



Quintana Toledo, Elena. 2020. Data-driven learning in the academic writing classroom: Citation and stance. *Revista de lenguas para fines específicos* 26.1, pp. 196-213.
<http://dx.doi.org/XXXXXXXXXXXXXX>
<https://ojsupdc.ulpgc.es/ojs/index.php/LFE/index>

Paper received: 19 November 2019

Paper received in revised form and accepted for publication: 08 April 2020



Data-driven learning in the academic writing classroom: Citation and stance

Elena Quintana Toledo¹

Independent Scholar

Former Lecturer ULPGC, Departamento de Didácticas Especiales · E35004 · Las Palmas de Gran Canaria (Spain)

ABSTRACT

Acquiring academic writing skills in English may be a true challenge, especially for undergraduate students for whom English is not their first language, but still have to accomplish writing tasks in a more or less successful way as a requirement to pass certain subjects and eventually qualify in higher education. Citation is perhaps the most distinctive property of academic writing and, at the same time, a feature which requires learners to understand how citation structures and reporting verbs can express the writer's stance towards the imported information and the source authors themselves. This paper seeks to explore the use of a corpus of research papers written by native speakers of English in an English for Academic Purposes classroom with non-native speakers of the language. The proposal aims at boosting the students' academic writing skills by providing them with a set of data-driven learning activities which promote both reflection and practice on a range of citation strategies. Teaching them how to use citation structures and reporting verbs effectively will ultimately allow them to take an adequate stance in their academic papers by shaping their claims in an appropriate way within a given disciplinary discourse community.

Keywords: Citation, reporting verbs, academic writing, stance, data-driven learning

RESUMEN

La adquisición de habilidades de escritura académica en inglés puede ser un auténtico reto, especialmente para los estudiantes de grado cuya lengua materna no es la lengua inglesa, pero aún así deben presentar tareas escritas de manera más o menos exitosa como requisito para superar ciertas asignaturas y poder finalmente graduarse. El empleo de las citas es quizás la cualidad más característica de la escritura académica y, al mismo tiempo, representa la necesidad de los estudiantes de entender cómo las estructuras de las citas y los verbos de habla pueden expresar la actitud del escritor hacia la fuente de información y la información misma. Este trabajo describe de qué forma se puede utilizar un corpus de artículos científicos escritos por hablantes nativos de inglés en una clase de Inglés para Fines Académicos con hablantes no nativos del idioma. El objetivo de la propuesta consiste en mejorar las habilidades de escritura académica de los estudiantes utilizando para ello un set de actividades sustentadas en el aprendizaje basado en datos destinadas a fomentar la reflexión y la práctica sobre el uso de diversas estrategias para referenciar otros textos y a sus autores. En última instancia, la enseñanza del manejo efectivo de las citas y los verbos de habla permitirá a los estudiantes adoptar una actitud adecuada en sus ensayos académicos al tiempo que emiten sus opiniones atendiendo a las convenciones de una comunidad discursiva disciplinar concreta.

Palabras clave: Cita, verbos de habla, escritura académica, actitud, aprendizaje basado en datos

1. Introduction

The spread of English as the preferred language for international communication is beyond dispute today. Its status as a lingua franca can be easily seen in a number of domains which range from entertainment and social

¹ Corresponding author · Email: elena.quintana@ulpgc.es



media to finance, politics and technology. Mauranen (2010, p. 7) includes the academia in this list and points out that

Academia is in many ways the typical ELF domain: it is international, mobile and its dependence on English has skyrocketed in the last few decades. Academia is thoroughly dependent on cooperation across national borders and internationally negotiated standards, especially in science, where cutting edge research teams operate in several countries and recruit from anywhere in the world.

While it is true that there is still a heated debate about the desirability of this phenomenon on several grounds, i.e. accuracy, intelligibility or linguistic diversity (Baird and Baird, 2018; Chapman, 2015; Roux, 2014; Ammon, 2006), it has been also taken for granted that being able to communicate successfully in English is considered to be beneficial for future success in the academic and professional world, not only for scholars and scientists, but also for tertiary education students, no matter what their mother tongue is.

Within the scholarly panorama, English is the vehicular language for the dissemination of researchers' findings and knowledge exchange. It certainly enjoys a privileged position in the publishing industry, so if scholars report their results in English effectively, they will somehow guarantee that their work reaches a wider audience and, consequently, attains visibility. However, it should be borne in mind that they will only gain membership in the academic discourse community provided that they have considerable expertise in "the strategic manipulation of various rhetorical and interactive features" (Hyland, 1999, p. 341), or in other words, they know about the conventions and expectations of academic writing and put them into practice.

Regarding international undergraduate and postgraduate students particularly, current trends seem to point towards an increasing demand for English-speaking countries as the places they would opt for when choosing a university (Vandermensbrugge, 2004). Additionally, in some non-anglophone countries the English language has become the medium of instruction at higher educational levels (Lasagabaster and Doiz, 2018; Cosgun and Hasirci, 2017; Vu and Burns, 2014). This is not surprising in the globalisation era: becoming thoroughly proficient in English means being ready for international mobility and employability.

One of the requirements for qualifying in higher education is to achieve adequate academic writing skills as writing tasks are considered to be compulsory components within the assessment criteria in many, if not all, of the subjects within degree programmes. Starting in their first year, university students have to adapt to "new ways of knowing: new ways of understanding, interpreting and organizing knowledge" (Lea and Street, 2000, p. 32). Under normal circumstances, this is a highly demanding endeavour for most of them, but when it comes to doing written assignments in English, the task becomes a major hurdle for Spanish students of English as a Second Language.

The academic paper is a core genre in higher education. It may be considered to be "knowledge-creating" (Hyland, 2004, p. 12) in the sense that students engaged in writing one are supposed to display originality to some extent (Cryer, 2006, p. 193), frequently because they deal with a new perspective about existing research on a specific scientific topic. It follows that making reference to the contributions made by previous scholars within the discipline turns out to be an important feature of academic writing. As a matter of fact, according to Swales (2014, p. 119), it "is the most overt and most immediately obvious indication that a text is indeed academic". In this regard, and referring to "knowledge-creating genres", Hyland (2004, p. 12) states that they should exhibit the following characteristics:

- establish the novelty of one's position
- make a suitable level of claim

- acknowledge prior work and situate claims in a disciplinary context
- offer warrants for one's view based on community-specific arguments and procedures
- demonstrate an appropriate disciplinary ethos and willingness to negotiate with peers

As such, this characterisation stresses the interactional dimension of citation in academic writing, which can be readily perceived on account of the writer's choice of interpersonal strategies including how citations are framed and the use of reporting verbs. Some studies reveal that many undergraduate students still fail to recognise and put into practice these interpersonal strategies in their academic papers (Lee, Hitchcock and Casal, 2018; Luzón, 2015; Borg, 2000), which may result in accusations of plagiarism. In this light, "we should expect courses in academic writing to sensitize students to the choices that are available to them when they decide to refer to other texts", as Thompson and Tribble (2001, p. 91) put it.

This paper seeks to explore the use of a corpus of research papers written by native speakers of English in an English for Academic Purposes classroom with non-native speakers of the language. The proposal aims at boosting the students' academic writing skills by providing them with a set of data-driven learning activities which promote both reflection and practice on a range of citation strategies. The theoretical framework for the formal and functional characterisation of citations draws on Swales' (1990) distinction between integral and non-integral citations, and on Thompson and Tribble's (2001) further elaboration of Swales' (1990) distinction. Hyland's (2013) model of reporting verbs also proves to be helpful in the detection of the process and evaluative functions expressed by them. Teaching students how to use citation structures and reporting verbs effectively will ultimately allow them to take an adequate stance in their academic papers by shaping their claims in an appropriate way within a given disciplinary discourse community.

2. Citation in academic writing: Stance, engagement and evidentiality

The fact that academic writing stands as an example of interactional discourse is widely acknowledged (Livnat, 2012; Lorés-Sanz, Mur-Deñás and Lafuente-Millán, 2010; Hyland, 2008, 2005a, 2005b, 2004). It cannot be seen any longer as a monolithic presentation of facts in a detached and impersonal way, but rather as a space where authors evaluate their own ideas and the ideas of others and, by doing so, they create bonds with their readers. In fact, much has been written during the last three decades about the way writers interact with their audience and about the linguistic/rhetorical strategies they make use of when adopting an attitude towards their findings and other academics' investigations. In this sense, academic writing does not operate in a vacuum and has to be considered as a "social process" (Hyland, 2004, p. 12).

The interactional character of academic writing becomes apparent thanks to a wide variety of resources and strategies. According to Hyland (2005a, p. 176), there are two means by which writers can handle how relationships are manifested in a text, namely, "stance" and "engagement", which may be taken as two sides of the same coin. The former is an attitudinal facet of academic writing which allows authors to get self-represented in their text while "convey[ing] their judgements, opinions and commitments" (Hyland, 2005a, p. 176). Writers can opt for adopting an explicit position towards the information presented, or for masking their actual position by keeping their authorial involvement at a low level. The latter, in contrast, "is an alignment dimension" (Hyland, 2005a, p. 176) which has to do with audience awareness and reader involvement. This means that authors grant the existence of potential readers, leading them to certain interpretations and acknowledging, at the same time, the possibility that they may hold conflicting views regarding the topic under research.

Citation, understood here as “attributing propositional content to another source” (Hyland, 2004, p. 20), can be considered to be a signal of both stance and engagement. On the one hand, stance, as an attitudinal, writer-oriented trait in academic writing is recognisable through the author’s assessment of propositional information which may be attributed to themselves or to others. A case in point is the writers’ frequent presentation of their research by evaluating its novelty in the corresponding disciplinary field, claiming their authority by virtue of its originality. Nevertheless, the notion of originality cannot be taken here as “the expression of autonomous self divorced from existing theories and discourses; rather it always involves, to a large extent, already legitimated thoughts within a certain theoretical or conceptual framework” (Kubota, 2003, p. 80). Along the same lines, Hyland (2002, p. 215) observes that writers should try “to balance claims for the significance, originality, and correctness of their work against the convictions and expectations of their readers”. Therefore, there is a need for contextualising their work by making reference to others, either by appraising or simply judging their suitability for the purpose of their investigation. Citation, then, serves as support when authors have to write critically about the extent to which their contribution is a worthy one.

On the other hand, citation may be simultaneously an indicator of engagement in academic writing, at least in a general sense. References to different sources of information are symptomatic of the writer’s compliance with the rhetorical conventions and expectations of the disciplinary discourse community. Thus, they strengthen a shared sense of collegiality between writer and readers. As an engagement strategy, they appear to be essential to draw readers to intended interpretations since other scholars’ previous work may provide arguments or counterarguments for the writer’s ideas. Were citation practices not to have an impact on how readers shape meanings, the process of sorting out “what and who to cite, why to cite, how to cite, and even when not to cite at all” (Polio and Shi, 2012, p. 95) would not be as critical as it actually is. Lastly, the use of citation comes to represent a dialogic exchange between the writer and the source authors that are integrated in his/her paper; somehow s/he “must respond to the author’s ideas and/or words” (Mori, 2017, p. 3).

Citation is also inextricably related to the concept of evidentiality whose relevance in linguistics dates back to Boas’ work on the obligatory morphological marking for “source of information” that some languages possess (Aikhenvald, 2006, p. 283). Whenever evidentiality is brought to the foreground, it unavoidably goes hand in hand with epistemic modality. Their relationship has been approached from different conflicting angles and, so far, scholars have not come to terms with the most suitable way to deal with the two notions.

Some authors view one of them as being subsumed within the other, i.e. inclusion, which leads us to the broadest interpretation of evidentiality to date. Chafe (1986, p. 271) conceptualises the term “broadly to cover any linguistic expression of attitude toward knowledge”, including (the type of) evidence and a set of epistemological factors like reliability, probability and mode of knowing. Drawing attention to mode of knowing, too, Willet (1988, p. 55) defines evidentiality as “the linguistic means of indicating how the speaker obtained the information on which s/he bases an assertion” and assigns degrees of reliability to the information obtained depending on the source. In like manner, Mithun (1986, p. 89) argues that evidential markers contribute to assessing the reliability of information as indicators of “the source of evidence on which statements are based, their degree of precision, their probability and expectations concerning their probability”.

A dissimilar standpoint, albeit still inclusive in nature, is that hold by Palmer (2001, p. 8) for whom propositional modality comprises two categories, that is, epistemic modality and evidential modality, the difference between them being that epistemic applies to speakers’ “judgements about the factual status of the proposition”, whereas evidential involves “the evidence they have for its factual status”. It is worth noting that the sense of factuality

used in this framework refers to facts assumed to be true by the speaker, and that the degree of commitment to the truth of the proposition expressed is often invoked.

Van der Auwera and Plungian (1998) are the leading exponents of the intersective approach where evidentiality and epistemic modality are regarded as interrelated but notionally distinct. For these authors, the intersecting point of encounter is inferential evidentiality. In other words, modality markers do not necessarily indicate the source of information nor do all evidential markers necessarily indicate epistemic values; if, and only if, these markers convey both source of information and epistemic value, they are said to overlap (Faller, 2002, p. 83; Carretero and Zamorano-Mansilla, 2014, p. 320).

Linguists applying the disjunctive methodology endorse the view that there should be a clear-cut distinction between evidentiality and epistemic modality. Some representative cases are De Haan (2005), Pietrandrea (2005) and Nuyts (2001). In this model, the differentiation between evidentiality and epistemic modality stems from the very notion of evidence and how it is used when defining each category: while evidentiality asserts the existence of some sort of evidence for a given proposition, epistemic modality qualifies the evidence in terms of truth-values for the corresponding proposition. Following Cornillie (2009) and Cornillie and Delbecque (2008), Alonso-Almeida (2015a, 2015b, 2012) mainly sees these two categories as independent from one another, but also admits that “functional overlapping” (2015a, p. 34) may occur at times. For him, evidentials have to do “with information source and the way in which knowledge is construed” and they are stance markers insofar as they characterise “the role of the speaker in the formulation of predication and its relation to the information presented in terms of authorial responsibility” (Alonso-Almeida, 2015b, p. 122).

Studies in academic discourse, which is the concern of this paper, offer interesting insights into this issue. Hyland (2005a, p. 178) claims that evidentiality is one of the three constituents of stance together with “affect” and “presence”. More specifically, evidentials are rhetorical features pertaining to the interactive dimension in his model of metadiscourse (Hyland, 2005b, p. 49), which brings together a rather heterogeneous group of rhetorical features all of which signal authorial attitudes in one way or another with respect to the text itself and to discourse participants. Evidentials are classified with transition markers, frame markers, endophoric markers and code glosses primarily as discourse organisational devices, although their function is not limited to that alone (Hyland, 2005b, p. 50-52). They are defined as references to previous texts which pursue a twofold aim: to help readers in the interpretation process and to substantiate the writer’s authority on the subject matter. Hyland (2005b, p. 51) goes on to explain that evidentiality “may involve hearsay or attribution to a reliable source; in academic writing it refers to a community-based literature and provides important support for argument”.

As pointed out earlier, academic writing learners have to prove that they are able to handle source materials efficiently. This certainly involves developing adequate citation skills as well as putting into practice other evidential strategies like the use of reporting verbs. They undoubtedly qualify as carriers of evidential values which fall squarely into Aikhenvald’s (2007, p. 211) quotative category, that is, “reported information with an overt reference to the quoted source”. They are of special interest to studies on academic writing since they embody the “nature of the channel of evidence drawn upon” (Breeze, 2017, p. 296), and they come to represent how writers conceptualise disciplinary knowledge and how readers construe the writers’ epistemological positioning.

In the light of the discussion presented in this section with reference to evidentiality and epistemic modality, it is obviously necessary to indicate at this stage which of the approaches is taken in this work. To the author of this paper, it seems clear that evidentiality and epistemic modality are conceptually different, but this does not

mean that (un)certainty and (un)reliability overtones cannot coexist with the indication of the source of information. I believe that the most inclusive approach is amenable to provide a more exhaustive account of the interactional character of citation and reporting verbs and hence it is the approach that I side with. On a more practical level and taking into consideration the theoretical-targeted sessions on citation and evidentiality that students are expected to attend to, there is no need for them to make fine-grained distinctions between evidential and epistemic meanings of reporting verbs, although such sophisticated theoretical reasoning will be welcomed when discussing functions.

2.1. Citation typologies and reporting verbs

A relatively straightforward way to categorise citations is by pinpointing their surface realisations. Swales (1990) puts forward a classification distinguishing between integral and non-integral citations, which differ in how the name of the cited author is given. In integral citations, the names are integrated in the citing sentence as “subject [...], passive agent [...], as part of a possessive noun phrase [...] and as what Tadros (1985) calls ‘an adjunct of reporting’” (Swales, 1990, p. 148). In non-integral citations, in contrast, the name is given outside the sentence, either in a parenthesis or superscript. Integral and non-integral citations are said to differ depending on whether the writers wish to place the emphasis on the cited author or on the cited information itself, respectively.

Thompson (2001, 2005) and Thompson and Tribble (2001) further elaborate Swales’ classification by specifying subcategories within each type combining both formal and functional criteria of the citations. Integral citations can be (i) verb controlling, (ii) naming, and (iii) non-citation. Consider the following examples:

- (1) Thompson (2001) defines verb controlling citation as...
- (2) This research follows the definition of naming citation as in Thompson (2001).
- (3) The citation typology proposed by Thompson differs slightly from the one we use in this paper.

(1) is an example of a verb controlling citation where the citation itself fulfils the function of an agent which controls the lexical verb “defines”. (2) illustrates naming citation because it takes the form of a noun phrase or part of it. Particularly, this is an instance of “reification” as “the noun phrase signifies a text, rather than a human agent” (Thompson and Tribble, 2001, p. 96). (3) exemplifies non-citation: the cited author’s name is given but the year is not provided. This is a frequent practice in case the reference is given earlier in the text and the writer decides not to include it again.

Non-integral citations, on their part, can be (i) source, (ii) identification, (iii) reference, and (iv) origin. Take (4) to (7) below:

- (4) Making reference to others’ research is a major difficulty for academic writing learners (Borg 2000).
- (5) Reporting verbs have been analysed as indicators of how writers evaluate previous research in the introduction section of scientific papers (Thompson and Ye 1991).
- (6) Cases of self-citation have been discarded for the purposes of the present study (see Hyland 2001).
- (7) Data extraction was done with AntConc (Anthony 2005).

The citation shown in (4) points to where the propositional information comes from, i.e. source. (5) contains a citation which identifies Thompson and Ye as the authors of the analysis being referred to. (6) stands as an example of reference citation in which the citation is used in conjunction with the directive “see” to signal further reference to another text. The last citation in the set specifies who coined a certain concept or who created a certain product, i.e. origin.

Apart from the explicit attribution features that are exhibited through the use of integral and non-integral citations and its various types, the role of reporting verbs as devices which indicate the authors' attitudes to their own claims and to those of others has been also stressed in earlier literature. For example, Hunston (1993) draws attention to the consequences of choosing a reporting verb as this choice characterises the reported information as "received knowledge". Tadros (1993) demonstrates that writers can select certain reporting verbs in order to show distance from the reported information and the source authors themselves.

Thompson and Ye (1991, p. 369-373) take their analysis of reporting verbs a step further and offer a comprehensive taxonomy. They catalogue them into two main classes: denotation and evaluation. The verbs falling into the denotation category are said to pertain to either author acts or writer acts. It should be noted that author acts refer to the "quoted voice" and writer acts are related to the "quoting voice" (Jafarigozar and Mohammadkhani, 2015, p. 2492). Author acts are classified into (i) textual (verbs referring to processes of verbal expression, i.e. *point out, state, write*); (ii) mental (verbs referring to cognitive or mental processes, i.e. *believe, focus on, think*); and (iii) research (verbs referring to mental or physical processes in any research activity, i.e. *calculate, demonstrate, find*). Writer acts are categorised into (i) comparing (different source texts are set against one another, i.e. *correspond to*); and (ii) theorising (a source text is set against the citing text, i.e. *support*).

Within evaluation, Thompson and Ye (1991, p. 369-373) establish three subcategories: (i) author's stance, which can be positive (the reported proposition is taken as true, i.e. *accept*), negative (the reported proposition is taken as false or wrong, i.e. *question*), or neutral (i.e. *undertake*); (ii) writer's stance, which can be factive (the writer agrees with the reported proposition, i.e. *show*), counter-factive (the writer disagrees with the reported proposition, i.e. *confuse*), or non-factive (the writer remains neutral, i.e. *propose*); and (iii) writer's interpretation, which includes author's discourse interpretation verbs (they enable the writer to relate the author's proposition to the source text, i.e. *add, remark*), author's behaviour interpretation verbs (they show the writer's perception of the reason why the author expressed a certain proposition, i.e. *admit, warn*), status interpretation verbs (they show the writer's judgement regarding how the reported proposition is related to the text, i.e. *account for, conform to*), and non-interpretation verbs (they allow the writer to report the information objectively, i.e. *employ, use*).

As may be inferred from the above description, Thompson and Ye's taxonomy is not only exhaustive but also overlapping at times and so, a single reporting verb may be classified in more than one category because it may encode varied interpersonal meanings and functions. Besides, Hyland (2013) notes that their taxonomy is based on a study of the introductory sections of research articles, which means that it probably overlooks fine-grained differences arising from how reporting verbs are used in different sections. As an alternative, he offers a classification system of the functions of reporting verbs which is partially based on Thompson and Ye (1991) and Thomas and Hawkes (1994), and which aims to be less confusing.

Hyland (2013, p. 118-119) draws a line between process functions and evaluative functions. The former covers reporting verbs referring to a certain activity type such as (i) "Research (Real World) Acts", i.e. *analyse, calculate, explore, show*, (ii) "Cognition Acts", i.e. *assume, believe, conceptualise, view*, and (iii) "Discourse Acts", i.e. *ascribe, discuss, report, state*. Even though they are roughly equivalent to Thompson and Ye's "research", "mental" and "textual" acts, this system explicitly allows for reporting verbs to fall into more than one category. In such circumstances, one of them would count as a "primary aspect [...] of the research process" (Hyland, 2013, p. 118).

The evaluative functions of reporting verbs are directly derived from the activity types mentioned above. Research Act verbs report on either “findings” or “procedures” and the kind of evaluation attached to them can be (i) factive, i.e. *confirm, demonstrate*, (ii) non-factive, i.e. *find, observe*, or (iii) counter-factive, i.e. *fail, ignore*. The associated authorial attitudes to Cognition Act verbs fluctuate between (i) positive, i.e. *agree, concur*, to (ii) critical, i.e. *disagree, dispute*, including (iii) tentative, i.e. *believe, suppose*, and (iv) neutral positionings, i.e. *anticipate, conceive*. Discourse Act verbs can express doubt in (i) tentative, i.e. *hypothesise, suggest*, or (ii) critical ways, i.e. *exaggerate, not account*, and assurance in (i) factive or (ii) non-factive ways.

3. Methodology: Data-driven learning

Data-driven learning (DDL) is defined by Johns and King (1991, p. iii) as “the use in the classroom of computer-generated concordances to get students to explore regularities or patterning in the target language, and the development of activities and exercises based on concordance output”. It is rooted in other language-learning methodologies like “language awareness” and “consciousness-raising” (Granger and Tribble, 1998, p. 200). In a DDL approach, learners are given direct access to corpus data in order that they can explore language forms and functions for themselves. It differs essentially from other language-learning methodologies, i.e. grammar-translation, in that learners will be examining authentic samples of the target language while carrying out the tasks.

It stands to reason that one of the main advantages of DDL is that learners can practice with a considerable amount of naturally-occurring language (Boulton, 2009, p. 84) which, in turn, contributes to increasing their sensitivity to certain language patterns (Gabrielatos, 2005). Secondly, it promotes learner autonomy as the exploration of the corpus and the use of the concordance tools which facilitate it may well extend beyond teaching hours (Boulton, 2009, p. 86). And thirdly, it may be useful in learners’ self-correction by allowing comparison between their own written productions and those produced by native speakers (Guilquin and Granger, 2010).

In spite of the fact that DDL is not devoid of some weaknesses, i.e. it is time-consuming and learners are supposed to fulfil high requirements, it has the potential to bring about benefits in learning situations. Ackerley (2017) and Aston (2015) describe how DDL is applicable to teaching phraseological items to English as a Second Language learners through DDL activities based on written and aligned speech corpora, respectively. Curado-Fuentes (2017) and Chujo et al. (2013) address the issue of teaching grammar points to non-native speakers of English in a university setting. While the proposal of Curado-Fuentes is aimed at intermediate to upper-intermediate students, Chujo et al.’s is designed for a beginner-level group. However, both studies comment on significant gains in student motivation and linguistic competence. Another area where the positive effects of DDL have been found to be worthwhile is vocabulary instruction as in Barabadi and Khajavi (2017) and Yilmaz and Soruç (2015). In both experimental studies, traditional and DDL methodologies to teach vocabulary are compared, and the authors make use of materials gathered from the *Corpus of American Contemporary English* to carry out tasks with the help of its concordance programme. As to the feasibility of DDL in writing skills, some works evince its positive outcomes, for instance, Chen and Flowerdew (2018) report on the use of teacher-built corpora representing different disciplines for PhD students to perform hands-on tasks with AntConc so as to identify discourse strategies in academic texts. Larsen-Walker (2017) looks into the use of the *Michigan Corpus of Upper Level Student Papers*, Key Word in Context searches and instructional materials to improve learners’ accuracy of the use of linking adverbials.

The proposal presented in this paper is aimed at Spanish students in their third year in the degree in Modern Languages at the University of Las Palmas de Gran Canaria. This degree seeks to provide students with proficiency in two languages, i.e. itinerary English-French or itinerary English-Chinese, and some notions of a third language, i.e. Arabic, Chinese, French, German or Greek, depending on the itinerary chosen by the student. With the exception of the subjects which are taught in the languages mentioned above, the rest of them are taught in English.

The degree programme for their third year encompasses a compulsory subject, namely, Inglés VI, which, according to the syllabus, “is the last in a series of six subjects corresponding to the obligatory component ‘English language: Theoretical and practical aspects’” (Fernández-Martínez and Vizcaíno-Ortega, 2018-2019). Students are strongly recommended to have reached a B2 level of English language proficiency according to the *Common European Framework of References for Languages* (Council of Europe, 2001) before enrolling into it. One of the main goals of the course is learning how to write an academic paper; then, there are both theoretical and practical sessions devoted to tackle this point of the syllabus throughout the semester. At the end of it, students must submit an academic paper whose assessment counts towards the final mark of the subject.

Learners are expected to do practice activities in a relatively controlled environment where both a corpus of academic writing and a corpus tool are necessary. An English native speaker corpus of 50 linguistics research papers has served as a source of examples in the design of the activities. It is part of a larger compilation which contains scientific articles also covering two more register domains, i.e. engineering and medicine. They have been further organised according to specific topics within each discipline whenever possible. In the case of linguistics, for example, some of the topics touched upon include pragmatics, prosody, semantics or syntax, among others. The choice of texts has been made randomly, except for the fact that they had to be published in high impact factor journals between 2016 and 2019. Some of these journals are *Applied Linguistics*, *Linguistic Inquiry*, *Journal of Semantics* and *Language and Literature: International Journal of Stylistics*. This compilation has been carried out for the purposes of the research project “Identification and Analysis of Metadiscourse Strategies in Spanish and English” (IAMET), which aims to describe cross-linguistic variation in the use of metadiscourse strategies in the scientific register. The suitability of the English native speaker corpus may be called into question, but I concur with Swales (1990, p. 113) on the idea that reflecting upon what expert writers do, and in this particular case about how they acknowledge others’ research, may serve as a good starting point for novice academic writers.

Students have to operate a concordancer so as to explore the corpus and, needless to say, this is a major drawback for them: seldom are they acquainted with corpus tools. Despite the fact that the IAMET corpus comes along with a specific tool, i.e. METOOL, which allows for the identification of metadiscourse devices including those with evidential and epistemic values (Carrió-Pastor, 2019), other queries related to integral and non-integral citations such as the use of quotation marks, parenthesis and dates cannot be made by its search engine. As a result, METOOL is not suitable in the envisaged classroom scenario and AntConc will be used instead. This is a freeware corpus analysis toolkit which has been designed not only for researchers, but also for classroom use (Anthony, 2004, 2005). It makes available to users “a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis and a word distribution plot” (Anthony, 2004, p. 7). Learners will be given a thorough introduction to the tool in order to show them how the interface works, paying special attention to its limitations and to detailed instructions on the use of wild cards, blank space characters and case sensitive searches.

4. DDL activities

The activities presented in this section are aimed at providing students with examples which illustrate how citation and reporting verbs are used in academic writing as well as at promoting observation of the texts to reflect on the various means writers have at their disposal when they have to make reference to the work of other writers in their own work. Hopefully, the activities themselves and the exploration of the corpus will provide them with models which they can progressively incorporate into their written assignments until they have gained enough academic literacy skills and have developed their own authorial voice with confidence.

1. Classify the following citations into integral and non-integral. Specify the function they fulfil according to the formal and functional criteria shown in Thompson and Tribble's (2001) taxonomy, i.e. (i) verb-controlling, naming and non-citation, and (ii) source, identification, reference, and origin.

- (1) The standard approach to mark-up and coding is generally manually driven, so is highly labour-intensive (Fanelli et al. 2010: 70) and often subjective [Engappli2016_4]
- (2) The final stage of grammaticalization may appear in Lithuanian where, according to Nevis & Joseph (1993), the reflexive marker, historically a clitic and parallel in its positioning to the endoclitics of Pashto, is synchronically fully affixal (but see also Stolz 1989) [Engling2016_2]
- (3) Whether the facts regarding island (in)sensitivity support this analysis is debatable: see Abels 2011 and Barros, Elliott, and Thoms 2014 for critical discussion of the notion of island repair [Englingsynta2016_9]
- (4) By means of a computer-based keyword analysis procedure using WordSmith Tools (Scott 2004), the potential for human bias in selection of 'key concepts' for further investigation was mitigated [Engappli2016_3]
- (5) Derwing and Munro (2009) define English speakers' accents in relation to the local variety [Engappli2016_2]
- (6) We note, however, that Fox's analysis of ellipsis fundamentally does not translate over to the only examples discussed here, as shown in Kehler & Buring (2007) [Englingseman2016_3]
- (7) In Mian the contrast is only in agreement: éil 'pig' takes masculine agreement for the male and feminine for the female (Fedden 2011: 170), while the epicene koból 'cassowary (a flightless bird)' takes feminine agreement, irrespective of sex (Fedden 2011: 171) [Engling2016_1]
- (8) For Dalrymple & Mycock (2011) the availability of syntactic information to prosody and vice versa required by such phenomena is limited to the marking of syntactic and prosodic constituent edges, which are required to match up to a greater or lesser degree (similarly, outside LFG, Bermúdez-Otero 2012) [Engling2016_2]
- (9) Both POSS-i and SELF-i have received attention in the sign language literature (for POSS-i, see Abner 2012; for SELF-i, see Fischer & Johnson 2012; Koulidobrova 2009) [Englingseman2016_3]
- (10) The analyses were carried out using STATA 13 (StataCorp., 2013) [Engmedhealthcareedu2_2016]
- (11) Hanna and Turville-Petre (2014: 136–137) extend Duggan's recommendations to the text of the alliterative *Morte Arthure*, nominating 11b and several other b-verses for the editorial chopping block [Engstyli2016_6]
- (12) As a result, V1 declaratives were lost (Simonenko & Hirschbühler 2012), as were sequences like (13), as shown in Donaldson (2012) [Engling2016_5]

Students will be given a handout with the teacher-generated list of citations above or, alternatively, can be given instructions to make specific queries in the concordancer so that they can generate the list themselves. While commenting on their conclusions regarding the classification of the citations, the teacher is supposed to promote discussion about how integral and non-integral citations shape authorial attitudes and intentions. This activity may be enriched by asking students to look for integral and non-integral citations on their own which, in turn, would stimulate the autonomous exploration of the corpus.

2. Group the citations below according to the activity types denoted by the reporting verbs used in each of them, i.e. research acts, cognition acts and discourse acts. Refer to the type of evaluation expressed as appropriate.

- (13) A straightforward case of a numberless three-person system is that of Jarawa (Kumar 2012). Its pronouns are shown in table 5 (allomorphs omitted). Kumar **states** explicitly that these pronouns are numberless and cover both singular and plural persons [Englingmorph2016_8]
- (14) Dalrymple & Mycock (2011) **assume** that category edge information is passed down the c-structure and is associated with s-string elements, where it must be available at the interface with prosody [Engling2016_2]
- (15) Both van Benthem (1986: 125) and Heim & Kratzer (H&K 1998: 178) **find** a type mismatch problem in semantically interpreting expressions like (1b) [Englingseman2016_2]
- (16) Hardt (2006) and Grant (2008) **present** even more evidence against a MaxElide account of the contrast between (ia) and (ib); for this reason, we limit ourselves to discussing configurations involving movement and leave the constraints on the availability of sloppy interpretations as a topic of future research [Englingsynta2016_9]
- (17) Similarly, Stockwell (2002) **treats** negated metaphors as forms of sentence metaphor, with the same low level of visibility [Engstyli2016_8]
- (18) Duffell (2000a) **shows** in some detail what Chaucer's caesura owes to cosmopolitan metrical culture [Engstyli2016_6]
- (19) Lasnik and Park (2013) **note** that long-distance object extraction from a finite clause is impossible even with intervening focus, contra the prediction of MaxElide [Englingsynta2016_9]
- (20) Thus, in many ways, our central theoretical result is that Bergen et al.'s model **predicts** embedded implicatures in non-monotone and downward-monotone contexts if it is combined with a full theory of semantic composition [Englingseman2016_4]
- (21) Birkenes et al. (2014) **find** evidence for this in a corpus of German narrative texts written between the 17th and the 19th centuries [Engling2016_1]
- (22) Derwing and Munro (2009) **discuss** accent in terms of its salience, or how different it is perceived to be from the local dialect, and comprehensibility, that is 'the listener's perception of how easy or difficult it is to understand a given speech sample' (478) [Engappli2016_2]
- (23) Hinzelin (2007: 232) **concludes** that postposed clitics occur as a result of dialect-internal variation [Engling2016_5]
- (24) Gambarotto & Muller (2003) **demonstrate** that many spatial relations can be defined using regions as the only spatial type [Englingseman2016_1]

This task prompts learners to think about reporting verbs not only in terms of the types of activities attached to them in Hyland's (2013) framework, but also to ponder those cases where categories overlap. *Conclude*, for instance, can be taken to express either a cognition or a discourse act; for this reason, learners should be encouraged to provide justifications for their answers when assigning activity types to categories. Teachers should emphasise the need to give details related to the writers' evaluation of the imported material as seen through the choice of these reporting verbs.

3. Think of a synonym or a near-synonym for the reporting verbs in bold. Some of the verbs have already been changed for a synonym or a near-synonym and do not coincide with the writers' actual choice. Interrogate the corpus to find out what their actual choice was.

- (25) Like Steen (2008), Gavins (2007: 152) **claims** that the norm is for the target-world to be in the foreground and the source-world to be in the background [Engstyli2016_8]
- (26) Seedhouse and Almutairi (2009: 324–6) **carried out** a multimodal analysis of the meaning conveyed by A in line 20 and how this meaning is understood by Y and C [Engappli2016_4]
- (27) Principled exceptions to this generalization **are noted** by Farkas and Giannakidou (1996), Kennedy (1997), and Kayne (1998), but they do not undermine the data discussed here, which control for these exceptions [Englingsynta2016_9]

- (28) Hinzelin (2007, also 2009), who **examined** legal charters from the 11th to the 14th centuries, does not report specifically on SV declaratives, but his data from unmarked affirmative main declaratives (the types subsumed under his type Ix, see Hinzelin 2007: 34) [Engling2016_5]
- (29) Grice (1975) famously **provided** a philosophical framework for understanding the driving forces behind such pragmatic enrichment [Englingseman2016_4]
- (30) Hanna and Turville-Petre (2014: 136–137) **extend** Duggan’s recommendations to the text of the alliterative *Morte Arthure*, nominating 11b and several other b-verses for the editorial chopping block [Engstyli2016_6]
- (31) Myklebust (2012: 406–408) **outlines** a different, metrically systematic scansion of the Ellesmere-Hengwrt reading at *Canterbury Tales* VII 3418 [Engstyli2016_6]
- (32) Apparent subject lowerings in Chamorro (Chung 1990) and Tagalog (Sabbagh 2005) **raise** similar issues [Englingproso2016_6]
- (33) Cumming et al. (2006) and Banerjee et al. (2007) both **focus** exclusively **on** written task types (in the TOEFL iBT and IELTS, respectively) [Engappli2016_1]
- (34) Bergen et al. (2014) briefly **consider** giving false states tiny positive probability [Englingseman2016_4]

The choice of words matters in any piece of writing, but even more so when you are dealing with an academic paper. This exercise is precisely meant to raise awareness of the role of synonyms as tools writers can use to provide their texts with lexical cohesion (Bailey, 2006, p. 109). Students have to establish relationships between lexical items and opt for the most appropriate reporting verb to achieve accuracy of meaning as intended by the writers. Eventually, they will be able to learn the reporting verbs from close examination of the context in which they appear.

4. The list below contains both integral and non-integral citations. Transform the integral citations into non-integral ones and vice versa by paraphrasing the information using a suitable reporting verb when necessary.

- (35) Dryer (2011a, b) indicates that of 1326 languages investigated for a study of minor morphological means of signalling negation, only 10 languages involve non-concatenative morphological processes [Engling2016_3]
- (36) The Rouergue data are of interest because Hinzelin (2007: 161) finds postverbal clitics to be more frequent than preverbal clitics in affirmative SV declaratives, unlike in charters from other regions, where preverbal clitics predominate [Engling2016_5]
- (37) For instance, Lasnik and Park (2013) observe that subextraction from a DP in the object position in VP-ellipsis is unsalvageable, even when the object is contained in the highest VP in the ellipsis site, as in (32); (33) again shows that overt parallel extraction is different [Englingsynta2016_9]
- (38) Alsharif & Sadler (2009) examine negation in Modern Standard Arabic [Engling2016_3]
- (39) This is the case for example in Tamil, where gender assignment is strictly semantic, i.e. male humans and gods are masculine, female humans and goddesses are feminine and everything else is neuter (Asher 1985: 136–137; Corbett 1991: 8–10) [Engling2016_1]
- (40) Considerable dialectal variation existed within the OOc zone (Bec 1963: 34–58), and few if any OOc texts are dialectally pure (Lafont 1967: 23; Smith & Bergin 1984: 4) [Engling2016_5]
- (41) A similar construction is possible in Italian (16) (Benni 1976[2011]: 27), but its context is restricted, being limited to athletes and their injuries [Engling2016_1]
- (42) The digitisation of public domain books has led enterprising readers to dig into the ‘great unread’, the deluge of fiction from the Victorian era that accompanied the mechanisation of the printing press (Cohen, 2002) [Engstyli2016_9]
- (43) In TBLT, a task is designed to ensure that meaning is primary, there is a communication problem to solve, there is a relationship to comparable real-world activities, completion has some priority, and assessment is in terms of outcome (Ellis 2003) [Engappli2016_4]

As stated in subsection 2.1, Swales’ (1991) integral and non-integral citations have different rhetorical functions: while the former foreground the source author, the latter give prominence to the source author’s contribution. With this activity, learners will practise shifting prominence from the referenced author to the given information and the other way around and will consider what shifts in prominence reveal about the writers’ attitude.

5. Interrogate the corpus to find out the most frequent grammar patterns used with the reporting verbs *argue*, *propose*, *suggest*, *assume* and *find*.

- (44) Lillo-Martin & Meier (2011) argue against this view, pointing to examples of exceptional first-person forms, as well as a number of syntactic effects of directional verbs [Englingseman2016_3]
- (45) On the other hand, the necessity of formal variables has been contested in semantic theory; in particular, Jacobson (1999) **argues for** a Variable-Free Semantics (VFS), grounded on the observation that variables are not logically necessary for expressive purposes [Englingseman2016_3]
- (46) Szabo (2008) **argues that** extensionalists can avoid Parsons's apparently problematic prediction [Englingseman2016_5]
- (47) Takahashi and Fox (2005) **propose an account** in terms of ellipsis parallelism, an identity condition on ellipsis, and MaxElide, a constraint that favors deletion of the largest constituent possible; crucially, the domain of application of MaxElide is the domain defined by the parallelism constraint, called a parallelism domain (PD) [Englingsynta2016_9]
- (48) Kratzer (1998, 2009) and Heim (2008) **propose that features** on bound pronouns do not exist at LF, but instead are inherited via syntactic agreement [Englingseman2016_3]
- (49) The findings also **suggest a listening comprehension advantage** for test takers who are familiar with accents [Engappli2016_2]
- (50) Scholars **suggest that underdeveloped language proficiency** is partly due to teachers' tendency to neglect language during content instruction, and form-focused instruction (FFI) has been suggested as a possible remedy [Engappli2016_5]
- (51) Alternatively, we could **assume a more direct linkage** between pragmatic and prosodic principles [Englingproso2016_6]
- (52) This departs from Quine (1960), Steedman (1985) and Jacobson (1999), who all **assume that combinators** exist in the grammar, but only apply to lexical heads and to the outputs derived by operators [Englingseman2016_3]
- (53) This observer could be conceptualized more abstractly as a POV (Levinson 1996, 2003; Kracht 2008; Bohnemeyer 2012). In examples such as (43) in which no observer is overtly specified, the observer is usually **assumed to be** the speaker [Englingseman2016_1]
- (54) She did not **find a significant relationship** between scores on the MC listening assessment and the speakers who delivered the inputs [Engappli2016_2]
- (55) Like Chemla & Spector, we **find that scalar terms in non-monotone environments** support implicatures inferences (though these seem not to be the preferred or most salient construals) [Englingseman2016_4]

The point of this form-focused activity is to bring about the students' familiarity with the grammar patterns that are available to them when using the reporting verbs specified above. The reason why these verbs (and not others) have been chosen is their high frequency of occurrence in the corpus. The AntConc concordancer facilitates the observation of the recurring patterns as the concordance lines show the collocates of the search term highlighted in different colours.

6. Interrogate the corpus to find out the top collocates of the subject pronouns *I* and *we*. Make a list of those collocates which are reporting verbs and comment on the most frequent type of activity denoted by them.

- (56) I agree with these authors that [...] [Engappli2016_3]
- (57) In this section, I approach the question from the opposite side [...] [Englingseman2016_3]
- (58) First, I argue against prominent extensionalist analyses given by Parsons and Szabo [Englingseman2016_5]
- (59) Following Link (1983), I assume a relation between portions of matter and individuals [Englingseman2016_1]
- (60) however, we believe that such examples could be made to fit within a structural view of parallelism [...] [Englingsynta2016_9]

- (61) we conclude that its particular approach to enrichment is at odds with the patterns for embedded implicatures
[Englingseman2016]
- (62) Rather we find that lexical entries in that part of speech select for just one of the values
[Engling2016_1]
- (63) However these questions are resolved, the conclusion we stress here is that the observations [...]
[Englingproso2016_6]

The objective of the last task is to let learners discover how writers themselves use reporting verbs to talk about their research and how this contributes to shaping their voice by expressing doubt or assurance.

5. Concluding remarks

This paper has addressed the way in which DDL activities may be integrated into an English for Academic Purposes classroom in order to teach citation structures and reporting verbs as strategies which make manifest the writer's attitudes towards both the source author and the information presented. It seems that there is a need for learners to acquire efficient citation skills because throughout degree programmes they have to face the task of writing academic papers and failing to comply with the conventions of the academic discourse community in this sense would bring about accusations of plagiarism and negative consequences in the assessment of the subjects.

The proposal combines hands-on and hands-off activities based on a corpus of linguistics research papers and they are aimed at Spanish students in their third year in a degree in Modern Languages. Hands-on activities will let them analyse the corpus in search for integral and non-integral citations as well as reporting verb structures. The teacher's role will be to promote not only their self-sufficiency in the learning process by providing them with an extensive training in the use of AntConc, but also as a discussion stimulator when necessary. Hands-off activities are targeted to work on the same points, but they will somehow help balance the time devoted to the students' operation of the corpus tools, since that may definitely become overwhelming for newcomers to computer software analysis. In any case, this proposal appears to be worth exploring as the gains in learning outcomes may outweigh the disadvantages.

About the author

Elena Quintana-Toledo completed her PhD in English Philology in 2013 at the University of Las Palmas de Gran Canaria (Spain), where she has taught undergraduate and postgraduate courses in English as a Foreign Language and Corpus Linguistics. She has also taught a number of courses and seminars in English for Specific Purposes at the Technical State University of Quevedo (Ecuador). Her research interests include the analysis of English scientific discourse from a pragmatically-oriented perspective, and she has been recently engaged in the study of the relationship between modality and evidentiality, and its pedagogical implications.

Acknowledgements

The research reported here has been funded by the Spanish Ministry of Economy, Industry and Competitiveness as part of the research project “Identification and Analysis of Metadiscourse Strategies in Spanish and English Research Papers” (Grant FFI2016-77941-P). This grant is hereby gratefully acknowledged.

References

- Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology*, 21(3), 195-216.
- Aikhenvald, A. (2006). Evidentiality. In K. Allan (Ed.), *Concise Encyclopedia of Semantics* (pp. 282-287). Amsterdam, Netherlands: Elsevier.
- Aikhenvald, A. (2007). Information source and evidentiality: What can we conclude? *Rivista di Linguistica*, 19(1), 209-227.
- Alonso-Almeida, F. (2012). Sentential evidential adverbs and authorial stance in a corpus of English computing articles. *Revista Española de Lingüística Aplicada*, 1, 15-32.
- Alonso-Almeida, F. (2015a). On the mitigating function of modality and evidentiality. Evidence from English and Spanish medical research papers. *Intercultural Pragmatics*, 12(1), 33-57.
- Alonso-Almeida, F. (2015b). The functions of *seem* and *parecer* in early medical writing. *Discourse Studies*, 17(2), 121-140.
- Ammon, U. (2006). Language conflicts in the European Union. On finding a politically acceptable and practicable solution for EU institutions that satisfies diverging interests. *International Journal of Applied Linguistics*, 16(3), 319-338.
- Anthony, L. (2004). AntConc: A learner and classroom friendly multi-platform corpus analysis toolkit. *IWLeL 2004: An Interactive Workshop on Language e-Learning*, 7-13.
- Anthony, L. (2005). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. *2005 IEEE International Professional Communication Conference Proceedings*, 729-737.
- Aston, G. (2015). Learning phraseology from speech corpora. In A. Lénko-Szymanska & A. Boulton (Eds.), *Multiple Affordances of Language Corpora for Data-Driven Learning* (pp. 65-84). Amsterdam and Philadelphia: John Benjamins.
- Bailey, S. (2006). *Academic Writing. A Handbook for International Students* (Second edition). London and New York: Routledge.
- Baird, R. & Baird, M. (2018). English as a lingua franca: Changing “attitudes”. In J. Jenkins, W. Baker & M. Dewey (Eds.), *The Routledge Handbook of English as a Lingua Franca* (pp. 531-543). London: Routledge.
- Barabadi, E. & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students’ learning of English vocabulary. *Cogent Education*, 4(1). <<https://www.tandfonline.com/doi/pdf/10.1080/2331186X.2017.1283876?needAccess=true>> [19/07/2019]
- Borg, E. (2000). Citation practices in academic writing. In P. Thompson (Ed.), *Patterns and Perspectives: Insights into EAP Writing Practices* (pp. 27-45). Reading: University of Reading.
- Boulton, A. (2009). Data-driven learning: Reasonable fear and rational reassurance. *Indian Journal of Applied Linguistics*, 35(1), 81-106.
- Breeze, R. (2017). Exploring evidential uses of the passive of reporting verbs through corpus analysis. In J. I. Marín-Arrese, G. Haßler & M. Carretero (Eds.), *Evidentiality Revisited: Cognitive Grammar, Functional and Discourse-Pragmatic Perspectives* (pp. 295-311). Amsterdam and Philadelphia: John Benjamins.
- Carretero, M. & Zamorano-Mansilla, J. R. (2014). Annotating English adverbials for the categories of epistemic modality and evidentiality. In J. Marín-Arrese, M. Carretero, J. Arús-Hita & J. Van der Auwera (Eds.), *English Modality: Core, Periphery and Evidentiality* (pp. 317-358). Berlin and Boston: De Gruyter.

- Carrió-Pastor, M. L. (2019). Phraseology in specialised language: A contrastive analysis of mitigation in academic papers. In G. Corpas-Pastor & R. Mitkov (Eds.), *Computational and Corpus-Based Phraseology. Third International Conference, Europhras 2019. Málaga, Spain, September 25-29, 2019. Proceedings* (pp. 61-72). Switzerland: Springer.
- Chafe, W. (1986). Evidentiality in English conversation and academic writing. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology* (pp. 261-272). Norwood, New Jersey: Ablex Publishing Corporation.
- Chapman, R. (2015). The deceiving ELF? Can English really fulfil the role of a lingua franca? *Lingue e Linguaggi*, 15, 113-127.
- Chen, M. & Flowerdew, J. (2018). Introducing data-driven learning to PhD students for research writing purposes: A territory-wide project in Hong Kong. *English for Specific Purposes*, 50, 97-112.
- Chujo, K., Anthony, L., Oghigian, K., & Yokota, K. (2013). Teaching remedial grammar through data-driven learning using AntPConc. *Taiwan International ESP Journal*, 5(2), 65-90.
- Cornillie, B. (2009). Evidentiality and epistemic modality: On the close relationship of two different categories. *Functions of Language*, 16, 44-62.
- Cornillie, B. & Delbecque, N. (2008). Speaker commitment: Back to the speaker. Evidence from Spanish alternations. *Belgian Journal of Linguistics*, 22, 37-62.
- Cosgun, G. & Hasirci, B. (2017). The impact of English medium instruction (EMI) on students' language abilities. *International Journal of Curriculum and Instruction*, 9(2), 11-20.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Cryer, P. (2006). *The Research Student's Guide to Success*. United Kingdom and United States: Open University Press.
- Curado-Fuentes, A. (2017). Form-focused data-driven learning for grammar development in ESP contexts. *Revista de Lenguas para Fines Específicos*, 23(1), 12-30.
- De Haan, F. (2005). Encoding speaker perspective: Evidentials. In Z. Frajzyngier, A. Hodges & D. S. Rood (Eds.), *Linguistic Diversity and Language Theories* (pp. 379-397). Amsterdam and Philadelphia: John Benjamins.
- Faller, M. (2002). *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. Stanford: Stanford University. <<https://personalpages.manchester.ac.uk/staff/martina.t.faller/documents/thesis-a4.pdf>> [10/07/2019]
- Fernández-Martínez, D. & Vizcaíno-Ortega, F. (2018-2019). *Guía docente de Inglés VI*. Facultad de Filología (Grado en Lenguas Modernas). Universidad de Las Palmas de Gran Canaria. <https://www2.ulpgc.es/aplicaciones/proyectosdocentes/pdf.php?id_proyecto=55296&NUEVA=1> [04/07/2019]
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *Teaching English as a Second Language – Electronic Journal*, 8(4), 1-35.
- Granger, S. & Tribble, C. (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer* (pp. 199-209). London and New York: Longman.
- Guilquin, G. & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-369). London: Routledge.
- Hunston, S. (1993). Evaluation and ideology in scientific writing. In M. Ghadessy (Ed.), *Register Analysis: Theory and Practice* (pp. 57-73). London: Pinter.
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341-367.
- Hyland, K. (2002). Directives: Argument and engagement in academic writing. *Applied Linguistics*, 23(2), 215-239.
- Hyland, K. (2004). *Disciplinary Discourses. Social Interactions in Academic Writing*. United States: The University of Michigan Press.
- Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192.
- Hyland, K. (2005b). *Metadiscourse. Exploring Interaction in Writing*. London: Continuum.
- Hyland, K. (2008). Persuasion, interaction and the construction of knowledge: Representing self and others in research writing. *International Journal of English Studies*, 8(2), 1-23.
- Hyland, K. (2013). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 115-130). London and New York: Routledge.

- Jafarigohar, M. & Mohammadkhani, A. (2015). Reporting verbs in applied linguistics research articles by native and non-native writers. *Theory and Practice in Language Studies*, 5(12), 2490-2496.
- Johns, T. & King, P. (1991). Classroom concordancing. *English Language Research Journal*, 4, 27-45.
- Kubota, R. (2003). Striving for original voice in publication? A critical reflection. In C. P. Casanave & S. Vandrick (Eds.), *Writing for Scholarly Publication. Behind the Scenes in Language Education* (pp. 73-83). New Jersey and London: Lawrence Erlbaum Associates.
- Larsen-Walker, M. (2017). Can data driven learning address L2 writers' habitual errors with English linking adverbials? *System*, 69, 26-37.
- Lasagabaster, D. & Doiz, A. (2018). Language errors in an English-medium instruction university setting. How do language versus content teachers tackle them? *Porta Linguarum*, 30, 131-148.
- Lea, M. R. & Street, B. V. (2000). Student writing and staff feedback in higher education: An academic literacies approach. In M. R. Lea & B. Stierer (Eds.), *Student Writing in Higher Education: New Contexts* (pp. 32-46). United Kingdom and United States: Open University Press and Society for Research into Higher Education.
- Lee, J. J., Hitchcock, C. & Casal, J. E. (2018). Citation practices of L2 university students in first-year writing: Form, function and stance. *Journal of English for Academic Purposes*, 33, 1-11.
- Livnat, Z. (2012). *Dialogue, Science and Academic Writing*. Amsterdam and Philadelphia: John Benjamins.
- Lorés-Sanz, R., Mur-Dueñas, P. & Lafuente-Millán, E. (2010). *Constructing Interpersonality: Multiple Perspectives in Written Academic Genres*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Luzón, M. J. (2015). An analysis of the citation practices of undergraduate Spanish students. *Journal of Academic Writing*, 5(1), 52-64.
- Mauranen, A. (2010). Features of English as a lingua franca in academia. *Helsinki English Studies*, 6, 6-28.
- Mithun, M. (1986). Evidential diachrony in northern Iroquian. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology* (pp. 89-112). Norwood, New Jersey: Ablex Publishing Corporation.
- Mori, M. (2017). Using the Appraisal framework to analyze source use in essays: A case study of engagement and dialogism in two undergraduate student's writing. *Functional Linguistics*, 4(11), 1-22.
- Nuyts, J. (2001). Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics*, 33(3), 383-400.
- Pietrandrea, P. (2005). *Epistemic Modality. Functional Properties and the Italian System*. Amsterdam and Philadelphia: John Benjamins.
- Palmer, F. R. (2001). *Mood and Modality* (Second edition). Cambridge: Cambridge University Press.
- Polio, C. & Shi, L. (2012). Perceptions and beliefs about textual appropriation and source use in second language writing. *Editorial/Journal of Second Language Writing*, 21, 95-101.
- Roux, P. (2014). English as an international language: The debate continues. *Polyglossia*, 26, 45-57.
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, J. (2014). Variation in citational practice in a corpus of student biology papers: From parenthetical plonking to intertextual storytelling. *Written Communication*, 31(1), 118-141.
- Tadros, A. A. (1985). *Prediction in Text*. Birmingham: English Language Research.
- Tadros, A. A. (1993). The pragmatics of text averral and attribution in academic texts. In M. Hoey (Ed.), *Data, Description, Discourse* (pp. 99-114). London: Harper Collins.
- Thomas, S. & Hawes, T. (1994). Reporting verbs in medical journal articles. *English for Specific Purposes*, 13, 129-148.
- Thompson, G. & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365-382.
- Thompson, P. (2001). *A Pedagogically-Motivated Corpus-Based Examination of PhD Theses: Macrostructure, Citation Practices and Uses of Modal Verbs*. Unpublished doctoral dissertation. Reading: University of Reading.
- Thompson, P. (2005). Points of focus and position: Intertextual reference in PhD theses. *Journal of English for Academic Purposes*, 4(4), 307-323.

- Thompson, P. & Tribble, C. (2001). Looking at citations: Using corpora in English for Academic Purposes. *Language Learning & Technology*, 5(3), 91-105.
- Van der Auwera, J. & Plungian, V. (1998). Modality's semantic map. *Linguistic Typology*, 2, 79-124.
- Vandermensbrugge, J. (2004). The unbearable vagueness of critical thinking in the context of the anglo-saxonisation of education. *International Education Journal*, 5(3), 417-422.
- Vu, N. T. T. & Burns, A. (2014). English as a medium of instruction: Challenges for Vietnamese tertiary lecturers. *The Journal of Asia TEFL*, 11(3), 1-31.
- Willet, T. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language*, 12(1), 51-97.
- Yilmaz, E. & Soruç, A. (2015). The use of concordance for teaching vocabulary: A data-driven learning approach. *Procedia – Social and Behavioral Sciences*, 191, 2626-2630.