Teaching and learning stance in specialized English for computing.

A Corpus Linguistics Approach

Francisco Alonso Almeida[1,2], Belén López Brito[3], – Elena Quintana Toledo[1]

[1] Instituto para el Desarrollo Tecnológico y la Innovación en Comunicaciones
Universidad de Las Palmas de Gran Canaria
Campus de Tafira
Las Palmas de Gran Canaria
35017 - Las Palmas (Spain)

[2] Grupo de Investigación "Las Matemáticas Avanzan"
Despacho 23 - Facultad de Filología – c/ Pérez del Toro, 1
E35004 - Las Palmas de Gran Canaria
Universidad de Las Palmas de Gran Canaria

[3] Grupo de Investigación "Las Matemáticas Avanzan"
Departamento de Matemáticas
Universidad de Las Palmas de Gran Canaria
Campus de Tafira
Las Palmas de Gran Canaria
35017 - Las Palmas (Spain)

francisco.alonso@ulpgc.es

Abstract. The use of large corpora in the study of languages is a well established tradition. In the same vein, scholarship is also well represented in the case of the study of corpora for making grammars of languages. This is the case of the COBUILD grammar and dictionary and the case of the Longman Grammar of Spoken and Written English. This means that corpora have been analyzed in order to identify patterns in languages that can be later practised by learners following those patterns described and exemplified with real instances. The way in which we approach the use of corpus linguistics and teaching is rather different but it is by no means new. We want to address the way in which large corpora of specialized languages can be used in the classroom to develop language skills, namely *writing, speaking, listening, reading* and *interacting*

ISBN: 978-84-617-1860-3

with a focus on the use of stance language, i.e. the expression of point of view. Interaction stands as the most problematic of the five language skills because, while the rest show a long tradition in second language acquisition studies, interaction has been only recently under focus. This said, our objective is to offer directions of use of Corpus Linguistics (CL) in second language acquisition, especially in the teaching of ICT English. We hypothesize that CL may provide learners with a better framework for the study of specific varieties through interrogation and analysis of those corpora. In this paper, we argue that CL helps to learn the phraseology concerning stance.

Keywords. English learning, English teaching, specialized English, corpus linguistics, stance language

1 Introduction

During the last decade, there has been a growing interest in incorporating discourse elements other than lexical in the syllabus of ICT courses, a big deal concerned with the field of mathematics, in particular, and ESP courses in general in Spain. Thus, in the last decade, students were very well prepared in aspects related to the stock of technical vocabulary of specific disciplines. This was also complemented with the teaching of textual genres, such as the curriculum vitae and the letter, which put an emphasis on job seeking rather than the professional and the academic side of the discipline. The situation has fortunately changed, and aspects concerning the learning and the practice of those academic and research genres (the abstract, the scientific article and the project) are now part of the university syllabi.

This involves rhetorical and textual aspects as well as specific linguistic features. Among these features, we are interested in stance expressions. To our knowledge, the teaching of stance expressions in Spain includes modal verbs and some metadiscourse devices. Within metadiscourse, hedging stands as the most favorite device to show the authors's negotiation of meaning with their readers. A related device is evidentiality. This refers to the expression of the authors's source or mode of information. This is frequently associated with epistemicity to show the lack of commitment towards the proposition manifested. In this paper, we will show that this relationship between epistemicity and evidentiality does not always hold. This has some implications for the teaching of these two important aspects of scientific discourse, and we will suggest some ways in which we approach them in the language classroom.

The structure of the paper is, as follows. First, we present an overview of the current literature concerning stancetaking, especially evidentiality and epistemic modality, and the way in which these two concepts relate. Finally, we explain the type of

 ISBN: 978-84-617-1860-3

activities that may help lecturers of ICT English to develop the topic in class. The data is taken from *evycorpe*, a multi-register corpus of English scientific papers collected at the University of Las Palmas de Gran Canaria.

2 The expression of stance

Stance has been defined as the expression of speakers or writers' "personal feelings, attitudes, value judgements, or assessments" [1]. It has been studied from different angles, and thus the concept seems to be interpreted as an umbrella term to refer to diverse interpersonal language strategies which reveal the author's position with respect to their text. Englebretson's claim [2] that "stance is by no means a monolithic concept" is evinced in the number of multiple approaches in which the speaker or writers' position to their texts are looked at. Thus, stance covers the study of *evaluative language* [3, 4], *evidentiality* [5], *affect* [6], and *hedging* [7, 8], among others.

Stance strategies are semantically classified into *epistemic stance*, *attitudinal stance* and *style of speaking stance* [1]. Epistemic stance is concerned with "comments on the status of information in a proposition"; attitudinal "stance markers… report personal attitudes or feelings"; style of speaking stance includes "comments on the communication itself". The linguistic marking of stance can be realised in several ways, namely lexical marking, grammatical marking, and person marking [1]. Whereas lexical marking relies on evaluative lexical items, grammatical marking consists in the use of stance grammaticalized phenomena, and this includes, according to Biber et al., (a) stance adverbials, (b) stance complement clauses, (c) modals and semi-modals, (d) stance noun and prepositional phrase, and (e) premodifying stance adverbials (stance adverb + adjective or noun phrase). Person marking implies the study of the use of pronouns, the passive voice, and indirect speech, among other linguistic items, as suggested in Precht [9].

Marín-Arrese [10], following Langacker's [11], divides the domain of stance into two main types, namely effective and epistemic. She thereby attempts to identify the language strategies used to register the position of the writer/speaker in relation to their texts. Her proposal not only reconciles controversial issues, such as the relationship between epistemic modality and evidentiality, but also facilitates the analysis of other interpersonal devices, such as the expression of subjectivity and intersubjectivity. Epistemic modality and evidentiality are differently treated in the literature, since evidentiality is seen either as a subdomain of epistemic modality or as an independent category. Whereas the first perspective refers to evidence as a sign of authorial commitment towards the truth of the proposition manifested, the second one shows no connection between evidence for the proposition and its truth. Marín-Arrese [10] clearly considers these two concepts as methodologically distinct even when there might be pragmatic implications concerning the authors's stance towards their texts.

2.1 Effective stance

Effective stance is defined as follows: "Effective stance pertains to the realization of events and situations, to the ways in which the speaker/writer expresses the neces-

ISBN: 978-84-617-1860-3

sity or possibility of the event occurring, or his/her inclination, decision or intention to carry out an event, or his/her emotive/affective position with regard to the event" [10]. According to this author, this type of stance may be realized by the set of linguistic resources, below; examples are taken from the same source (descriptions of categories follow).

- Deontic modality: And we *must* all act against the presence of Al Qaeda in Iraq.
- Volitive modality: I *will not* be party to such a course.
- Participant-internal and participant-external possibility: Those who have fallen behind in the basics *can* use the time for extra literacy or numeracy lessons.
- Participant-internal and participant-external necessity: But now as we develop what will be our 10 year Children's Plan we *need* to move to the next stage in the transformation of standards in education in Britain.
- Attitudinal expressions: It will now be presented to the parties as Abu Mazen is confirmed in office, *hopefully*.
- Communicative evidentials: *I urge* all nations to implement the International Compact to renew Iraq's economy, to participate in the Neigh…
- Imperative mood: And *let* nobody say that Academies aren't helping the poorest children, when a third of those attending…

2.2 Epistemic stance

Epistemic stance "refers to the knowledge of the speaker/writer regarding the realization of the event and/or to his/her assessment of the validity of the proposition designating the event" [10, 12]. She has included several evidential strategies in the epistemic stance domain. The concepts of epistemic modality and evidentiality are methodologically treated as two distinct categories, but not as a general policy [13]. The former is related to the expression of truth concerning the proposition manifested (cf. Halliday [14]), whereas the latter is related to the mode of knowing in the sense described in Cornillie [15]. This allows these two concepts to be identified, while avoiding confusing approaches, which consider evidentiality simply as a subdomain of epistemic modality within propositional modality [16]. The four strategies in this group are:

- Epistemic modality: …talking to an Iraqi exile and saying to her that I understood how grim it *must* be under the lash of Saddam.
- Communicative evidentials: *I have said* education is my passion.
- Experiential evidentials: So, when *I see* the superb work being achieved at this Academy, I believe we should also…
- Cognitive evidentials: And *this is my belief*: that world class performance come from consistent brilliance from teachers in every classroom.

As we have already pointed out, epistemic modality is closely associated to the idea of certainty, and so epistemic devices are used to show differing degrees of au-

 ISBN: 978-84-617-1860-3

thorial stance concerning propositional truth-values depending on the owned evidence.

The second category of epistemic stance is communicative evidentials. This category refers to expressions of self-reference and self-attribution with respect to knowledge and the degree of certainty regarding the proposition they accompany, as in *I have said* above. Experiential strategies report on how information has been gained through senses, including inferential reasoning by the observation of the evidence. Instances of this category are: *I see, We witnessed, It appears…, That shows...* [10]. Cognitive evidentials include cognitive verbs and expressions representing mode of knowing, such as *This is my belief, We think, I consider…*

While all the devices described by Marín-Arrese give a clear picture of the position of authors with respect to their texts, we will concentrate on aspects concerning knowledge rather than desires. The complete picture would be just too much for ICT students, and we take the risk of making our class too linguistic and technical for them to follow.

3 Teaching epistemic modality and evidentials

One way in which mitigation of claims is expressed is by means of hedges. Hedges and evidentials constitute one of the most problematic areas for students. While we keep telling them that scientific language should be free of bias, and objectivity is reflected in the use of language, we also tell them that there are ways in which authors both mitigate and disguise their claims in order to prevent possible future face threatening acts. We also insist that mitigating a claim is not necessarily mitigating its truth. In many aspects, mitigating a claim might reflect decorum and professional culture, and it changes from one discipline to another. It is also a cultural aspect. In previous studies we have shown that English speakers are more prone to use epistemic modality and Spanish speakers tend to use evidential devices.

A possible reason for this is the Spanish tendency to assert claims even when these claims are fruit of rumours and hearsay. This is also seen by many as a way to protect their public self-image. However, using evidentials is not necessarily a mitigation of the propositional content it accompanies. In a sentence like *Mary told me she repaired the computer*, the speaker is not really mitigating the proposition 'she repaired the computer'; placing the communicative evidential expression *Mary told me* means to indicate third party authority, but also third party responsibility for the claim.

The distinction among mitigating a claim, responsibility and commitment requires some extra effort from students, which are faced with real examples to detect possible uses of devices showing any of these three aspects. A first attempt is guided and students work with the teacher to disambiguate possible implications obtained from the use of evidentials and epistemic devices in selected sentences. This comes after a short presentation of the concepts in which technicalities have been kept to a minimum. After this initial exercise, students are asked to identify possible evidential and epistemic structures in real texts taken from the *evycorpe* database. All students have the same text and they can share their opinions concerning their categorisation of the

ISBN: 978-84-617-1860-3

detected devices. At this point, it is good that while correcting the activity the students give their view with respect to the way in which the authors use evidentiality and epistemic modality to develop their text. This part is followed by specimens of English texts written by Spanish speakers to see whether cultural differences emerge. This type of activities allows students to become aware of the real differences between both users of the language, the native speaker and the Spanish speaker of English.

The creative process starts once students understand both the concepts of evidentiality and epistemic modality (and related notions), and their pragmatic implications. The first task has to do with a set of matrices, adverbials and modals serving as evidentials or epistemic devices. These must be correctly placed in sentences, which correspond to given situations. Another exercise is transformation, and so students are given sentences and they must transform them to indicate either evidence for the claim made or differing degrees of commitment.

Because our students are first year, we consider this approach fulfils their communicative needs. It is essential however that they identify and detect these devices in real texts so that they can correctly interpret the messages underneath a given sentence, and also are able to codify mitigation and assertion in a variety of ways following the English-culture tradition.

4 Examples of corpus-based activities to teach and learn stance

A corpus of scientific papers like *evycorpe* is useful insofar it allows the search of stance formulas to students of specialized English. To put it simple, these formulas can be categorized as (a) fixed formulas and (b) more flexible formulas although the range of categories may be expanded to detail in order to include specific types of formulaic language. Wray and Perkins define formulaic language as ‘a sequence, continuous or discontinous, of words or other meaning elements, which is, or appears to be prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar’ [17]. This definition is in line with the category of fixed formulas.

The definition does not entirely exclude dynamic templates, which may be accommodated to needs, as Wray and Perkins [17] argue, in the light of the evidence put forward in Pawley and Syder [18]. The category of fixed formulas includes stance devices, such as *presumably, in my opinion, arguably, safely*, among others. The second category comprises language that can be adapted to fit a particular context and communicative needs. This is the case of matrices of the *think*-type, e.g. *I think that…, I thought that…, I consider that…,* as well as inferential devices of the *seem*-type, such as *it seems that…*

Another important issue in learning formulaic stance is function. In this line, students may also collect cases to illustrate their interactional intention. Students may learn different ways in which they may convey degrees of probability and possibility in discourse and they may also learn how to indicate academic modesty and polite behaviour to accommodate the standards of scientific language. A corpus based approach is very convenient in this case since students may gather, store and learn por-

 ISBN: 978-84-617-1860-3

tions of real language that they can find potentially useful in similar communicative situations.

4.1 Activities

This section presents some activity types that focus on learning stance language using a methodology based on corpus linguistics. The examples used for illustration are excerpted from *evycorpe*.

4.1.1. Concordance exercises

These activities require that students have learnt to use corpus tools in order to inquiry the compilation and retrieve the information requested.

— *Find out cases of the form* seem *in the corpus of computing texts and describe frequency of co-ocurrence*.
- Examples from *evycorpe*:
  - This **seems to contradict** the fact
  - …this **seems to create** few problems
  - …which, thus far, **seems to have** escaped attention
  - … the technique **seemed to have** no discernible effect
  - … the MSE **seems to have** very little
  - Although the diagram **seems to imply** it should
  - These simulations **seem to indicate** that
  - Fig. 10(c) **seems to indicate** that…
  - Fig. 10(d) **seems to indicate** that…
  - …these flat regions **seem to occur** near the global maximum
  - This **seems to perform** quite well on a simple three-source separation problem

4.1.2. Cloze texts and context analysis

These activities are aimed to practice specific uses of language according to context and communicative needs. In the example given below, the student is asked to fill in the gaps using either the modal *can* or the modal *may*. Afterwards, the student has to interrogate the corpus for these specific examples to get the correct answers. The last instruction concerns the search of a manageable size of *may* and *can* in the corpus so that students may understand how native speakers use these modals following an inductive reasoning.

— *Give* may *or* can *in the blank spaces in the instances below:*
- Intrinsic reconfiguration _____ be used for fault recovery so long as it finishes before the mandatory recovery deadline.
- In some cases, it _____ be possible to do the recovery in stages rather than all at once.

ISBN: 978-84-617-1860-3

- Most of the functions prefer an exemplar without the tongue. This _____ be because of the high contrast between pixels projected dimly by the inside of the mouth and those projected brightly by lip and tongue.
- Once we have the approximate probability distributions for the bit assignment updates, we _____ calculate the probability of making an erroneous update at the site indexed.
- Such a visual resemblance _____ suggest a folding and stretching of the state space happening in our case due to repeated updating and normalization described by the equilib- rium model.
- From a computational standpoint, the memory _____ be used to iteratively recall previously stored binary patterns when pre- sented with their perturbed or noisy variants as input.
- To ensure that the LSD _____ respond to illumination changes at 15 frames per second (fps), however, we have used some additional optimizations.
- These statements _____ show logical correctness, but without comparing the reconfiguration time against a deadline there is no proof of temporal correctness.

— *When you have finished, search evycorpe for these examples and check your answers*.

— *Now, interrogate the corpus for the terms* may *and* can *and choose* 10 instances of each. Analyse them and rule out how writers use these modal verbs.

4.1.3. Frequency activities

Detecting how frequent a particular word or expression is in scientific English makes students aware of how important it is the notion of idiomaticity when they learn the language. Students very often tend to translate chunks of information from their language without noticing that, even if they are highly proficient in the target language, there are occasions in which linguistic productivity does not work for a variety of reasons to the extent that a translated text can result in a meaningless chunk in English. In the case of specialized English, there are certain discourse conventions that pertain to each field of knowledge but not to others.

— *Search the corpus and find out ways in which authors express their point of view. Use keywords, such as* opinion, think, consider *and* believe, *to retrieve examples. Make a list of these expressions from the most frequent to the least frequent.*
  - Examples:
    - **We believe that** information about the likelihood of false positives, that is erroneously believing that we have a prediction system, also warrants investigation.
    - **It may be thought that** hypervolume is not, after all, a practical idea as a metric for comparing fronts.
    - **We assume that** the processing times are negligible.
    - **We conjecture that** this algorithm will find such nonnegative well-grounded independent sources, under reasonable initial conditions. While the algorithm has proved difficult to analyze in the general case, we give some analytical

 ISBN: 978-84-617-1860-3

results that are consistent with this conjecture and some numerical simulations that illustrate its operation.

5 Conclusion

This paper has addressed the applicability of CL to ESL by focusing on aspects related to teaching and learning those stance elements typically found in scientific literature. In order to do so, we rely on Marín Arrese's proposal for the categorization of stance into effective and epistemic strategies. This model stands as a clear one which students who are not familiarized with linguistic terminology may find easy to understand and apply. A corpus-based methodology allows students to deal with authentic language samples so that at the end of the course they are able not only to correctly understand the epistemic and evidential uses of elements such as modal verbs, but also to put them into practice in a given communicative situation. We finally suggest some activities for students to learn and practice the use of stance formulas as they occur in *evycorpe*.

References

1. Biber, D.; Johansson, S.; Leech, G.; Conrad, S.; Finegan, E.; Longman Grammar of Spoken and Written English. Pearson Education, Ltd, Harlow, Essex, 1999
2. Englebretson, R.; Stancetaking in Discourse: Subjectivity, Evaluation, Interaction. John Benjamins, Amsterdam, 2007
3. Hunston, S.; "Evaluation and organisation in a sample of written academic discourse", Couthard, M. (ed.) Advances in written text analysis, Routledge, London, 1994; 191-218
4. Hunston, S.; Thompson, G.; Evaluation in text. Oxford University Press, Oxford, 2000
5. Anderson, L.B.; "Evidentials, paths of change and mental maps: typologically regular asymmetries", Chafe, W., Nichols, J. (eds.) Evidentiality: The Linguistic Coding of Epistemology, Ablex, New Jersey, 1986
6. Ochs, E.; "The Pragmatics of Affect". Special Issue of *Text*. Mouton de Gruyter, 1989
7. Hyland, K.; "Nurturing hedges in the ESP curriculum". System 24, 1996; 477-490
8. Hyland, K.; Hedging in scientific research articles. John Benjamins, Amsterdam, 1998
9. Precht, K.; Patterns of stance in English. PhD. Northen Arizona University, 2000
10. Marín Arrese, J. I.; "Effective vs. Epistemic Stance, and Subjectivity/ Intersubjectivity in Political Discourse. A Case Study", Tsangalidis, A.; Facchinetti, R. (eds.) Studies on English Modality. In Honour of Frank R. Palmer, Peter Lang, Bern, 2009; 23-52
11. Langacker, R. W.; "Control and the Mind/body Duality: Knowing vs. Effecting", Tabakowska, E.; Choinski, M.; Wiraszka, L. (eds.) Cognitive Linguistics in Action. From Theory to Application and Back, Mouton de Gruyter, Berlin, 2010; 165-207

ISBN: 978-84-617-1860-3

12. Marín-Arrese, J.I.; "Evidential and epistemic qualifications in the discourse of fact and opinion: A comparable corpus study", Marín-Arrese, J. I. (ed.) Perspectives on Evidentiality and Modality. Editorial Complutense, Madrid (2004)

13. Kranich, S.; "Epistemic modality in English popular scientific texts and their German translation". *Trans-kom* 2, 2009; 26-41

14. Halliday, M.A.K.; "Functional diversity in language, as seen from a consideration of modality and mood in English". Foundations of Language 6, 1970; 322-361

15. Cornillie, B.; "Evidentiality and epistemic modality: on the close relationship of two different categories". Functions of language 16, 2009; 44-62

16. Palmer, F.; Modality and the English modals. Longman, London, 2001

17. Wray, A.; Perkins, M.R.; "The functions of formulaic language: an integrated model". Language & Communication 20, 2000; 1-28

18. Pawley, A., Syder, F.H.; "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency", Richards, J. C.; Schmidt, R. W. (eds.) Language and Communication, Longman, London, 1983; 191-225

 ISBN: 978-84-617-1860-3