



HAL
open science

A didactic comparison of online French-English lexical resources

Laura M. Hartwell

► **To cite this version:**

Laura M. Hartwell. A didactic comparison of online French-English lexical resources. *Etudes en didactique des langues*, 2020, *Transmission & vecteurs/vectors*, 34, pp.7-22. hal-04011925

HAL Id: hal-04011925

<https://ut3-toulouseinp.hal.science/hal-04011925v1>

Submitted on 2 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A didactic comparison of online French-English lexical resources

Laura M. HARTWELL

Professeure des Universités

LAIRDIL, Université de Toulouse 1 Capitole



Introduction

Recent technologies allow dictionary and concordancer publishers to transmit large amounts of lexical information to anyone with a cellphone or a computer and Internet access. Literally, *transmission* refers to the “act, process, or instance of transmitting”, that is to say “[sending or conveying] from one person or place to another” (Merriam-Webster Inc., 2020), such as through radio waves. A primary synonym of *transfer* is *communicate*, an example in context being: “technology allows data to be transmitted by cellular phones” (*ibid*). Thus, the sender first controls the nature and form of the data transmitted. However, the actual reception may be modified by the user’s practices.

New forms of transmission of lexical information modify user practices that may or may not be understood by the teaching community. Furthermore, a range of inequalities among these online lexical resources complicate their use in a teaching and learning environment. Concerning dictionaries, Rundell (2012) highlights the gap of attention to new “look-up” practices, especially in English for Academic Purposes (EAP) or English for Specific Purposes (ESP) communities.

The commonest forms of look-up (meaning, spelling, pronunciation), have already become subsumed into the larger enterprise of “search”, where the starting point is typically Google, not a dictionary. Even this can be problematic, as many online dictionaries marry cutting-edge technology with horribly outdated dictionaries. Meanwhile, far more needs to be done to meet the receptive and productive needs of users in EAP or ESP environments (Rundell, 2012: third paragraph).

Dictionaries have a long history within language teaching and learning. To the contrary, concordancers have only recently been introduced into teaching and learning environments, so that teachers may be unaware of the possibilities that they offer in parallel to dictionaries.

A concordancer is a piece of software, either installed on a computer or accessed through a website, which can be used to search, access and analyse language from a corpus. They can be particularly useful in exploring the relationships between words and can give us very accurate information about the way language is authentically used (Peachey, 2020).

Popular easily-accessed resources, such as *WordReference* and *Linguee*, are already being used by many university students for foreign language comprehension or expression. However, the plethora of available options or the quality of the language of these resources vary. The inquiry here focuses on the actual characteristics of these readily available resources in order to help language teachers who wish to encourage better practices of consultation.

The present comparison of online lexical resources is inspired by Francis Grossmann's presentation, "*Faut-il rappeler cette évidence? Usages des formules de l'évidence dans l'écrit scientifique*" [Must one recall this piece of evidence? Uses of formulaic expressions of evidence in academic writing] (2016). Evidence is a key aspect of academic, scientific, and legal domains and discourse. The translation of the French *évidence* to English is problematic because of the many idiomatic expressions in French as well as because of slightly different meanings of the English "evidence". In order to understand these cultural and contextual differences, which are typical of the difficulties related to acquiring vocabulary in a foreign language, I draw upon specialized corpora as a means to delve into this lexical complexity and the "transmission" or lack of transmission of meaning, via online lexical resources.

Transmission and reception

If one evokes a possible lack of constructive transmission, the question of actual reception by users becomes central. Boulton & Tyne (2014) remind us of the traditional use of resources such as dictionaries, but which take on new computerized options. New technologies of information and communication offer vast possibilities that may or may not be exploited or exploitable in a teaching and learning context.

L'enseignement des langues va profiter particulièrement de ce que les technologies permettent de stocker, d'organiser, de présenter les données: ainsi, les grammaires, les dictionnaires et les manuels constituent des outils *a priori* assez anciens (qui apparaissent essentiellement à partir du moment où l'imprimerie rend leur reproduction possible), mais désormais indissociables de la plupart des programmes d'apprentissages ou d'enseignement de langues [...] On se heurte donc à un problème: les technologies existent dans la société et se retrouvent ensuite appliquées au monde de l'éducation, mais la question qui se pose est de savoir ce qu'on peut faire concrètement avec ces technologies (*ibid.*: 30-31).

Much attention has been paid to accessing vocabulary, such as through word lists, specialised dictionaries or corpora. Coxhead (2013) posits that vocabulary knowledge is essential for learners of English for specific purposes (ESP). Thus, classroom time should be dedicated to language needs, including key concepts and the language of the field, as understanding and using "special purposes vocabulary" to engage in disciplinary knowledge is necessary to become "full-fledged members of a particular community". For ESP learners, this lexical acquisition may represent an "extremely large learning task" (Coxhead, 2013: 116). Dictionaries are traditionally a printed source for language learning and teaching, many of which have given rise to a parallel online form. They are central to language learning, especially in language

courses for specific purposes because of the richness and complexity of domain-related vocabulary and the autonomy they provide to learners.

Building a dictionary

A brief reminder of the construction of dictionaries, and notably the use of corpora in doing so, will help to understand how editors gather and transmit information. A dictionary can be defined as

a reference source in print or electronic form containing words usually alphabetically arranged along with information about their forms, pronunciations, functions, etymologies, meanings, and syntactic and idiomatic uses (Merriam-Webster, Inc., 2020).

Many dictionaries and other online resources offering lexical information offer a plethora of information, which may be confusing for some users.

Historically, corpora and corpora studies have contributed to generating this information, with varying degrees of accuracy, depending notably on the size and quality of the corpus. Many dictionary publishers continue to rely upon corpus data: the *Pearson Longman* dictionaries are compiled from the *Longman Corpus Network* (Pearson ELT, 2019), the *Macmillan* dictionaries from data of the *World English Corpus* (Springer Nature Limited, 2019), and those of *Oxford* based on the *Oxford English Corpus* (Oxford University Press, 2019). Thus, corpus data has been a historical, but hidden foundation to constituting dictionaries. Technical advances allow corpus data to be accessible as transmitted by dictionary sources.

One major corpus source for many language reference materials is the *Collins Birmingham University International Language Database (Cobuild)* (Collins Language, 2020). The *Cobuild* corpus subsequently gave rise to both the *Bank of English* corpus (4.5 billion words) of general English and numerous dictionaries, including the *Collins Cobuild English Language Dictionary*, now freely available online (Collins, 2020). According to Stubbs (2007), the original Cobuild project was of little use to learners because the examples were too complex, but its methodology and approach were essential to the birth of the modern approach to lexicography. The capacity to draw upon corpora has developed in parallel with the expansion of technological capacities. However, this data may not be identical across resources, depending upon choices and means of transmission.

For example, the technology of Sketch Engine, now integrated into the *WordBanks Online Corpus* used by the *Collins Dictionary*, was initially developed to identify authentic examples from corpora for an electronic version of the *English Macmillan Dictionary*. This technology allows one to target examples of general English that illustrate frequent patterns, but retain adequate readability, that is, without the use of complicated terminology, which would hinder comprehension of brief dictionary entries (Kilgarriff *et al.*, 2008). The technology of Sketch Engine is now used to write other dictionaries, including *Dictionary.Com* and *Le Robert* (Lexical Computing, 2020). Thus, general dictionaries may miss uses or meanings that students or teachers of specialized English need for advanced academic or professional contexts. Students and teachers should be aware of the variations

between the different dictionaries and concordancers in order to best serve their learning or teaching objectives.

Some bilingual resources may offer translations that are reduced to little more than a one-word equivalent, such as found on *Google Translate*. *Google Translate* is frequently used by teachers and students alike as it is easily available on the Internet and offers accessible information in a matter of seconds. For example, a query of the French term *évidence* on *Google Translate* (2020) offers the English translation “evidence” and a sound recording of both words. Below this are listed synonyms in French (*preuve, évidence, témoignage, déposition, signe, marque*) and the unique English synonym “obviousness”. The latter is a legal construction related to patent law, any obvious object or idea being unlawful to patent. This specific legal usage is of very little use for the vast majority of English language learners, who may actually be distracted from other, more suitable, translations of the term.

Further tests reveal the oversimplification of *Google Translate*. A query of the English term “assets” displays the predominant translation *atouts*, tagged with the label “Community verified”. A more appropriate translation for economic purposes, the singular *actif* in French, is mentioned below this as a second possible translation, but without any information allowing the reader to distinguish between the two. For the English speaker attempting to translate “Master”, with a capital M as found when referring to university degrees, *Google Translate* proposes *maîtresse* and *maître*, with a secondary translation *le Jésus-Christ* (sic). This example demonstrates the unreliability of *Google Translate*. This oversimplification and a reference to Christian tradition is unlikely to be of use to students in advanced academic settings.

Some sources, such as *Google Translate*, rely upon frequency of occurrences found on the Internet. This leads to an over-representation of certain unacknowledged economic or other interests, such as those of the pornography industry or political agendas. To the contrary, sources based on corpora, such as *TradooIT*, are representative of actual discourse, as represented in the established corpus. Other resources propose, as this article demonstrates, more complex information, but which may require a more detailed consultation to be of use to language teachers and advanced learners. Helping students to become aware of options and the need to consult beyond the first translation are essential steps to building life-long autonomous dictionary-related skills.

Methodology

In order to explore the varying characteristics of these resources for language learning, this paper focuses on four of the most popular open-access bilingual resources: *Linguee*, *Reverso*, *WordReference*, and *TradooIT*. The *Linguee* website, developed by the German DeepL company, describes itself as “your bilingual dictionary”, offering some one billion translations between English and French, as well as translations of other languages (DeepL, 2020). The *WordReference* English-French dictionary incorporates the *Collins Dictionary* and their own evolving dictionary of some 100,000 words and expressions per language, augmented by a community forum (Kellogg, 2020). The multilingual website *Reverso*, offering several

options, such as “Translation” or “Context”, includes the *ReversoDictionary*, which builds upon the 2005 *Collins Dictionary* and contributions from the *Reverso* community (Reverso, 2020). Finally, *TradooIT*, the only resource of this study self-labeled as a *bilingual concordancer*, also offers a translation memory and a term bank (Okidoo Inc., 2019). Although not a dictionary, *TradooIT* is included in this study in order to offer a more complete range of freely available bilingual resources.

To compare these resources and help teachers understand their differences, I analyze their translations into English of the French word *évidence*. This term was chosen for its frequent occurrence in fixed expressions (*i.e. mettre en évidence* or *de toute évidence*) and its range of possible translations, despite the seemingly transparent English word “evidence”. Second, the English term “evidence” will be analyzed as it appears in two specialized corpora, the *Scientext* corpus of published science texts in English (Hartwell, 2013; Lidilem, n.d.) and the *United States Supreme Court Opinions Corpus* (130 million words) (Davies, 2018; United States Supreme Court of the United States, 2018). Specifically, the construction [ADJ + evidence] reveals contrasting visions of this vital notion intrinsic to both specialized domains. This analysis is important for language teachers in specialized domains, for whom the exact meaning is critical to understanding. These collocations are then compared with those of the three dictionary-type resources in order to evaluate the range of citations and their contextualization, which may or may not be useful to language teaching and learning. For this, I begin by reviewing some similarities and differences between the French term *évidence* and the English “evidence”.

The French *évidence* and the English “evidence”

In order to appreciate the nature of the differences between these cognates (that is to say, here words related by descent from Latin), it is necessary to return to the concept or essential qualities of the term (Roche 2007), notably the Late Latin *evidentia* “proof” and the Latin *evidens* “obvious” or “apparent”. As Grossmann (2016) points out, the French term *évidence* has evolved over the centuries, stemming from a philosophical meaning and moving to a more rhetorical one, while preserving the visual nature of the Latin etymology. From an analysis of the *Scientext corpus*, Grossmann notes the existence of adverbial phrases, such as *de toute évidence* (literally “of all evidence”) and verbal phrases, such as *mettre en évidence* (literally “put in evidence”). More importantly, his study of the word’s meaning identifies the existence of linguistic routines related to the three discursive motives of the term *évidence*; false evidence (1), as expressed by:

- 1) est loin d’être une évidence pour X,
[is far from being self-evident for X];

an disciplinary or empirical evidence (2 & 3) formulated by:

- 2) s’appuie sur l’évidence,
[relies on the evidence];
- 3) se fonde sur une évidence,
[is based on evidence/proof];

or a “forgotten” one (4), such as in the expression:

- 4) nier l'évidence,
[to deny the evidence].

As Grossmann (2016) reminds us, one translation of *évidence* is the English “obviousness”. While the English noun “evidence” also exists, it can be defined as an “outward sign: indication; something that furnishes proof: testimony” (Merriam-Webster Inc., 2020). In a legal context, the English expression “to give evidence” means “to talk and answer questions about something especially in a court of law while formally promising that what one is saying is true, to testify” (*ibid.*). Šeškauskienė & Stepančuk (2014) find that, within three oral arguments of the United States Supreme Court, two thirds of the metaphors refer to law as an OBJECT, notably with the nouns, “evidence”, “testimony”, and “law”. They consider “evidence” as an abstract notion that is objectified, such as in the metaphorical “to give evidence”, mirrored by the transitive verb “evidence”, meaning, “to offer evidence of” (Merriam-Webster Inc., 2020).

Hence, the English meaning tends to be concrete and object-oriented similar to the Late Latin *evidentia*. In comparison, the rhetorical French construct, closer to the Latin *evidens* “obvious” or “apparent”, is a visually-oriented reference to the nature of an object or a concept. These cognates can be considered “false friends” as they cannot be used interchangeably, although they appear similar.

Principal translations

The four resources discussed in this article present slight variations across their principal translations of *évidence*. All of the resources, except *TradoolT*, which is a concordancer instead of a dictionary, begin by noting that *évidence* is a feminine noun and incorporating an audio recording of its pronunciation, highlighting the dictionary-like quality of these three resources.

Linguee then proposes the English translation “obviousness”, followed by three “less common” alternatives: “commonplace”, “blatancy”, “patency”. However, none of these four terms can be found in the 30 sentences displayed in the “external sources not reviewed” of the French word *évidence*. In fact, while the entries for the English words “commonplace”, “blatancy”, and “patency” mention *évidence* as a translation, none of the examples include these three terms. Many of the sources of “obviousness” are from the International Association for the Protection of Intellectual Property. Thus, it appears that *Linguee* begins by focusing on translations from English to French of mostly legal documents containing the term “obviousness”, resulting in a misleading translation for students searching for meanings outside of patent law.

Although both *Reverso* and *WordReference* rely upon the *Collins Dictionary*'s translations of *évidence* as “evidence” or “proof” (Collins, 2020), their entries highlight slightly different meanings of the term. *Reverso* proposes two blocks of text for *évidence*. The first block focuses on the English translation “obviousness”, which is followed by a series of expressions containing “evidence”, while the second block contains only the translation “obvious fact”. *Reverso* does not offer examples of use, the proposed equivalent being immediately followed by expressions containing

évidence, which will be discussed in the following section. For the principal translations of *évidence*, *WordReference* suggests two meanings: *chose flagrante* [flagrant thing] (translated as “evidence”, “obviousness”, and “demonstrability”) and *chose certaine* [certain thing] (translated as “foregone conclusion”, “self-evident fact”, and “obvious”). Thus, the entries for *Reverso* and *WordReference* are similar by the presence of “obviousness” in the first principal translation and the words “obvious” and “fact” in the secondary translation.

Finally, the fourth resource discussed here, *TradooIT*, lists ten “grouped translations” by order of frequency in their corpus, beginning with “obviously (7501)”, “clearly (4023)”, “highlight (2692)”, and “evidence (1089)”. The translation by “obviousness” is not one of the ten grouped translations, but is mentioned as employed within intellectual and theoretical law documents. Contrary to the dictionary sources, *TradooIT* offers neither a literal translation of the word, nor an audio representation as a standard bilingual dictionary would do, but focuses on translations found in the integrated corpora. The *TradooIT* website has options for selecting the corpora, which allows the user to select a relevant genre or domain. Thus, these four resources contrast in their first principal translation, which appears dependent on the sources they have relied upon, notably an over-reliance on documents related to patent law to the detriment of other contexts.

While the French noun *évidence* is quite frequent, the English noun “obviousness” is not at all. In fact, while the *British National Corpus* (Center for Translation Studies, 2013), rich with some 100 million lexical units, contains only 21 occurrences of the word “obviousness”, the English part of the *Sciencetext* corpus of scientific texts (Lidilem, n.d.), composed of 35 million lexical units, contains none. In other words, the term “obviousness” exists, but is rarely used by Anglophones. This reliance upon the often-legal “obviousness” produces an overall underrepresentation of other possible translations of *évidence*, thereby substantially reducing their usefulness for the vast majority of language learners.

Expressions containing *évidence*

These resources also propose multi-word expressions containing the term *évidence*, followed by possible translations. Stubbs (2007) previously drew attention to the importance of the specific meanings of expressions. A word found in a particular expression may have a different meaning than when used alone. The presence of these expressions is thus an indicator of the range of meanings transmitted to the user, if they take the time to read them.

For *Linguee*, twelve expressions are labeled *examples*. The thirteen expressions of *WordReference* are found under the French “*formes composées*” and are followed by a much longer list compiled by members of the *WordReference* forum. *Reverso* lists eleven expressions, followed by a collaborative community dictionary that largely repeats and expands upon the given expressions. *TradooIT* does not explicitly list expressions.

As can be seen in Appendix 1, there is variation across the quantity of both the French expressions and the proposed English equivalents. Only four of the 22 expressions (*à l'évidence*, *mettre en évidence*, *de toute évidence*, *en évidence*) are found in the

three online dictionary resources, but none of the first proposed English equivalents are strictly identical due to lexical and grammatical variation. For example, *en évidence* is translated as “conspicuously”, “conspicuous”, or “in evidence”. Another five expressions associating *évidence* are found only in the *ReversoDictionary* and *WordReference* resources, but with contrasting translations.

WordReference proposes a rather curious example (5) of *évidence* in context and its translation:

5) *L'évidence de cette preuve ne me saute pas aux yeux.*

The evidence of the proof isn't jumping out at me.

A search on Google of the segment *l'évidence de cette preuve* produces twenty results (excluding those referring to *WordReference* or language learning flashcards), almost all dating from the 17th to 19th century and related to biblical issues (Google, 2020). A Google search of the English of “the evidence of the proof” reveals only five results (excluding expressions with “proof sheet” or “proof of claim”) (Google, 2020). In other words, the primary example of *évidence*/“evidence” in *WordReference* is of limited pertinence to modern day language use. Moreover, this jaded quote, “the evidence of the proof”, is incongruently matched with the more modern fixed expression *ne saute pas aux yeux*/“isn't jumping out at me”. Thus, while the example might initially appear authentic to the user, its existence as an authentic example of language use appears doubtful. However, online language learning supports have incorporated the quote, such as in the online vocabulary *Quizlet* (Mohit, 2015).

The frequency or rarity of use, related to the relevant importance of a term, is thus often an important pedagogical issue. However, this consideration is not always taken into account by online dictionaries. Students should be thus encouraged to consult multiple sources, especially for expressions or specific terminology that they need for comprehension or expression.

Frequency of expressions

As a concordancer instead of a dictionary, *TradooIT* does not explicitly list frequent expressions employing *évidence*, although they may be part of the aligned corpus examples displayed within the results of a given word. With *TradooIT*, it is also possible to search individually for the frequency of chosen expressions. Drawing upon this, Table 1 lists the occurrence frequencies of selected expressions from Appendix 1 and their most frequent equivalents in English as found in *TradooIT*. For expressions employing a verb, *TradooIT* queries were conducted for all tenses and the results combined. The results confirm that the frequency of the given expressions varies widely, from over 16,000 occurrences of “*de toute évidence*” to only seven occurrences of “*s'imposer comme une évidence*”.

All three resources list “highlight” as a possible translation of *mettre en évidence* (cf. Table 1). *WordReference* proposes the most frequent equivalents, according to *TradooIT* data, of both *à l'évidence* [obviously] and *c'est l'évidence même* [it is obvious].

Other expressions with limited frequency, such as *se mettre en évidence* ($n = 12$) has no single frequent equivalent in English, the translations for this expression ranging from “come to the fore” and “showcase itself”. These results highlight the frequency

of expressions in French containing the term *évidence* and which are often translated using terms from other grammatical categories, be it an adverb such as “obviously” or “clearly” or containing a verb such as “highlight” or “showcase”. In other words, the cognates *évidence*/“evidence” have different meanings according to lexico-grammatical context. This is often a difficult phenomenon for language learners who believe that similarly spelled words will have common meanings and uses.

Expressions listed in TradooIT	Most frequent <i>TradooIT</i> translation	Listed in
<i>de toute évidence</i> - 16,397	obviously - 6,019	<i>Reverso</i>
<i>en évidence</i> - 12,657	highlighted - 3,075	<i>Linguee, Reverso</i>
<i>mettre en évidence</i> - 4,922	highlight - 1,318	All three
<i>à l'évidence</i> - 2,929	obviously - 627	<i>Reverso</i> <i>WordReference</i>
<i>se rendre à l'évidence</i> - 349	face it/the facts - 31	<i>WordReference</i>
<i>c'est l'évidence même</i> - 136	it is obvious - 16	<i>WordReference</i>
<i>nier l'évidence</i> - 86	deny the evidence - 21	<i>Reverso</i>
<i>être en évidence</i> - 24	in plain view - 14	X
<i>se mettre en évidence</i> - 12	come to the fore - 2 showcase itself - 2	X
<i>s'imposer comme une évidence</i> - 7	be self-evident - 2	<i>WordReference</i>

Table 1 – Frequency of occurrence of selected expressions containing *évidence* in TradooIT and their most frequent equivalents in English

[ADJ] + evidence] collocations in context

Thus, as hinted at by the *Linguee* entry for the French-to-English translation of *évidence* discussed *supra*, the [ADJ] + evidence] collocation is particularly productive in English, by both the quantity and the variety of adjectives. Indeed, *Linguee* displays a series of collocations in which an adjective precedes the noun “evidence” (*i.e.* “scientific evidence”) within the entries for the translation from the French *évidence* to the English “evidence”. Collocations are recognized as building blocks to academic discourse (Hartwell, 2013).

An analysis of the entries for the translation from the English “evidence” to the French *évidence* confirms the tendency for “evidence” to be qualified by an adjective in English. *Linguee* proposes the principal translation of the noun “evidence” as the French *preuve*, and less commonly *témoignage*, *indice*, *témoignage*, *indication*, *certificat*, before offering a series of 47 examples of use, of which 34 (72%) are an [ADJ] + evidence] collocation. *WordReference* proposes the principal translations of *indice*, *indication*, *preuve*, *témoignage*, *déposition*, before displaying 67 *formes composées*, of which 45 (67%) are an [ADJ] + evidence] collocation. These figures for adjectives include a few compounds,

such as “illegally obtained evidence” and “blood test evidence”. *Reverso* suggests *preuve* and *témoignage* as the two principal translations of “evidence”. In the main entry for “evidence”, the only collocations suggested are “forensic evidence” and “hearsay evidence”, the other examples being mainly expressions with a verb, such as “to show evidence of”. The main entry is followed by the collaborative dictionary, of which 41 of the 131 (31.3%) entries concern an [ADJ + evidence]. However, many of the entries concern parallel translations, such as “demonstrativeness” or “to come to prominence”.

The concordancer *TradooIT* offers the option of searching for an unknown word using a question mark sign; thus “? evidence” allows one to search for words preceding “evidence”. However, this option did not produce useful results in this case. The simple presence of collocations does not ensure the relevance to language learning. In order to understand if and how these examples of collocation are pertinent, it is necessary to compare them with those found within specialized discourse.

[ADJ + evidence] in corpora

I then examined the frequency of [ADJ + evidence] collocations in two complementary contemporary corpora, one of scientific content and the other of legal content, in order to better define the English nature of “evidence” and to evaluate the examples displayed in the three online dictionaries (Table 2).

First, I consulted the *Scientext* corpus (35 million words), which is composed of 7,564 published texts in medicine and biology dating from 1991 to 2002 (Tutin & Grossmann 2014; Hartwell & Jacques 2014). The *Scientext* corpus identifies 14,030 occurrences of the term “evidence”, of which 5,100 are part of [ADJ + evidence] collocations, accompanied by 440 different adjectives.

Second, a collection of United States Supreme Court opinions is freely available on the website of the official *Supreme Court* website and the *Corpus of United States Supreme Court Opinions* (Davies, 2018). A corpus of similar size and contemporary to the *Scientext* corpus is possible by limiting the search to the years 1990-2010. This 23-million-word corpus of Supreme Court opinions contains 26,253 occurrences of “evidence”, of which 522 different adjectives are found composing 8,632 [ADJ + evidence] collocations.

The analysis of these two corpora demonstrates that although the scientific corpus is larger, it contains some 10,000 fewer occurrences of the word “evidence”. There are also approximately 20% (18.6%, $n = 82$) more different adjectives in the legal corpus. Table 2 displays the most frequent adjectives by corpus and in parentheses the quantity of the same adjective linked to “evidence” in the other corpus. For example, “strong” is found 328 times in collocation with “evidence” in the *Scientext* corpus, but only 151 times in the US Supreme Court corpus. This collocation is listed only in *Linguee*.

The results displayed in Table 2 show little overlap by discipline of the ten most frequent adjectives preceding “evidence”. Only the adjective “direct” is found in both corpora among the ten most frequent collocations with “evidence”: with 172

occurrences in the Scientext corpus and 183 occurrences in the Supreme Court corpus. “Direct evidence” is also mentioned in *WordReference*. For the Scientext corpus of scientific English, two of the most frequent adjectives, “experimental” and “clinical” are completely absent from the collocations in the US Supreme Court corpus. These two adjectives reflect the scientific context, where evidence is often created as part of the scientific process. In parallel, “mitigating” and “relevant” are present only in the corpus of the Supreme Court. These adjectives mirror the consequential nature of evidence, which influences the Court’s decision. Surprisingly few of these collocations are actually found as examples within the online resources, *Linguee* being the most represented with only six of the 20 collocations.

Scientext (USSC corpus)	Listed in	US Supreme Court (Scientext corpus)	Listed in
strong - 328 (SC - 151)	<i>Linguee</i>	mitigating - 1368 (Sci - 0)	X
experimental - 267 (SC - 0)	X	convincing - 441 (Sci - 95)	<i>Linguee</i>
recent - 199 (SC - 8)	<i>Linguee</i>	new - 334 (Sci - 50)	<i>Linguee</i>
further - 187 (SC - 82)	<i>Linguee</i>	other - 314 (Sci - 58)	X
direct - 172 (SC - 183)	<i>WordReference</i>	substantial - 299 (Sci - 60)	<i>Reverso</i>
available - 168 (SC - 58)	X	sufficient - 269 (Sci - 45)	X
little - 161 (SC - 11)	X	relevant - 35 (Sci - 0)	X
good - 146 (SC - 13)	X	additional - 215 (Sci - 99)	X
clinical 120 (SC - 0)	<i>Reverso</i>	historical - 197 (Sci - 5)	<i>Linguee</i>
clear - 118 (SC - 70)	X	direct - 183 (Sci - 172)	<i>WordReference</i>

Table 2 – Top ten most frequent adjectives collocated with “evidence” in the two corpora and their inclusion within the three online dictionaries

The different contexts are also reflected by the synonyms “recent” and “new” (*cf.* Table 2), “recent evidence” being employed in cutting-edge scientific research, while “new evidence”, in comparison to “former evidence”, is often essential in determining the admissibility or the outcome within a legal context. Specifically, the construction [AD] + evidence] reveals contrasting visions of “evidence”, a vital construct intrinsic to both specialized domains, which language learners would benefit from understanding.

The notion of evidence is so different between the scientific and legal domains, that many of the collocations present in one corpus are not found in the other, as listed in Table 3. Certain adjectives, found only within the Scientext sciences corpus (*experimental, clinical, scientific, little, fossil, and genomic*), give witness to the importance of scientific methodology. In contrast, the Supreme Court corpus, containing adjectives

absent within the sciences, such as *mitigating*, *prima facie*, *new*, *competent*, *incriminating* and *exculpatory*, highlight the often unpredictable, but moderating nature of evidence within legal proceedings.

Scientext	Listed in	US Supreme Court	Listed in
experimental - 267	X	mitigating - 1368	X
clinical - 120	X	relevant – 235	X
current - 84	X	exculpatory – 162	<i>WordReference</i>
first - 79	X	circumstantial – 158	<i>WordReference, Reverso</i>
suggestive - 59	X	<i>prima facie</i> – 124	<i>WordReference, Reverso</i>
biochemical - 50	X	incriminating - 92	<i>WordReference, Reverso</i>
genetic - 37	X	persuasive - 78	X
histological - 34	X	probative - 66	X
such - 28	X	forensic - 62	All three
much - 26		overwhelming - 60	<i>Reverso</i>

Table 3 – Adjectives in collocation with “evidence” present in only one corpus

Again, as listed in Table 3, these collocations are scarcely represented in the online lexical dictionaries. Furthermore, only adjectives from the legal domain are displayed in the online dictionaries. The types of examples in context are dependent on the corpora from which they are drawn. It appears that legal or governmental documents are more often represented in the sources exploited for these dictionaries than scientific documents, thereby influencing the types of examples and collocations. Even for non-specialized resources such as these, an absence of documents of scientific or legal content may result in an absence of productive examples that portray the intrinsic qualities of the target term. This lack may weaken their pedagogical value when used to understand or write more specialized documents.

Conclusion

As we have seen, there are several definitions, approaches, and options transmitted via the popular English-French online lexical resources, *Linguee*, *Reverso*, *WordReference*, and *TradooIt*. These sources present contrasting means of processing terms, as reflected in the proposed translations and the examples of the term in context. Teachers may want to examine and compare these resources or specific entries in order to find a resource that transmits more appropriate and accessible information, especially for the needs of learners of specialized languages. Students

will gain in autonomy if they understand the differences these resources offer for them when reading or writing.

It becomes clear that the presence of the target term, here “evidence”, within a multi-word expression may significantly modify meaning, for example, *mettre en évidence*. It is important that language learners be attentive when translating word-by-word, but to take larger sequences into account. This notion is difficult for many learners, who often limit queries to a single word. These resources may be an entryway into understanding these differences. Many students rely upon translation for understanding, but translations should not be confined to a given grammatical category or the literal translation of a given term. Students may need a reminder on how to best search for their particular terms or expressions, especially as related to their discourse community. These practices include: reading the full entry for a word and not simply the first translation, comparing examples across multiple sources, and translating a term back from the target language to the original language.

Furthermore, teachers should be aware of the corpus that constitutes the foundation of these lexical resources. Teachers of specialized language may want to verify if their target discourse community is represented. If not, the specialized meanings or uses may not be available, leaving advanced students unaware that they may find partial or misleading information. In turn, this lack of awareness may lead to poor or partial understanding.

In conclusion, this critical analysis offers insights into the possible obstacles and assets of these frequently consulted online bilingual resources. In class, I encourage students to consult their cellphones with the objective of building positive life-long practices of dictionary or concordance consultation as acquiring vocabulary is an essential aspect to language learning. I ask students to report to the class on which resource they are using and the proposed translations so that they and other students may also become aware and critical of the available resources. The question here is not which resource is better, but which resources transmit dependable information that is receivable by individual language learners and teachers.

Bibliography of resources

(Free resources are indicated by “Available at”, paid resources by “See”).

CENTER FOR TRANSLATION STUDIES. 2013. *IntelliText: British National Corpus*. Available at: <http://corpus.leeds.ac.uk/itweb/htdocs/Query.html#>.

COLLINS. 2020. *Collins English Dictionary*. Available at: <https://www.collinsdictionary.com/dictionary/english>.

COLLINS LANGUAGE. 2020. *Collins Online Corpus Service – WordBanks Online*. See: <http://www.collinslanguage.com/2017/03/16/collins-online-corpus>.

DAVIES, MARK. 2018. *The Corpus of US Supreme Court Opinions*. Available at: <https://corpus.byu.edu/scotus>.

DEEPL COMPANY. 2020. *Linguee English-French Dictionary*. Available at: <https://www.linguee.com/english-french/about#about>.

GOOGLE. 2020. *Google*. Available at: <https://www.google.com>.

GOOGLE TRANSLATE. 2020. Available at: <https://translate.google.fr>.

KELLOGG, MICHAEL. 2020. *About WordReference.com*. Available at: <https://www.wordreference.com/english/AboutUs.aspx?dict=enes>.

LEXICAL COMPUTING CZ S.R.O. 2020. *Sketch Engine*. See: <https://www.Sketchengine.eu/#blue>.

LIDILEM. Not dated. *Scientext Corpus*. Available at: <https://corpora.aiakide.net/scientext20>.

MERRIAM-WEBSTER INC. 2020. *Merriam-Webster On-line Dictionary*. Available at: <https://www.merriam-webster.com>.

MOHIT, JAIN. 2015. *Examen 4 Quizlet*. Available at <https://quizlet.com/109134555/examen-4-flash-cards>.

OKIDOO INC. 2019. *TradooIT*. Available at: <https://www.tradooit.com>.

OXFORD UNIVERSITY PRESS. 2019. *English Oxford Living Dictionary: The Oxford English Corpus*. See: <https://en.oxforddictionaries.com/explore/oxford-english-corpus>.

PEARSON ELT. 2019. *Longman Corpus Network*. See: <http://www.pearsonlongman.com/dictionaries/corpus>.

REVERSO. 2020. *Reverso Online dictionary: English-French translation of words and expressions, definition, synonyms*. Available at: <https://dictionary.reverso.net/english-french>.

SPRINGER NATURE LIMITED. 2019. *Macmillan Dictionary: Corpus*. See: <https://www.macmillandictionary.com/corpus.html>.

UNITED STATES SUPREME COURT OF THE UNITED STATES. 2018. *Oral arguments*. Available at: https://www.supremecourt.gov/oral_arguments/argument_transcript/2019.

References

- BOULTON, Alex & Henry TYNE. 2014. *Des document authentiques aux corpus: démarches pour l'apprentissage des langues*. Paris: Didier.
- COXHEAD, Averil. 2013. Vocabulary and ESP. Brian PALTRIDGE & Sue STARFIELD, *The Handbook of English for Specific Purposes*, London: Wiley-Blackwell, 115-134.
- GROSSMANN, Francis. 2016. Faut-il rappeler cette évidence? Usages des formules de l'évidence dans l'écrit scientifique. *Lildilem seminar: autours des écrits académiques*, Grenoble (France), June 10 2016.
- HARTWELL, Laura M. 2013. Corpus-informed descriptions: English verbs and their collocates in science abstracts. *Études en didactique des langues* 20, 79-94.
- HARTWELL, Laura M. & Marie-Paule JACQUES. 2014. Authorial presence in French and English: "Pronoun + Verb" patterns in Biology and Medicine research articles. *Discours* 15. Available at: <http://discours.revues.org/8941>.
- KILGARRIFF, Adam, Miloš HUSÁK, Katy MCADAM, Michael RUNDELL & Pavel RYCHLÝ. 2008. GDEX: Automatically finding good dictionary examples in a corpus. *Proceedings of EURALEX*, Barcelona (Spain), 2008.
- PEACHEY, Nik. 2020. *Concordancers in ELT*. British Council: *Teaching English*. Available at: <https://www.teachingenglish.org.uk/article/concordancers-elt>.
- ROCHE, Christophe. 2007. Le terme et le concept: fondements d'une onto-terminologie. *Terminologie & ontologie: théories et applications*. Annecy (France), June 1 2007, 1-22.
- RUNDELL, Michael. 2012. *Who needs dictionaries?* British Council: *Teaching English*. Available at: <https://www.teachingenglish.org.uk/article/who-needs-dictionaries-0>.
- ŠEŠKAUSKIENĖ, Inesa & Julija STEPANČUK. 2014. Evidence speaks for itself: metaphors in courtroom hearings. *Filologija* 19, 102-120.
- TUTIN, Agnès & Francis GROSSMANN (dirs.). 2014. *L'écrit scientifique: du lexique au discours. Autour de Scientext*. Rennes: Presses de l'Université de Rennes
- STUBBS, Michael. 2007. Quantitative data on multi-word sequences in English: the case of the word world. Michael HOEY, Michaela MAHLBERG, Michael STUBBS & Wolfgang TEUBERT. *Text, discourse and corpora*. London: Bloomsbury Publishing, 163-190.

Appendix 1 – Expressions containing évidence and their equivalents in English

	<i>Linguee</i>	<i>Reverso</i>	<i>WordReference</i>
<i>à l'évidence</i>	clearly	evidently/ obviously	obviously/evidently/ manifestly
<i>METTRE en évidence (V.)</i>	bring to light/ highlight	show/highlight/ bring to light	highlight/make evident/bring to light there is clear evidence that /
<i>de toute évidence</i>	clearly	apparently/obviously	all the evidence points to / quite evidently in evidence / conspicuously
<i>en évidence</i>	conspicuously	conspicuous	
<i>c'est l'évidence même</i>	X	it's completely obvious	it's/that's obvious
<i>ÊTRE en évidence</i>	X	be clearly visible	be clearly displayed / visible
<i>SE RENDRE à l'évidence</i>	X	bow to the evidence	recognize/ acknowledge that ignore /
<i>NIER l'évidence</i>	X	deny the evidence	deny the obvious / deny the evidence manifestation /
<i>mise en évidence</i>	X	highlighting/ emphasis / revealing	demonstration / illustration
<i>non-évidence</i>	non-obviousness	X	X
<i>évidence scientifique</i>	scientific fact	X	X
<i>évidence empirique</i>	empirical evidence	X	X
<i>vérité d'évidence</i>	truism	X	X
<i>autre évidence</i>	other evidence	X	X
<i>évidence directe</i>	direct evidence	X	X
<i>évidence vidéo</i>	video-evidence	X	X
<i>c'est une évidence</i>	X	obvious fact	X
<i>SE METTRE en évidence</i>	X	X	get yourself noticed/seen
<i>S'IMPOSER comme une évidence</i>	X	X	be self-evident
<i>bien en évidence</i>	X	X	in full view
<i>force de l'évidence</i>	X	weight of evidence	X
<i>d'une grande évidence</i>	X	very obvious	X