learners become independent writers". In *Feedback in Second Language Writing*. Ed. by Ken Hyland and Fiona Hyland, 123-39. New York: Cambridge University Press.

Partington, Alan. 2008. *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam: John Benjamins.

Römer, Ute. 2008. "Corpora and language teaching". In *Corpus Linguistics*, vol. 1: *An International Handbook*. Ed. by Anke Lüdeling and Merja Kytö, 112-30. Berlin: Mouton de Gruyter.

Sha, Gouquan. 2010. "Using Google as a super corpus to drive written language learning: A comparison with the British national corpus". *Computer Assisted Language Learning* 23 (5): 377-93.

Shei, Chi-Chiang. 2008a. "Discovering the hidden treasure on the Internet: Using Google to uncover the veil of phraseology". *Computer Assisted Language Learning* 21 (1): 67-85.

Shei, Chi-Chiang. 2008b. "Web as corpus, Google, and TESOL: A new trilogy". *Taiwan Journal of TESOL* 5 (2): 1-28.

Tribble, Chris, and Glyn Jone. 1990. *Concordances in the Classroom*. London: Longman.

Wu, Shaoqun, Margaret Franken, and Ian H. Witten. 2009. "Refining the use of web (web search) as a language teaching and learning resource". *Computer Assisted Language Learning* 22 (3): 249-68.

Zanca, Cesare. 2013. "Online learning and data driven learning in translation training and language teaching". In *Creativity and Innovation in Language Education*. Ed. by Carmen Argondizzo, 317-34. Bern - New York: Peter Lang.

Zanca, Cesare. 2014. "Developing translation strategies and cultural awareness using corpora and the web". http://lodel.irevues.inist.fr/tralogy/index. php?id=116&format=print.

Zanca, Cesare. 2017. "When the going gets tough… the tough get googling". In *Digital Resources, Creativity and Innovative Methodologies in Language Teaching and Learning*. Ed. by Adriana T. Damascelli, 111-25. Cambridge: Cambridge Scholars Publishing.

Zanca, Cesare. 2018. "Corpora, Google e roba simile: per quale ragione gli studenti di una lingua straniera dovrebbero perderci tempo?". In *Translation and Interpreting for Language Learners (TAIL). Lessons in Honour of Guy Aston, Anna Ciliberti, Daniela Zorzi*. Ed. by Laurie Anderson, Laura Gavioli, and Federico Zanettin, 201-20. Milano: AItLA.

Zanettin, Federico. 2001. "Swimming in words: Corpora, translation, and language learning". In *Learning with Corpora*. Ed. by Guy Aston, 177-97. Bologna: CLUEB.

Zanettin, Federico. 2002. "Corpora in translation practice". In *Language Resources for Translation Work and Research, LREC 2002. Workshop Proceedings*. Ed. by Elia Yueste-Rodrigo, 10-14. Las Palmas de Gran Canaria.

ABSTRACT

Unlike a few decades ago, using our phones, tablets, phablets or computers, all sorts of foreign language authentic materials are now easily available and accessible outside our language centres. Nevertheless, learners might

find it difficult to select what is more effective for their learning process and be daunted by some complex features of naturally occurring language. This paper draws on previous studies and personal teaching experience to suggest that, in order to fully exploit these resources, language centres should aim at helping learners increase their ability in dealing with online authentic materials inside and outside the centre's premises. In this perspective, they might consider the introduction of a relatively new approach based on corpora and online resources to enhance the learner's autonomy and confidence when dealing with online authentic and unfiltered materials in a foreign language.