Elastic language in scientific writing: Evidence from the *Corpus of English Life Sciences Texts*

FRANCISCO J. ÁLVAREZ-GIL
Universidad de Las Palmas de Gran Canaria

ELENA QUINTANA-TOLEDO
Universidad de Las Palmas de Gran Canaria

Abstract
This paper sets out to explore the expression of vagueness in eighteenth and nineteenth century scientific writing, specifically in the *Corpus of English Life Sciences Texts*. Following Zhang's (2015) pragmatic-oriented approach to vague language, vague expressions are viewed through the lenses of elasticity. This notion applies to the strategic use of vagueness insofar as it provides a space for the negotiation of pragmatic meanings in writer-reader interaction. The exploration of elastic language involves four lexical categories: (i) approximate stretchers, i.e., approximators and vague quantifiers, (ii) general stretchers, i.e., general terms, placeholders and vague category markers, (iii) scalar stretchers, i.e., intensifiers and softeners, and (iv) epistemic stretchers. The *Coruña Corpus Tool* has been used to elicit data, but classification and contextualised interpretations have necessarily relied on manual analysis. Findings reveal that elastic language fulfils a variety of relational functions in scientific writing, including making generalisations when the information is either not available or relevant for the purposes of communication, marking shared knowledge and group membership, or self-protection.
Key words: vagueness, elastic language, approximate stretcher, general stretcher, scalar stretcher, epistemic stretcher, scientific writing

1. Introduction

Vagueness is pervasive in virtually all forms of human communication and a widely acknowledged core feature of natural languages (Channell 1994, Sainsbury 1996, Fairclough 2003, Van Rooij 2011, McCarthy 2020). Perfect or nearly perfect accuracy has typically been regarded as a highly desirable quality of every single communicative situation. This idea, however, started to be challenged in the 1980s and 1990s by a pragmatic approach which emphasised the dynamic nature of vague language use as a means of ensuring successful interaction between interlocutors (Channell 1994, Cheng & Warren 2001, Ruzaite 2004, Cutting 2007, Murphy 2010, Zhang 2015). According to this view, a lack of precision may be particularly apposite in some contexts when trying to strategically achieve certain communicative purposes.

Channell's (1994) influential work paved the way for systematic research on vagueness by drawing attention to the pivotal role of appropriateness and its relationship with context: whenever we communicate, we can tentatively evaluate the extent to which any vague item would be adequate bearing in mind the

surrounding circumstances or, as Kyburg and Morreau (2000:577) put it, "a speaker can adjust the extension of a vague expression to suit his needs, relying on the hearer to recognize his intentions and to accommodate him". As participants in a round table discussion at a conference, for example, we are probably expected to attain a high degree of precision when commenting on quantitative research. On the other hand, in a pleasant chatter with colleagues during the coffee break, we could reasonably be less precise without causing any communicative breakdown.

The role of vagueness has been analysed in a number of contexts, for instance, Channell (1983) describes vague language use by native speakers of standard English with the aim of proposing a theoretical framework which may be applied not only to other varieties of English, but also to other languages. Her exploration of the semantics/pragmatics interface of vague expressions in conversational data enables her to offer a complete picture of the phenomenon which, in a later work (Channell 1994), is specifically referred to as a part of native speakers' communicative competence. Along the same lines, Urbanová (1999) studies spontaneous telephone conversations and confirms the need to consider the semantic/pragmatics interface of vague language to adequately account for its functions. Research on vagueness in native-speaker language has been extended to written discourse as well. Channell (1990), Myers (1996) and McCarthy (2020) focus on the academic sphere; they highlight how contextual variables shape the strategic use of vague expressions.

The strategic dimension of vagueness has been further raised in the fields of Foreign and Second Language Acquisition. Cheng and Warren (2001:98) state that vagueness use stands as a measure of language proficiency, in particular, of the communicator's "strategic competence" and "sociolinguistic competence" (as understood by Canale & Swain (1980)), as it brings together the effective accommodation of (im)precision and the sociocultural factors of communication. Other works in the same line include Hyland and Milton (1997), Hinkel (2002, 2005), Hasselgreen (2005) or Neary-Sundquist (2013).

When it comes to scientific communication, writers are frequently asked to avoid vague language in order to communicate their work as objectively as possible (Gilbert & Mulkay 2003, Wallwork 2016, Alley 2018); in Towns' (1990) words, "[v]agueness, ambiguity and inability to express clearly and succinctly are intolerable in a scientist". It has not been until relatively recently that some scholars have looked more closely into scientific discourse to confirm not only the presence of vague language, but also, and most importantly, that it can fulfil a "relational" function (O'Keeffe, McCarthy & Carter 2007:159) by facilitating interaction through the negotiation of pragmatic meanings. This line of research has been explored by Myers (1989, 1996), Salager-Meyer (1994, 1997), Hinkel (1997), Banks (1998), Hyland (1996, 1998, 2000), Ruzaite (2004) and Cusen (2019), among others.

From a diachronic standpoint, which is our concern in this paper, the role of vagueness in scientific texts has been scrutinised in some studies. Carroll (2007) and Ortega-Barrera (2012) concentrate on the presence of extenders in Middle and

early Modern English letters and early Modern English recipes, respectively. Both authors aim at establishing the form and function of phrases such as *or the lik*e which occur at the end of lists. In a later study, Carroll (2009) partially adopts Channell's (1990) taxonomy of vague language to analyse the fourteenth century recipe collection *The Forme of Cury*. Textual evidence supports the addition of three extra categories to Channell's (1990) classification, e.g., flexibility, superordinacy and omission. Borrowing Zhang's (2015) notion of elasticity and related categories, Quintana-Toledo (2017) explores vague language in another subcorpus of the *Coruña Corpus of English Scientific Writing*, namely, the *Corpus of History English Texts*. Finally, hedging may well be the most extensively researched pragmatic function of vagueness from a diachronic perspective, notably in the works by Salager-Meyer, Defives & Hamelinsck (1996), Skelton (1997), Salager-Meyer & Defives (1998), Alonso-Almeida (2012), Puente-Castelo & Mónaco (2016), Álvarez-Gil (2018a; 2018b), Álvarez-Gil & García-Alonso (2019) or Moskowich & Crespo (2019).

This study aims to contribute to the diachronic approach to vagueness by exploring the phenomenon in late Modern English scientific writing. To this end, the *Corpus of English Life Sciences Texts*, a subcorpus of the *Coruña Corpus of English Scientific Writing*, will be analysed. The approach taken involves Zhang's (2015) pragmatic-oriented conceptual framework of vagueness, where this notion is preferably dealt with in terms of elasticity. The elastic language categories examined here include (i) approximate stretchers, i.e., approximators and vague quantifiers, (ii) general stretchers, i.e., general terms, placeholders and vague category markers, (iii) scalar stretchers, i.e., intensifiers and softeners, and (iv) epistemic stretchers. The automatic searches of the items in these categories have been conducted using the *Coruña Corpus Tool*, which provides initial identification and quantification. Manual disambiguation has been essential for classification purposes and for the identification of pragmatic functions in the specific context of use because elasticity allows items in these categories to perform functions as varied as making generalisations when the information is not available or even relevant for the purposes of communication, marking shared knowledge and in-group membership, or self-protection.

Based on all the above, the research questions which will guide our analysis are the following:

1. Can the major categories and subcategories of elastic language be attested in eighteenth and nineteenth century *Life Sciences* texts?
2. If so, what pragmatic functions do they fulfil?

In order to meet the main objective and answer the research questions, the paper is structured as follows: section 2 presents the theoretical framework for the analysis including earlier views on vagueness and its conceptualisation in Zhang's (2015) model. Section 3 contains the description of the corpus and the methodology. Section 4 offers a discussion of the findings and the implications drawn from the study are presented in section 5.

2. Vagueness in language

Paradoxical as it may seem, there is no general agreement on how to define the concept of vagueness in language. Urbanová's (1999:99) generic definition, e.g., "a semantic manifestation of indeterminacy" may be taken as a starting point for an overview of previous research on the topic. The truth is that definitions with countless subtle nuances and categorisations have proliferated since Channell's (1994) pioneering analysis on the semantics and pragmatics of vague expressions. Her broad definition reads "an expression or word is vague if: a. it can be contrasted with another word or expression which appears to render the same proposition; b. it is 'purposely and unabashedly vague'; c. its meaning arises from the 'intrinsic uncertainty' referred to by Pierce" (Channell 1994:20). The author further notes that "there are a number of ways in which speakers can avoid being precise or exact" (Channell 1994:17), suggesting a correlation between vagueness and imprecision or inaccuracy. In like manner, Crystal and Davy (1975:11) associate vagueness with "lack of precision" and Cook (2007:21), for his part, uses the term to refer to "the absence of that quality [being 'precise']".

Kempson (1977:124-128) makes a clear distinction between ambiguity and vagueness. On the one hand, ambiguity applies to items which have more than one unrelated meaning; on the other, vagueness denotes "lack of specification", that is, while the meaning of the item is clear, it is only generally specified. The term *neighbour* exemplifies this lack of specification, for example, as for "sex, […] race or age, etc. It can be applied to people as disparate as a tiny, five-foot Welshman studying Philosophy, and a six-foot Ghanaian girl who has seven children and who only did four years of schooling" (Kempson 1977:125). Drave's (2002) and Cheng's and Warren's (2001, 2003) conceptualisation relies on "non-specificity" because it is an integral quality of vague language which, after all, "is non-specific regardless of the context in which is it uttered". In addition, Cheng and Warren (2001) allude to "inexplicitness" in their attempt to disentangle the confusion over vagueness. According to them, inexplicit expressions "achieve specific meaning from the negotiation of context between participants in conversation" (Cheng & Warren 2003:382). Ellipsis, deixis and reference, and substitution are considered standard cases of inexplicitness: whenever speakers use these forms, they do not intend "to be imprecise or ambiguous"; they are simply "combining language and context to convey her or his meaning in an inexplicit form in the expectation that the hearer can assign an unambiguous meaning to it" (Cheng & Warren 2003:391-392).

Explicitness and implicitness are also included in the inventory of interrelated concepts. Scholars hold dissimilar views in this respect, for instance, Koetser (2007:41) considers that vagueness is just the opposite of explicitness. Cutting (2007:4) takes vagueness and implicitness as overlapping to the extent that implicitness "can be expressed with [vague language] and other language features" and, at the same time, "[vague language] can express implicit meaning but it can be taken as its face value". Even other terms like "loose talk" (Sperber & Wilson 1991:540) or "lack of commitment" (Stubbs 1996:2002) have been used in

discussions about vagueness. Fortunately, despite this general terminological disagreement, Cheng and Warren (2003:384) notice that "the realizations of vagueness are more consistent across the various studies".

With regard to categorisations of vague language, Channell's (1994) classification is one of the earliest and, consequently, one of the most frequently quoted. Vague language items may fall into one of the following three categories: (i) approximators, (ii) vague category identifiers, and (iii) placeholder words. Approximators are items which express an estimated amount or quantity. They may be used with (i) numbers, e.g., *three or four books*, (ii) approximators properly speaking, e.g., *about/approximately/around five hundred people*, (iii) round numbers, e.g., *to be six feet tall*, (iv) faded numbers, e.g., *a couple of chairs*, and (v) partial specifiers, e.g., *at least twenty minutes*. Non-numerical quantifiers (positive, negative or neutral) are also approximators, e.g., *lots of*, *a bit* or *some*. Vague category identifiers comprise expressions such as *and things like that* and *or stuff like that*. They have a specific structure, e.g., *exemplar + tag* as in *I'll have some coffee or something like that*, where *coffee or something like that* refers to "any hot beverage". In vague category identifiers of this sort, the exemplar always comes first. Placeholders, which are normally used as pro-forms for persons and objects and less frequently for events or locations, are expressions like *thingy* and *whatshisname*.

Later classifications have been strongly inspired by Channell (1994), but in most of these the original categories have been rearranged and labelled differently. A case in point is Cotterill (2007:99), who uses the label "vague additives" to deal with both approximators and vague category identifiers. Placeholders and other vague quantifiers belong to "vagueness through lexical choice". Vague quantifying expressions with round numbers pertain to the last category which is referred to as "vagueness by implicature". Jucker, Smith and Lüdge (2003) rework Channell's (1990) proposal taking a more interactional approach which results in a four-type system: (i) representations of people and places which take the form of referential vagueness, e.g., the pronoun *it*; (ii) assigning events and experiences, which involves vague category identifiers, placeholders and downtoners, e.g., *a bit*, *kind of*, *sort of*; (iii) representations of amounts, frequencies and probabilities, where approximators, vague quantifying expressions, adverbs of frequency (e.g., *sometimes*, *usually*) and vague adverbs of likelihood (e.g., *maybe*, *probably*) are incorporated; and (iv) propositional attitudes, where items expressing lack of commitment are considered, e.g., adverbs like *presumably*, modal verbs like *may*, and parentheticals like *I think*.

The inclusion of the notions of possibility/probability and lack of commitment as implemented in Jucker, Smith and Lüdge (2003) represents a natural step in the development of the concept of vagueness, at least if we aim at having a more complete picture of this linguistic phenomenon. The resulting taxonomy is less restrictive than Channell's (1990), but it seems to have certain methodological limitations regarding functional overlap with categories, specifically, vague adverbs of likelihood in category (iii) and adverbs expressing lack of commitment

in category (iv), both of which can be certainly decisive in the derivation of propositional attitudes. Apart from that, overlapping categories stand as a potential threat when quantifying items in the analysis of any corpus.

In this study, a broader definition and a less limiting approach to vagueness are preferred so as to get a more comprehensive perspective on its strategic uses. Endorsing Trappes-Lomax's (2007) ideas, our working definition of vagueness covers "any purposive choice of language designed to make the degree of accuracy, preciseness, certainty or clarity with which a referent or situation (event, state, process) is described less than it might have been" (Trappes-Lomax's 2007:122). Concerning a less narrow approach, Zhang's (2015) views on vague language provide us with a solid theoretical framework for analysis. The author herself makes a distinction between conventional and liberal approaches to linguistic vagueness. Investigations following the conventional method are heavily influenced by the work of Channell (1990) and hence focus on vague quantifying items, vague category identifiers, placeholders and so on. Those following the liberal one, including Zhang (2015), take into account the items considered in the conventional approach plus others which are essentially epistemic.

Zhang's (2015:49-65) system of vague language is conceived as part of elasticity theory. The notion of elasticity refers to "the springy nature of language that makes it able to adjust readily to different contexts and communicative goals" (Zhang 2015:5), and it applies to those linguistic units which have "an unspecified meaning boundary, so that its interpretation is elastic in the sense that it can be stretched or shrunk according to the strategic needs of communication" (Zhang 2013:88). In other words, the interpretation of vague language items is always subject to the communicative needs of the situation, which will cause them to get negotiated throughout interlocutors' interaction. In elasticity theory, vague language interpretation is, in essence, a mutual negotiation process and elasticity itself is a strategic feature of vagueness. This is certainly not at odds with the view of vagueness as an interactional strategy supported by Jucker, Smith and Lüdge (2003).

Elastic language is governed by three interconnected principles, namely, fluidity, stretchability and strategy. Zhang (2015:57) spells them out as follows:

> Fluidity principle: The meanings of utterances are non-discrete, overlapping, context-dependent but context-irresolvable: that is, elasticity is a matter of degree, and EL [elastic language] is gradual rather than abrupt.
> Stretchability principle: Utterances can be stretched in various ways, and how far to stretch is governed by communicative needs. Appropriate stretching assures effective communication.
> Strategy principle: Fluid utterances are employed primarily to serve strategic purposes, performed through their pragmatic functions.

The typology of elastic language described in Zhang (2015:35-37) encompasses four lexical categories: (i) approximate stretchers, (ii) general stretchers, (iii) scalar stretchers and (iv) epistemic stretchers. Approximate, general and scalar stretchers

are content-focus items, whereas epistemic stretchers are stance-focus devices. Approximate stretchers are numerical and non-numerical approximators as well as quantifiers expressing inexact quantities, e.g., *about*, *a lot*, *a few*, *many*, *some*. General stretchers are "expressions with limited semantic specificity" (Zhang 2015:36), like general terms, placeholders and vague category markers, e.g., *something*, *thingy*, *and things like that*. Scalar stretchers include softeners and intensifiers, e.g., *a bit*, *quite*, *really*, *so*. Lastly, epistemic stretchers are expressions of the speaker's uncertainty and lack of commitment to the propositional content, e.g., *could*, *possible*, *probably*, *I guess*. It should be noted that there are endless ways in which elastic language may be linguistically manifested. Besides, due to the polypragmatic character of many elastic items, their classification and interpretation are not always straightforward (Zhang 2015:75). However, this taxonomy makes empirical research feasible in principle.

The following section offers an in-depth analysis of the use and pragmatic functions of vagueness as manifested through elastic language in the *Corpus of English Life Sciences Texts*.

3. Corpus description and methodology

The data for this study have been drawn from the *Corpus of English Life Sciences Texts* (henceforth CELiST), one of the four already released subcorpora of the *Coruña Corpus of Scientific Writing* (henceforth CC). CC has been compiled by the Research Group for Multidimensional Studies in English at the University of A Coruña and it stands as a powerful resource for diachronic research on scientific writing in English.

CELiST consists of forty samples of texts belonging to the discipline of Life Sciences, e.g., the scientific study of living organisms (human beings, animals and plants) covering topics like anatomy, biology, entomology or zoology. The first and the last eighteenth-century texts in CELiST date back to 1707 and 1795. The earliest text in the nineteenth century was published in 1804 and the last one was released in 1898. The complete list of the texts analysed is given in Appendix 1.

The texts are representative of scientific writing in the eighteenth and nineteenth centuries. Since representativeness and balance are two of the main principles which guide the compilation of all the subcorpora in CC, all samples in CELiST roughly contain the same number of words, around 10,000. In this sense, Biber (1993:249) argues that even smaller samples (1,000 words) can be representative enough of their language features but, taking into account that the level of standardisation of the scientific register in these two centuries was lower than it is at present, the compilers have considered that longer samples (10,000 words) can optimally reflect scientific writing at that time (Monaco & Puente-Castelo 2019:49). The material used here represents more than 400,000 words which are evenly distributed across the two centuries: 200,220 for the eighteenth century and 200,085 for the nineteenth century (Crespo & Moskowich 2020).

As in the other three subcorpora of CC (*Corpus of English Texts on Astronomy*, *Corpus of English Philosophy Texts* and *Corpus of History English Texts*), the texts

are classified according to genre or communicative format features (Crespo & Moskowich 2020:671). The genres represented in CELiST are treatises, essays, catalogues, lectures, textbooks, articles, guides and letters. It should be noted that the genres of catalogue and guide are only found in CELiST. In relation to author variables, geographical distribution and sex have been taken into account. The authors, both male and female, have varied linguistic and educational backgrounds, which may be linked to England, Scotland, Ireland or North America (Crespo & Moskowich 2020:673-679).

The *CC Tool* (Parapar & Moskowich-Spiegel 2007, Barsaglini-Castro & Valcarce 2020) is a query software which allows researchers to manage and exploit the CC by means of a concordancer. The identification of individual cases of elastic language is not always a straightforward task, not only because the notion of elasticity may be slippery at times, but also because it is liable to embrace many linguistic manifestations. As a result, we have decided to adopt the common approach of searching for a predetermined set of lexical items, which has been borrowed from Zhang (2015) in an attempt to facilitate the initial identification and quantification of elastic items (see Appendix 2). As Hickey (2003:11) rightly notes, "corpus interrogation can rarely be completely automatic and left to the computer". For that reason, manual disambiguation has been an important step in the adequate assessment of the semantics and pragmatics of elastic language.

4. Analysis and discussion

This section deals with the data obtained from the interrogation of CELiST after manual disambiguation was performed. As stated earlier in this paper, the categories and subcategories used for data classification cover (i) approximate stretchers, i.e., approximators and vague quantifiers, (ii) general stretchers, i.e., general terms, placeholders and vague category markers, (iii) scalar stretchers, i.e., intensifiers and downtowners, and (iv) epistemic stretchers. A discussion of relevant examples is offered afterwards, where elastic stretchers are analysed in relation to their pragmatic function in scientific writing. Source details, e.g., year and author, are provided between round brackets.

Table 1. Elastic language categories in CELiST.

Category	Total items	Items per 1,000 words	% of total elastic language
Approximate stretchers	**4049**	**10.1**	**44.4**
Approximators	711	1.7	7.8
Vague quantifiers	3338	8.3	36.5
General stretchers	**223**	**0.5**	**2.4**
General terms	202	0.5	2.2
Placeholders	0	0	0
Vague category markers	21	0.05	0.2
Scalar stretchers	**2726**	**6.8**	**29.8**
Intensifiers	2473	6.1	27.07
Softeners	253	0.6	2.8
Epistemic stretchers	**2135**	**5.3**	**23.3**

Table 1 summarises the data as for the occurrence of individual items of elastic categories and subcategories in the corpus in terms of raw figures, items per 1,000 words and distribution in percentages. Findings reveal very similar results to those obtained in Quintana-Toledo (2017) for the *Corpus of History English Texts*. Approximate stretchers outnumber the rest of the categories with 4049 items, accounting for nearly half of the total amount of tokens with 44,4%. The occurrence of scalar stretchers parallels that of epistemic stretchers with a slight difference of 6,5%. The results here contrast with those in Quintana-Toledo (2017) for the same elastic markers, where the frequency of epistemic stretchers is higher than that of scalar stretchers. Disciplinary variation might account for these differences, at least as regards the expression of epistemicity: Life Sciences texts are characterised by providing very detailed descriptions of natural elements and/or phenomena, frequently to try to come up with classifications. These descriptions are mostly based on observable, factual information, allowing for more categorical assertions. History texts, on the other hand, are more narrative and authors might be more prone to expressing their opinions in the form of epistemic judgements. The least frequent elastic category is general stretchers in which we have only registered 223 items.

Table 2. Elastic language categories in CELiST across the 18th and 19th centuries

Category	Total items		Items per 1,000 words		% of total elastic language	
	18th c.	**19th c.**	**18th c.**	**19th c.**	**18th c.**	**19th c.**
Approximate st.	**2109**	**1940**	**10.5**	**9.6**	**46.4**	**42.4**
Approximators	376	335	1.8	1.6	8.3	7.3
Vague quantifiers	1733	1605	8.6	8.02	38.1	35.1
General st.	**111**	**112**	**0.5**	**0.5**	**2.5**	**2.4**
General terms	95	107	0.4	0.5	2.1	2.3
Placeholders	0	0	0	0	0	0
Vague category	16	5	0.07	0.02	0.4	0.1
Scalar st.	**1553**	**1173**	**7.7**	**5.8**	**34.1**	**25.6**
Intensifiers	1416	1057	7.07	5.2	31.1	23.06
Softeners	137	116	0.6	0.5	3.01	2.5
Epistemic st.	**778**	**1357**	**3.8**	**6.7**	**17**	**29.6**

Table 2 shows the occurrence of elastic categories across the 18th and 19th centuries. Approximate and scalar stretchers seem to be evenly distributed across the two centuries, and general stretchers reveal almost the same results. The occurrence of epistemic stretchers is particularly worth commenting on as the findings show that the nineteenth-century writers tended to use them more than those in the previous century. In this respect, log-likelihood calculation, i.e., 159.49 with $p > 0.05$, confirms this overuse of epistemic stretchers in the nineteenth century subcorpus with respect to those in the eighteenth century subcorpus. The log-likelihood ratio, 0.8 in this case, corroborates this tendency to show more cases of these stretchers in the nineteenth century. A possible explanation for the increase of epistemicity in scientific texts is the ongoing professionalisation of science at the time: the large proliferation in scientific societies and journals contributed to a more formal, structured way of publishing and exchanging ideas for members of the growing scientific community. In practice, this meant that, in order to gain acceptance, authors had to find a balance when presenting facts and judgements. The expression of epistemicity is certainly a rhetorical strategy which might have let them make their own claims showing modesty and politeness to be eventually accepted by their colleagues.

Epistemic stretchers are not the only elastic language (sub)category which can be used primarily as a mitigating strategy in CELiST; that is also the case of approximators, e.g., vague quantifiers, and scalar stretchers, e.g., softeners. The level of imprecision conveyed by vague quantifiers and softeners suggests authorial reservations about the truth of the information presented and writers use them to safeguard their image. Intensifying scalar stretchers, on the other hand, emphasise the force of the arguments discussed, standing for a persuasive technique of the authors' position.

We have identified other functions for the categories investigated: sometimes authors simply lack the details or the details themselves are not a requisite under

certain circumstances, which entitles them to give just the right amount of information. Approximators and general terms serve this purpose. Approximators do also serve other purposes such as building writer-reader solidarity on the basis of shared knowledge because they are supposed to belong to the same community.

4.1 Providing the right amount of information

Elastic items can be used to express that precision is not necessary for the purpose of communication or, alternatively, that it is merely not attainable for whatever reason, e.g., the information is not available at the time of writing. In both cases, the writer strategically provides just the right amount of information and communication is still effective. Approximators and general terms generally fulfill this function in the corpus as shown in (1)-(5):

(1) When the roof of the Hot-house is covered, one fire will suffice for **about** 7 or 800 square feet; but there no covering is used, it will not give a proper heat to more than 5 or 600 feet (1786, Speechly)
(2) The narwal, when full grown, measures from thirteen to sixteen feet in length, exclusive of the tusk, and at the thickest part, which is two feet behind the fins, the circumference is **about** eight or nine feet (1828, Goodman)
(3) The epidermis may sometimes be so thick that it will proper to take it off, before the shell can be polished, for that purpose pour a proportion (**nearly** one tenth) of aqua-fortis mixed with common water, into a shallow bason or saucer, and place the shells therein (1794, Donovan)
(4) I could not find any great difference between the Coral found on the West Coast of England, and this in Jamaica, and very little between it and one I had given me by [Mr]. George Handisyd […] this last was a little higher, more branched, less crooked, and slenderer; That I gathered in Jamaica, was not **over** an Inch high with many very crooked Branches, smooth and solid (1707, Sloane)
(5)Whether this be true, I am now uncertain; tho' I had once like to have experienced its effects; when fitting in a chair one evening, and putting my hand behind me, I perceived **something** unusually cold, which I took to be the back of the chair, but soon after felt it move; when starting up, I perceived I had laid my hand on one of these Snakes (1769, Bancroft)

An approximator is one of the two subtypes of approximate stretchers; vague quantifiers stand as the other subtype. Approximators have been referred to as approximations (Channell 1980), approximators (Wachtel 1981, Channell 1994, Bibet et al. 1999) and rounders (Prince, Frader & Bosk 1982) in earlier literature. In our examples, the approximators *about*, *nearly* and *over* indicate numerical imprecise quantities.

As a preceding stretcher (Zhang 2015:81), *about* modifies the following item to provide an estimate of the surface area of a greenhouse which will be heated with one fireplace in (1), and of body girth of narwhals in (2). The pattern coincides with that identified by Channell (1980, 1990, 1994), that is, *approximator* + *exemplar number* + *measure noun*. The elasticity of *about* is reinforced by the presence of the conjunction *or* which coordinates two exemplar numbers in each case and explicitly signals the existence of two possibilities. The approximator together with

the variability attached to exemplar numbers, e.g., *7 or 800 square feet* and *eight or nine feet*, suggest that there is no need to achieve higher precision for the purposes of the current description.

Nearly in (3) and *over* in (4) constitute special cases of approximators, i.e., "partial specifiers" (Wachtel 1980). According to Jucker, Smith and Lüdge (2003:1760), they are different from other prototypical items because "they do not seem to mark a symmetrical interval around the exemplar number as do *around*, *about* or *like*". *Nearly* and *over* express a lower and an upper limit, respectively. Providing a fixed limit rather than the symmetrical interval in (3) appears to be enough for the reader to interpret the elastic approximator as indicating that the ratio of the two liquids in the mixture is not actually relevant, if and only if, it does not exceed one-tenth. Given the circumstances depicted in (4), the writer probably knows the exact size of the coral he is comparing with other specimens as he collected it himself in his voyage to Jamaica, yet he chooses to deliberately withhold the exact information, which contextually implies that precision is not important.

(5) contains an instance of a general term which, according to Zhang's (2015:36) classification, is a subtype of general stretchers together with placeholders and vague category markers. Ruzaite (2007:38) comments on their status as genuine vague language examples whose "interpretation […] depends largely on the hearer's framework of knowledge" as well as on their inherent potential for being "replaced by a more precise item". General terms like *something* in (5) are also referred to as non-numerical specifiers (Drave 2002); they have indefinite reference as they refer "to some entity or entities, but the identity of the referent(s) is either not known or not relevant to the message being conveyed" (Cruse 2006:86-87). In this context, *something* marks the author's lack of knowledge regarding the entity he had put his hand on. If the writer had had that information, he would have certainly provided it.

4.2 Rapport building

Another notable function of elastic language is to facilitate rapport in interaction. The notion of rapport, in turn, brings together certain aspects of communication like audience awareness and in-group solidarity which gain significance in the context of scientific discourse: in principle, writers and readers of scientific texts belong to the same community and, as such, the text itself is the space where meanings can be negotiated, and relationships can be established and maintained. Three (sub)categories of elastic items contribute to building up rapport in CELiST, namely, approximators, vague category markers and scalar stretchers. Consider (6)-(9):

(6) THE whole plant is white. The stem filiform, half an inch in length. The pileus hemispherical, and **about the diameter** of a rape seed (1789, Bolton)
(7) THE OUNCE is a most cruel Beast, bred in Libya, **about as big as** a Mastiff Dog; his Face and Ears being like a Lion's; his Body, Tail, Feet, and Nails, like a Cat's (1730, Boreman)

(8) The House-mice are the same here as those with us, and these and all other kinds of Mice are scarce here (except the Rear-mouse) which may reasonably be supposed from the great quantity of Vermine that continually destroy them, such as Hawks, Owls, Rattle-Snakes, Black-Snakes, **and the like** (1737, Brickell)
(9) It is **hardly** necessary to point out how unfit the human feet are for all purposes of prehension: but the hind limbs of the simiae really deserve the name of hands more than the front; and are more advantageously constructed for holding (1819, Lawrence)

The approximator *about* does not only collocate with *exemplar numbers* + *measure nouns* as commented on earlier in this section (examples (1) and (2)). Alternative patterns, though comparatively less significant, have been also identified, e.g., *about* + *noun phrase* as in (6) or *about* + *comparative structure* as in (7). Here the authors seem to appeal to some degree of shared knowledge with their readership: there must be some common ground which allows readers to safely infer the size of the referents, e.g., *rape seed* and *Mastiff Dog*. The level of shared knowledge is probably discipline-bound, i.e., writers and readers are insiders in the discipline of Life Sciences, so being more precise in the description turns out to be redundant.

Similarly, vague category markers such as *and the like* in (8) and scalar stretchers such as *hardly* in (9) stand as rapport builders. Regarding form, the former tends to have a formulaic structure consisting of *conjunctive or disjunctive conjunction* + *noun phrase*; it is often found at the end of lists. The latter conventionally involves degree adverbs which "describe the extent to which a characteristic holds" (Biber et al. 1999:554). The degree expressed may be greater or less than is commonly assumed, or greater or less than the degree of something else in its vicinity, and their interpretation is always determined by that of other stretchers on the same scale. In (8) *and the like* is preceded by four exemplars, e.g., *Hawks*, *Owls*, *Rattle-Snakes* and *Black-Snake*s, all of which belong to the wider category of *Vernime*. Again, some common knowledge must be shared, otherwise the reader cannot fill the lexical gap rendered by *and the like* with an appropriate exemplar for the overt prototype *Vernime*. *Hardly* in (9) qualifies the need to point out something evident to both writer and reader, e.g., *how unfit the human feet are for all purposes of prehension*. The stretcher evinces this sense of assumed shared understanding, acting as a catalyst for effective writer-reader relationships.

4.3 Mitigation, intensification and self-protection

We have decided to consider mitigation and intensification under the same section because they can be taken as two sides of the same coin: while mitigating items result in softening claims, intensifying ones strengthen claims. They both serve to modulate the weight writers give to their assertions, allowing them to provide differing degrees of commitment to textual information. As regards self-protection, it seems to us that it is unavoidably connected with mitigation because one of the reasons why writers mitigate their claims may be derived, in turn, from their need to shield themselves from potential criticism. Consequently, mitigating elastic items can serve as self-protecting devices too. In this respect, we should like to highlight the multifunctional nature of elastic language which is already

acknowledged in Zhang's (2015:152) framework: "a stretcher may be multifunctional, while at the same time a strategic function may be carried out by more than one stretcher […] functions are [not] discrete and clear-cut". To further complicate matters, polyfunctionality applies to epistemicity as well, as put forward by Hyland (1998:156) with special reference to hedging: "Any analysis of epistemic uses of language must confront the fact that hedging devices are polysemous and polypragmatic". In short, isolating the mitigating and self-protection elastic functions is a complex undertaking in practice, if feasible at all.

The (sub)categories of elastic language which may fulfil the functions above involve vague quantifiers, scalar stretchers and epistemic stretchers. Take the vague quantifiers in (10)-(12):

> (10) First, then, let it be observed, that these marine exuviae are not equally dispersed in every place; in many districts, **few** or none are found, in **many** districts again they are found in all parts (1758, Borlase)
> (11) Is inserted by a short and strong Tendon into the fourth Bone of the first Rank of the Carpus, placed upon the third; at **some** distance from its Termination, there goes a Ligament from this little Bone to the Os metacarpi minimi digiti, which **some** reckon to be a Continuation only of the Tendon of this Muscle (1707, Douglas)

(10) shows some instances of approximate stretchers, in particular, vague quantifiers. Both the assertive multal *many* and the negative *few* are used in making a generalisation about the geographical distribution of marine fossils in Cornwall and other places in the world. The elasticity of *many* and *few* allows the author to emphasise the complexity of the topic he is dealing with, something he announces earlier in the text, e.g., *yet in this particular we shall not find Cornwall singular, nor the subject sufficiently examined*. *Many* occurs in an assertive context, thereby expressing the author's certainty in the large number of districts where marine fossils can be found; *few* occurs in a negative one and so the author withholds complete commitment to the proposition. The contexts in which they are used comes to confirm their function as intensifying and downtoning elements, respectively.

The assertive quantifier *some* combines with an uncountable noun in *some distance* and acts as a pronoun in *some reckon* in (11). In both cases, the unspecified quantity is a moderate one which may well match the traditional lower-bounded implicature "at least *n*" and the upper-bounded implicature "not all", above all in *some reckon* which asserts that there are people (or authors) who consider a certain ligament to be a part of the tendons of the forearm, and simultaneously implies that that is not applicable to everybody (or all authors). The degree of imprecision expressed by the speaker is exploited for the purposes of attribution of information as marked by *reckon*, a verb of cognitive attitude whose interpretation renders "a proper qualificational reading" and whose original semantics is "related to the domain of computation and estimates" (Cappelli 2007:269). Schick (2014) shows the socialising potential of the indefinitely referential expression *some people*, as it

seeks to mitigate the impact of less than positive assessments. *Some* in (11) doubly contributes to elasticity as a mitigating and a self-protection item.

Examples (12) and (13) illustrate the functions fulfilled by scalar stretchers like *somewhat* and *very* as mitigating/self-protection and intensifying devices respectively:

(12) It is **somewhat** remarkable, that so sagacious a philosopher as the illustrious and learned RAY, who so clearly saw the object of migration in fishes, should not also have been led to a sight of it in birds (1824, Jenner)
(13) But let me tell you that five of their eyes are by no means so good as one of yours, and indeed though these red specks are essentially organs of sight, it is **very** doubtful how much they see with them (1859, Agassiz)

Somewhat is aimed at softening the mild criticism being made to *RAY*, a philosopher who could not successfully determine the reasons why birds migrate. It provides modulation to the impact of such an opinion on the scientific community and, at the same time, conveniently reduces the risk of imposition. *Very* combines with the epistemic adjective *doubtful*, strengthening the uncertainty over something being the case, e.g., that starfish are capable of seeing with their eyes. The writer ostensibly pursues assertiveness to persuade readers of the fact that starfish do not see with their eyes: he explicitly addresses his readership earlier in the excerpt, e.g., *let me tell you […] as one of yours*.

Epistemic stretchers are understood as epistemic stance markers in elasticity theory (Zhang 2015:105). Following (Biber et al. 1999:854), they are said to express the writer's assessment regarding the certainty and reliability of the proposition, and even the source of information:

(14) The genus Bruchus is found upon pea-blossoms. Its colour is black, speckled with white: the two fore-legs are reddish, and the hind-legs are armed with **a sort of** tooth, the purpose of which does not seem to be known. (Wakefield, 1816)
(15) Therefore the sublimest Genius must, with Humility, own, that we know but in part; and that by searching we cannot (till God is pleased father to unveil his now secret, tho' always wise Ends) find out the Causes and Effects of several Phaenomena, which future Ages **may possibly** discover (1750, Hughes)
(16) This method **might perhaps** in England be deemed expensive, as the great duty on Spirits has raised their price to an enormous height; but in a country where Rum is sold for ten pence sterling per gallon, the case is far different (1769, Bancroft)
(17) THIS Bird is of the Bigness of the Song Trush, the Figure is of the natural size; it hath a long straight Bill, flat Crown, long Head, and short Neck; the Head, **I think**, not so big in proportion to the Body, as in some others of this Genus (1743, Edwards)

Sort of in (14) indicates epistemic imprecision. Lakoff (1972) and Prince, Frader and Bosk (1982) observe that *sort of* is a hedging device which attenuates the membership of a certain expression to a certain category or prototype. Likewise, Rowland (2007:82) describes *sort of* as a hedge which brings about "the effect of blurring category boundaries or otherwise imprecise measures". Jucker, Smith and Lüdge (2003:1746) explain the hedging function of *sort of* from a relevance-

theoretic perspective as a loose use of language which "might indicate that the degree of interpretative resemblance is not as close as the hearer might otherwise expect". In (14) *sort of* suggests that what the beetle has in the hind-legs merely resembles (and only resembles) a tooth. It hedges the proposition by showing that the writer has some reservations as for the adequacy of the word *tooth* and, as a result, he is not completely committed to its truth.

May possibly in (15) and *might perhaps* in (16) illustrate harmonic combinations of hedges, that is, clusters of hedges, e.g., epistemic modal verbs and epistemic adverbs together. As pointed out by Hyland (1998:116), both *may* and *might* frequently exhibit the features of prototypical hedges. Collins (2007:478) states that *may* is primarily used subjectively, in which case it qualifies the speaker's attitude towards the propositional content expressing lack of knowledge about its truth. *Might*, as a distal form of *may*, is said to express tentativeness to a higher degree (Perkins 1983:59). Biber et al. (1999:491-493) add logical possibility to the possible interpretations of *might*. *Possibly* and *perhaps* are basically epistemic because they signal differing degrees of doubt and certainty. The epistemicity sense in (15) emanates from the prediction made by the writer as for future scientific discoveries; in (16), it comes from a potential consideration about the cost of a method for preserving bird specimens in England. Hedging in these two examples represents an explicit acknowledgement of the existence of other possibilities regarding the actualisation of the events. *May possibly* and *might perhaps* ultimately show authorial self-protection in these two excerpts.

Flanked by commas, the epistemic stretcher *I think* in (17) displays elastic properties. Stubbs (1986) refers to the modal and psychological interpretations this verb may have; Cappelli (2007), in the same vein, touches upon its qualificational and non-qualificational senses. Both authors stress that the different interpretations can only be construed in a contextualised manner. Biber et al. (1999:972) regard *think* as an epistemic stance marker which does not only signal lack of certainty but also source of knowledge. The authors contend that, compared with *know* and *suspect*, for example, *think* conveys less certainty than the former and more certainty than the latter. Zhang (2015:116) precisely concludes that "*think* is elastic in terms of its fluidity in form and function". Parenthetical *I think* in (17) primarily makes perspectivisation explicit since it frames a thought from the author's point of view, e.g., the head and body proportion of the black and white kingfisher with respect to other similar birds. Nevertheless, deviation from other perspectives is contextually implied probably because the writer lacks more than sufficient evidence to present the information with more assurance as indicated by the elastic generalisation *as in some others of this Genus*. *I think* safeguards the author from the stigma of being wrong and potential criticism.

5. Conclusion

This paper has addressed the presence of vague language in a corpus of eighteenth and nineteenth century scientific texts written in English. The search for vagueness has been guided by the theoretical framework of elasticity proposed by Zhang

(2015). Elastic items have been classified into four major lexical categories with their corresponding subtypes: (i) approximate stretchers, i.e., approximators and vague quantifiers, (ii) general stretchers, i.e., general terms, placeholders and vague category markers, (iii) scalar stretchers, i.e., intensifiers and softeners, and (iv) epistemic stretchers. All the major categories have been attested in the corpus.

The most frequent category is approximate stretchers; they are deployed in the expression of both numerical and non-numerical quantification. The pragmatic functions fulfilled by the items in this category are varied, for example, approximators may be used to hedge propositional content whenever it is contextually implied that writers do not need to achieve precision, e.g., it is manifest that both reader and writer share some knowledge. Additionally, approximators enhance writer-reader relationships by marking group membership. The generalisations brought about by vague quantifiers serve primarily to mitigate the impact of the authors' assessments on the academic community.

The occurrence of general stretchers is the least significant of all categories. In the examples analysed, vague category markers are functionally similar to approximators as they serve to build writer-reader relationships by strategically signalling group membership and solidarity. General terms, in contrast, mark a lack of precision probably derived from the author's unavailability of information.

Three different functions have been identified for scalar stretchers: depending on the context, they may (i) provide positive rapport on the basis of shared knowledge; (ii) mitigate authorial claims so as to avoid the risk of imposition; and (iii) express less than full commitment in the absence of supporting evidence. In the case of epistemic stretchers, the pragmatic effects resulting from their use range from hedging to showing perspectivisation.

Future research may explore elastic language across the different subcorpora of the CC in order to determine disciplinary differences in scientific texts on Astronomy, Philosophy, History and Life Sciences. The analysis of elasticity may even cover the extent to which collocational patterns are related to pragmatic functions.

References

Alley, Michael (2018), *The craft of scientific writing*. New York: Springer.

Alonso-Almeida, Francisco (2012), "An analysis of hedging in eighteenth-century English astronomy texts", in Moskowich, Isabel & Begoña Crespo (eds.), *Astronomy "playne and simple". The writing of science between 1700 and 1900*. Amsterdam & Philadelphia: John Benjamins, 199–220.

Álvarez-Gil, Francisco J. (2018a), "Epistemic modals in early Modern English history texts. Analysis of gender variation", *Revista de Lingüística y Lenguas Aplicadas,* 13:13-20.

Álvarez-Gil, Francisco J. (2018b), *Adverbs ending in -ly in late Modern English. Evidence from the Coruña Corpus of History English Texts*. Colección Estudios de Lingüística Aplicada, Núm. 4. Valencia: Editorial Universitat Politècnica de València.

Álvarez-Gil, Francisco J. & Natalia García-Alonso (2019), "Analysis of the linguistic expression *I think* in a corpus of late Modern English scientific texts", *Revista de Lenguas para Fines Específicos*, 25(2):44–57.

Banks, David (1998), "Vague quantification in the scientific journal article", *ASp. La Revue du GERAS*, 19-22:17–27.

Barsaglini-Castro, Anabella & Daniel Valcarce (2020), "The Coruña Corpus Tool: Ten years on", *Procesamiento del Lenguaje Natural*, 64:13–19.

Biber, Douglas (1993), "Representativeness in corpus design", *Literary and Linguistic Computing*, 8(4):243–257.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan (1999), *Longman grammar of spoken and written English*. New York: Longman.

Canale, Michael & Merrill Swain (1980), "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics*, 1(1):1–47.

Cappeli, Gloria (2007), *"I reckon I know how Leonardo da Vinci must have felt…" Epistemicity, evidentiality and English verbs of cognitive attitude*. Italy: Pari Publishing.

Carroll, Ruth (2007), "Lists in letters: NP-lists and general extenders in early English correspondence" in Moskowich-Spiegel, Isabel & Begoña Crespo-García (eds.), *Bells chiming from the past: Cultural and linguistic studies on early English*. Amsterdam & New York: Rodopi, 37–53.

Carroll, Ruth (2009), "Vague language in the medieval recipes of the *Forme of Cury*", in Peikola, Matti, Janne Skaffari & Sanna-Kaisa Tanskanen (eds.), *Instructional writing in English. Studies in honour of Risto Hiltunen*. Amsterdam & Philadelphia: John Benjamins, 55–82.

Channell, Joanna (1980), "More on approximations: A reply to Wachtel", *Journal of Pragmatics*, 4:461–476.

Channell, Joanna (1983), *Vague language: Some vague expressions in English*. University of York: Unpublished PhD thesis.

Channell, Joanna (1990), "Precise and vague quantifiers in writing on economics", in Nash, Walter (ed.), *The writing scholar: Studies in academic discourse (Volume 3)*. London & New Delhi: Sage Publications, 95–117.

Channell, Joanna (1994), *Vague language*. Oxford: Oxford University Press.

Cheng, Winnie & Martin Warren (2001), "The use of vague language in intercultural conversations in Hong Kong", *English World-Wide*, 22(1):81–104.

Cheng, Winnie & Martin Warren (2003), "Indirectness, inexplicitness and vagueness made clearer", *Pragmatics*, 13(3):381–400.

Collins, Peter (2007), "*Can*/*could* and *may*/*might* in British, American and Australian English: A corpus-based account", *World Englishes*, 26(4):474–491.

Cook, Guy (2007), "'This we have done': The vagueness of poetry and public relations" in Cutting, Joan (ed.), *Vague language explored*. London: Palgrave Macmillan, 21–39.

Cotterill, Janet (2007), "'I think he was kind of shouting or something': Uses and abuses of vagueness in the British courtroom", in Cutting, Joan (ed.), *Vague language explored*. London: Palgrave Macmillan, 97–114.

Crespo, Begoña & Isabel Moskowich (2020), "Astronomy, Philosophy, Life Sciences and History texts: Setting the scene for the study of modern scientific writing, *English Studies*, 101(6):665–684.

Cruse, Alan (2006), *A glossary of semantics and pragmatics*. Edinburgh: Edinburgh University Press.

Crystal, David & Derek Davy (1975), *Advanced conversational English*. London: Macmillan.

Cusen, Gabriela (2019), "Forms and functions of vague language in academic journal article abstracts", *Diacronia*, 9:1–8.

Cutting, Joan (2007), *Vague language explored*. London: Palgrave Macmillan.

Drave, Neil (2002), *Vague language in intercultural conversation*. City University of Hong Kong. Unpublished PhD thesis.

Fairclough, Norman (2003), *Analysing discourse. Textual analysis for social research*. London & New York: Routledge.

Gilbert, Nigel B. & Michael Mulkay (2003), *Opening Pandora´s box. A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press.

Hasselgreen, Angela (2005), *Testing the spoken English of young Norwegians: A study of test validity and the role of "smallwords" in contributing to pupils' fluency*. New York: Cambridge University Press.

Hickey, Raymond (2003), *Corpus Presenter. Software for language analysis. With a manual and A Corpus of Irish English as sample data*. Amsterdam & Philadelphia: John Benjamins.

Hinkel, Eli (1997), "Indirectness in L1 and L2 academic writing", *Journal of Pragmatics*, 27:361–386.

Hinkel, Eli (2002), *Second language writers' text. Linguistic and rhetorical features*. Mahwah, New Jersey & London: Lawrence Erlbaum Associates.

Hinkel, Eli (2005), "Hedging, inflating, and persuading in L2 academic writing", *Applied Language Learning*, 15:29–54.

Hyland, Ken (1996), "Talking to the academy: Forms of hedging in science research articles", *Written Communication*, 13(2):251–281.

Hyland, Ken (1998), *Hedging in scientific research articles*. Amsterdam: John Benjamins.

Hyland, Ken (2000), *Disciplinary discourses: Social interactions in academic writing*. New York: Longman.

Hyland, Ken & John Milton (1997), "Qualification and certainty in L1 and L2 students' writing", *Journal of Second Language Learning*, 6(2):183–205

Jucker, Andreas, Sara W. Smith & Tanja Lüdge (2003), "Interactive aspects of vagueness in conversation", *Journal of Pragmatics*, 35:1737–1769.

Kempson, Ruth (1977), *Semantic theory*. Cambridge: Cambridge University Press.

Koetser, Almut (2007), "'About twelve thousand or so': Vagueness in North American and UK offices" in Cutting, Joan (ed.), *Vague language explored*. London: Palgrave Macmillan, 40–61.

Kyburg, Alice & Michael Morreau (2000), "Fitting words: Vague language in context", *Linguistics and Philosophy*, 23:577–597.

Lakoff, George (1972), "Hedges: A study in meaning criteria and the logic of fuzzy concepts", *Chicago Linguistic Society Papers*, 8:138–228.

McCarthy, Michael (2020), "Vague language in business and academic contexts", *Language Teaching*, 53(2):203–214.

Monaco, Leida Maria & Luis Puente-Castelo (2019), "'A matter of both curiosity and usefulness': Compiling the *Corpus of English Texts on Language*", *Research in Corpus Linguistics*, 7:47–68.

Moskowich, Isabel & Begoña Crespo (2019), "'Arguments that could possibly be urged': Modal verbs and tentativeness in the Coruña Corpus", *Languages*, 4:57-69.

Murphy, Bróna (2010), *Corpus and sociolinguistics: Investigating age and gender in female talk*. Amsterdam & Philadelphia: John Benjamins.

Myers, Greg (1989), "The pragmatics of politeness in scientific articles", *Applied Linguistics*, 10:1–35.

Myers, Greg (1996), "Strategic vagueness in academic writing" in Ventola, Eija & Anna Mauranen (eds.), *Academic writing: Intercultural and textual issues*. Amsterdam & Philadelphia: John Benjamins, 3–17.

Neary-Sundquist, Colleen (2013), "The use of hedges in the speech of ESL learners", *Estudios de Lingüística Inglesa Aplicada*, 13:149–174.

O'Keeffe, Anne, Michael McCarthy & Ronald Carter (2007), *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.

Ortega-Barrera, Ivalla (2012), "A diachronic discussion of extenders in English remedies. Found in the *Corpus of Early English Recipes* (1350–1850)", in Sauer, Hans & Gaby Waxenberger (eds.), *English historical linguistics 2008. Selected papers from the fifteenth International Conference on English Historical Linguistics (ICHEL 15), Munich, 24-30 August 2008*. Amsterdam & Philadelphia: John Benjamins, 223–235.

Overstreet, Maryann (1999), *Whales, candlelight, and "stuff like that". General extenders in English discourse*. New York: Oxford University Press.

Parapar, Javier & Isabel Moskowich-Spiegel (2007), "The Coruña Corpus Tool", *Procesamiento del Lenguaje Natural*, 39:289–290.

Perkins, Michael R. (1983), *Modal expressions in English*. London: Frances Pinter Publishers.

Prince, Ellen, Joel Frader & Charles Bosk (1982), "On hedging in physician-physician discourse", in Di Pietro, Robert J. (ed.), *Linguistics and the profession*. Norwood, New Jersey: Ablex, 83–97.

Puente-Castelo, Luis & Leida María Mónaco, (2016), "The might of 'might': A mitigating strategy in eighteenth and nineteenth century female scientific discourse", *Revista Canaria de Estudios Ingleses*, 72:143–168.

Quintana-Toledo, Elena (2017), "Vagueness in the Corpus of Historical English Texts" in Alonso-Almeida, Francisco (ed.), *Stancetaking in late Modern English scientific writing. Evidence from the Coruña Corpus. Essays in honour of Santiago González y Fernández-Corugedo*. Valencia: Editorial Universitat Politècnica de València, 131–150.

Rowland, Tim (2007), "'Well maybe not exactly, but it's around fifty basically?': Vague language in Mathematics classroom", in Cutting, Joan (ed.), *Vague language explored*. London: Palgrave Macmillan, 79–96.

Ruzaite, Jurate (2004), "Academic precision reconsidered: A corpus-based account", *SKY Journal of Linguistics*, 17:217–247.

Ruzaite, Jurate (2007), *Vague language in educational settings: Quantifiers and approximators in British and American English*. Frankfurt am Main: Peter Lang.

Ruzaite, Jurate (2009), "Corpora for applied purposes: A case study of quantifiers in English and Lithuanian", in *Eesti Rakenduslingvuistika Ühingu Aastaraamat*, 5:239–250.

Sainsbury, Mark (1996), "Concepts without boundaries" in Keefe, Rosanna & Peter Smith (eds.), *Vagueness: A reader*. Cambridge: The MIT Press, 186–205.

Salager-Meyer, Françoise (1994), "Hedges and textual communicative function in medical English written discourse", *English for Specific Purposes*, 13(2):149–170.

Salager-Meyer, Françoise (1997), "*I think that perhaps you should*: A study of hedges in written scientific discourse" in Miller, Tom (ed.), *Functional approaches to written text: Classroom applications*. Washington: United States Information Agency, 105–118

Salager-Meyer, Françoise & Gérard Defives (1998), "From the gentleman's courtesy to the scientist's caution: A diachronic study of hedges in academic writing (1810–1995)", in Fortanet, Inmaculada, Santiago Posteguillo, Juan Carlos Palmer & Juan Francisco Coll (eds.), *Genre studies in English for academic purposes (Vol. 9)*. Castelló de la Plana: Publicacions de la Universitat Jaume I, 133–172.

Salager-Meyer, Françoise, Gérard Defives & Miguel Hamelinsck (1996), "Epistemic modality in 19th and 20th century medical English written discourse: A principal component analysis study", *Interface: A Journal of Applied Linguistics*, 10:163–199.

Schick, Laurie (2014), "Some people: From referential vagueness to social-moral socialization in middle school dance classes", *Pragmatics and Society*, 5(2):243–270.

Skelton, John (1997), "How to tell the truth in *The British Medical Journal*: Patterns of judgement in the 19th and 20th centuries", in Markanen, Raija & Hartmut Schröder (eds.), *Hedging and discourse. Approaches to the analysis of a*

pragmatic phenomenon in academic texts. Berlin & New York: Walter de Gruyter, 42–63.
Sperber, Dan and Deirdre Wilson (1991), "Loose talk", in Davis, Steven (ed.), *Pragmatics: A reader*. Oxford: Oxford University Press, 540–549.
Stubbs, Michael (1996), *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.
Towns, Peter (1990), "Is literacy important?", in *Australian Biochemical Society Newsletter*, 3:21.
Trappes-Lomax, Hugh (2007), "Vague language as a means of self-protective avoidance: Tension management in conference talks", in Cutting, Joan (ed.), *Vague language explored*. London: Palgrave Macmillan, 117–137.
Urbanová, Ludmila (1999), "On vagueness in authentic English conversation", *Brno Studies in English*, 25:99–107.
Van Rooij, Robert (2011), "Vagueness and linguistics", in Ronzitti, Giussepina (ed.), *Vagueness: A guide*. Dordrecht: Springer, 123–170.
Wachtel, Tom (1980), "Pragmatic approximations", *Journal of Pragmatics*, 4:201–211.
Wachtel, Tom (1981), "Distinguishing between approximations", *Journal of Pragmatics,* 4(3):201–211.
Wallwork, Adrian (2016), *English for writing research papers*. Cham: Springer.
Zhang, Grace (2013), "The impact of touchy topics on vague language use", *Journal of Asian Pacific Communication*, 23:87–118.
Zhang, Grace (2015), *Elastic language: How and why we stretch our words*. Cambridge: Cambridge University Press.

Appendix 1. Texts in CELiST.

YEAR	AUTHOR	TITLE
1707	Douglas, James	*Myographiae comparatae specimen: or, a comparative description of all the muscles in a man and in a quadruped. ... To which is added an account of the muscles peculiar to a woman, etc. M.D.*
1707	Sloane, Hans	*A Voyage to the islands Madera, Barbadoes, Nieves St Christophers and Jamaica; with the Natural History of the Herbs and trees, four footed Beasts, Fishes, Birds, Insects, Reptiles, &c. of the last of those Islands. To which is prefix'd an introduction, wherein is an account of the inhabitants, air, waters, diseases, trade, &c. of that place, with some relations concerning the neighbouring continent and islands of America.* In Two Volumes. Vol. I.
1717	Keill, James	*Essays on several parts of animal oeconomy. Essay IV: Of Animal Secretion.*
1720	Gibson, William	*The Farriers new Guide: containing first, the anatomy of a horse, being an exact and compendious description of all his parts; with their actions and uses: illustrated with figures curiously engrav'd on copper plates. Secondly, an account of all the diseases incident to horses, with their signs, causes, and methods of cure; wherein many defects in the farriers practice, are now carefully supply'd, their errors expos'd and amended, and the art greatly improv'd and advanc'd, according to the latest discoveries. The whole interspers'd with many curious and useful observations concerning feeding and exercise, &c.*
1723	Blair, Patrick	*Pharmaco-botanologia: or, an alphabetical and classical dissertation on all the British indigenous and garden plants of the new London Dispensatory. In which their genera, species, characteristik and distinctive notes are methodologically described; the botanical terms of art explained; their virtues, uses, and shop-preparations declared. With many curious and useful remarks from proper observation.*
1730	Boreman, Thomas (bookseller)	*A description of three hundred animals; viz. beasts, birds, fishes, serpents, and insects. With a particular account of the whale-fishery. Extracted out of the best authors, and adapted to the use of all capacities; especially to allure children to read.*
1737	Blackwell, Elizabeth	*A Curious Herbal, containing five hundred cuts, of the moſt uſeful plants, which are now uſed in the practice of*

		Physick. Engraved on folio copper plates after drawings, taken from the LIFE. To which is added a short description of ye plants and their common uses in PHYSICK. In Two Volumes. Vol. I.
1737	Brickell, John	*The Natural History of North-Carolina. With an account of the trade, manners, and customs of Christian and Indian inhabitants. Illustrated with copper-plates, whereon are curiously engraved the map of the country, several strange beasts, birds, fishes, snakes, insects, trees, and plants, &c.*
1743	Edwards, George	*A NATURAL HISTORY OF Uncommon BIRDS. And of some other rare and undescribed animals, quadupedes, fishes, reptiles, insects, &c. Exhibited in two hundred and ten copper-plates, from designs copied immediately from Nature, and curiously coloured after life. With a full and accurate description of each figure.* In Four Parts. Part I.
1750	Hughes, Griffith	*The Natural History of Barbados.* In Ten Books.
1752	Dodd, James Solas	*An Essay towards a Natural History of the Herring.*
1758	Borlase, William	*The Natural History of Cornwall. The Air, Climate, Waters, Rivers, Lakes, Sea and Tides; Of the Stones, Semimetals, Metals, TIN, and the Manner of Mining; The Constitution of the Stannaries; Iron, Copper, Silver, lead, and Gold, found in Cornwall. Vegetables, Rare Birds, Fishes, Shells, Reptiles, and Quadrupeds: Of the Inhabitants, Their Manners, Customs, Plays or Interludes, Exercises, and Festivals; the Cornish Language, Trade, Tenures, and Arts.*
1766	Pennant, Thomas	*The British Zoology. Class I. Quadrupeds. II. Birds.*
1769	Bancroft, Edward	*An essay on the Natural History of Guiana, in South America. Containing a description of many curious productions in the animal and vegetable systems of that country. Together with an account of the religion, manners and customs of several tribes of its Indian inhabitants. Interspersed with a variety of literary and medical observations. In several letters from a Gentleman of the Medical Faculty during his residence in that country.*
1774	Goldsmith, Oliver	*An History of the Earth, and animated Nature:* In Eight Volumes. Vol VIII.

1776	Withering, William	*A botanical arrangement of all the vegetables, naturally growing in Great Britain. With the descriptions of the Genera and species, according to the system of the celebrated Linnaeus. Being an attempt to render them familiar to those who are unacquanted with the learned languages. Under each species are added, the most remarkable varieties, the natural places of growth, the duration, the time of flowering, the peculiarities of structure, the common English names; the names of Gerard, Parkinson, Ray and Baubine. The uses as medicines, or as poisons; as food for men, for brutes, and for insects. With their applications in oeconomy and din arts, with an easy introduction to the study of botany. Shewing the method of investigating plants, and directions how to dry and preserve specimens.* In Two Volumes. Vol. I.
1786	Speechly, William	*A Treatise on the Culture of the Pine Apple and the Management of the Hot-house. Together with a Description of every Species of Insect that infest Hot-houses, with effectual Methods of destroying them by William Speechly. To which is added A method to preserve peach and nectarine trees from mildew &c. by Robert Browne.* With plates. Book I.
1789	Bolton, James	*An History of Fungusses, growing about Halifax. With forty-six copper-plates; or which are engraved sixty-four species of funguses, Including the Seven following GENERA, viz. CLATHRUS, HALVELLA, PEZIZA, CLAVARIA, LYCOPERDON, SPHAERIA, and MUCOR. Wherein their various appearances in the different stages of growth, are faithfully exhibited in about three hundred figures, copied with great care from the PLANTS, when newly gathered and in a state of perfection. With a particular DESCRIPTION of each SPECIES, in all its stages. From the first appearance to the utter decay of the plant; with the time when they were gathered; the soil and situation in which they grew; their duration; and the particular places mentioned, where all the new or rare species were found. The whole being a plain recital of FACTS, the result of more than twenty years observation.* In Three Volumes. Vol. III.
1794	Donovan, Edward	*Instructions for collecting and preserving various subjects of natural history: as animals, birds, reptiles, shells, corals plants, &c.: Together with a treatise on the*

		management of insects in their several states: selected from the best authorities.
1795	Smith, Sir James Edward	*English Botany; or coloured Figures of British Plants with their essential Characters, Synonyms, and Places of Growth.* In Thirty Six Volumes. Vol. IV.
1804	Jacson, Maria Elizabeth	*Botanical Lectures. By a Lady.* Altered from "Botanical Dialoques for the use of schools", and adapted to the use of persons of all ages, by the same author.
1808	Wilson, Alexander	*American Ornithology; or, The Natural History of the Birds of the United States. Illustrated with plates and colored from the original drawings taken from Nature.* In Nine Volumes. Vol. I.
1816	Wakefield, Priscilla	*An introduction to the Natural History and Classification of Insects, in a series of familiar Letters.* With Illustrative Engravings.
1819	Lawrence, William	*Lectures on Physiology, Zoology, and the Natural History of Man, delivered at the Royal College of Surgeons.*
1824	Jenner, Edward	"Some Observations on the Migration of Birds". By the late Edward Jenner, M.D. F.R.S.; with an Introductory Letter to Sir Humphry Davy, Bart. Pres. R.S. By the Rev. G. C. Jenner.
1828	Godman, John D.	American Natural History. In Three Volumes. Vol. III.
1832	Lincoln, Almira Hart Phelps	*Familiar Lectures on Botany, including practical and elementary Botany, with generic and specific Descriptions of the most common native and foreign Plants and a Vocabulary of botanical Terms. For the use of Higher schools and academies.*
1835	Jardine, Sir William	*The Naturalist's Library. In Forty Volumes. Mammalia Vol. III. Ruminantia. Part I. The Natural History of the Ruminating Animals containing Deer, Antelopes, Camels, &c. Illustrated by thirty-five Plates; with memoir and portrait of Camper.*
1840	Pratt, Anne	*Flowers and their Associations*
1848	Dalyell, Sir John Graham	*Rare and remarkable Animals of Scotland, represented from living Subjects: with practical Observations on their Nature.* In Two Volumes. Vol. II.
1859	Agassiz, Elizabeth	*A First Lesson in Natural History.*
1859	Darwin, Charles Robert	*On the Origin of Species by means of Natural Selection, or the Preservation of favoured Races in the Struggle for Life.*

1863	Huxley, Thomas Henry	*On the Origin of Species, or, The Causes of the Phenomena of organic Nature: a Course of six Lectures to working Men.*
1867	Spencer, Herbert	*The Principles of Biology.* In Two Volumes. V. II.
1876	Macalister, Alexander	*An introduction to animal morphology and systematic zoology.*
1879	Lankester, Phebe	*Wild Flowers worth Notice: A Selection of some of our Native Plants which are most attractive from their Beauty, uses, or Associations.*
1880	Balfour, Francis M.	*A Treatise on comparative Embryology.* In Two Volumes. Vol. I.
1889	Galton, Francis	*Natural Inheritance.*
1895	Gregory, Emily Lovira	*Elements of Plant Anatomy.*
1898	Packard, Alpheus Spring	*A Text-book of Entomology including the Anatomy, Physiology, Embryology and Metamorphoses of Insects. For use in agricultural and technical Schools and Colleges as well as by the working Entomologist.*

Appendix 2. List of elastic language items investigated (Zhang 2015).

Approximate stretchers	*a few* *a little* *a lot of* *about* *almost* *approximately* *around* *majority* *many* *most* *much* *nearly* *or so* *over* *some*
General stretchers	*and the like* *and things like that* *anybody* *anyone* *anything*

	or something *somebody* *someone* *something* *something like* *stuff* *thing*
Scalar stretchers	*a bit* *hardly* *kind of* *quite* *really* *so* *somewhat* *sort of* *too* *very*
Epistemic stretchers	*believe* *certain* *certainly* *could* *guess* *likely* *may* *maybe* *might* *perhaps* *possible* *possibly* *probable* *probably* *sure* *surely* *think*