

A corpus-assisted study of vague language in corporate responsibility reports of the cosmetics industry

Bixi Jin

Beihang University (China)

bixijin@outlook.com

Abstract

A corporate social responsibility (CSR) report is a form of information disclosure about environmental, health, and human rights issues. It serves as an essential bridge for communication between enterprises and stakeholders. Enterprises can use vague language (VL) to strategically communicate with the public, establish a positive corporate image, and legitimise their activities. In recent years, many scholars have studied the usage and communicative function of VL in general conversational discourse, yet VL in business discourse has not received much attention.

In this study, 30 social responsibility reports of 10 world-renowned cosmetic enterprises between 2015–2019 were collected to build a corpus. With the help of the corpus software AntConc, this paper explores VL, specifically in CSR reports, using a corpus-assisted approach. Further, the quantitative findings are complemented by the contextualised interpretation of the communicative functions of VL in CSR reports. The major findings of this research are as follows. There are four types of VL used in CSR texts: quantity, degree, time, and softening stance-taking. In addition, the use of VL can help achieve several communicative purposes of CSR report drafting (i.e., provide an appropriate amount of information, enhance persuasion, and self-protection). Some practical suggestions are discussed for improving the awareness of the strategic use of vagueness in CSR reports.

Keywords: vague language, corporate social responsibility report, corpus-assisted discourse analysis, cosmetics industry, business communication.

Resumen

Estudio asistido por corpus del lenguaje vago en los informes de responsabilidad social corporativa de la industria cosmética

El informe de responsabilidad social corporativa (RSC) constituye una forma de divulgar información sobre asuntos relacionados con el medio ambiente, la salud y los derechos humanos. Funciona como un puente esencial que posibilita la comunicación entre las empresas y los accionistas. Las empresas pueden emplear un lenguaje vago (LV) de manera estratégica para comunicarse con sus destinatarios, establecer una imagen corporativa positiva y legitimar sus actividades. En los últimos años, muchos investigadores han estudiado el uso y la función comunicativa del LV en el discurso conversacional general. Sin embargo, el LV apenas ha sido analizado en el discurso empresarial.

Para el presente trabajo se ha compilado un corpus de 30 informes de responsabilidad social de 10 empresas mundialmente conocidas de la industria cosmética publicados durante los años 2015-2019. Con la ayuda del gestor de corpus AntConc se ha explorado el LV en los RSC desde una perspectiva asistida por corpus. Los resultados cuantitativos se han complementado con una interpretación contextualizada de las funciones comunicativas desempeñadas por el LV en los informes de RSC. Entre los hallazgos obtenidos, cabe destacar que se han identificado cuatro tipos de LV en los textos analizados: de cantidad, de grado, de tiempo y de toma de postura de carácter atenuante. Además, este trabajo ha evidenciado que el uso del LV puede ayudar a alcanzar diferentes propósitos comunicativos en los informes de RSC: proporcionar una cantidad adecuada de información, mejorar la persuasión y autoprotegerse. También se ofrecen algunas sugerencias prácticas para que se tenga una mayor conciencia del papel estratégico de la vaguedad en los informes de RSC.

Palabras clave: lenguaje vago, informe de responsabilidad social corporativa, análisis de discurso asistido por corpus, industria cosmética, comunicación empresarial.

1. Introduction

Extensive research has been conducted to provide various definitions of corporate social responsibility (CSR) and related notions (e.g., sustainability reports), and the evolving conceptualizations of CSR over the decade reveal its inherent intricacies (Maignan, Ferrell, & Ferrell, 2005). A CSR report is officially defined as “the responsibility of enterprises for their impact on society” and “a process to integrate social, environmental,

ethical, human rights and consumer concerns into their business operations and core strategy in close collaboration with their stakeholders” (European Commission, 2011: 6). Though it is becoming a generalised practice to include a social responsibility section in annual reports, as the L’Oreal report (2018) indicates, this non-financial reporting section achieves the communicative purpose of presenting a clear picture of a company’s efforts and commitments in environmental sustainability, ethics, and corporate governance issues, seeking to obtain transparency and legitimacy (Ditlevsen, Nielsen & Thomsen, 2013). Further, it is a management strategy to construct a favorable and trustworthy corporate image (van Riel & Fombrun, 2007; Kushal, 2011; Fuoli & Hart, 2018), targeting a group of heterogeneous audiences, including “existing shareholders, potential shareholders, creditors, investment and credit analysts, the government and their agencies, employees and society at large” (Ditlevsen, 2012: 382). While CSRs tend to prioritise social responsibility at large, informed by stakeholder theory, the incorporation of stakeholder views and interests into CSRs has gained some prominence in view of the unforeseen outcomes emerging from business activities (Fordham & Robinson, 2018; Maignan et al., 2005).

As suggested by the well-known consulting company Ecovia Intelligence, consumers in the cosmetics industry have become extremely sensitive to the range of social and environmental initiatives mentioned in CSR reports. Growing consumer awareness prompts companies to be particularly concerned to explain to the public such facts as basic ingredients in the manufacturing process, animal testing, the certification of natural and organic cosmetics, sustainable packaging, and even financial assistance to disadvantaged communities. As a result, a company may organise CSR reports with a range of linguistic means to demonstrate its social and environmental initiatives. Use of vague language (VL), recognised by its communicative importance, is a drafting strategy in organizing specialised discourse such as laws, policy documents, and business communication (Handford, 2010; Li, 2017; McCarthy, 2019; Scott di Carlo, 2015). The strategic use of VL may also contribute to the delicate discursive negotiation in CSR reporting. The possible application of VL in CSRs motivates the current research.

Business contexts are appropriate places to observe the use of VL. On the one hand, it may be tempting to think that companies are expected to employ VL to emphasise or amplify their positive efforts in maintaining

positive corporate images and legitimizing corporate actions; on the other hand, VL may ensure the desired level of flexibility for companies to maintain control over private information or account for unexpected future contingencies of corporate activities (Choi & Triantis, 2010).

Therefore, this study aims to identify and explore the strategic use of VL in CSR reports in the cosmetics industry. The following questions are addressed: (1) What types of VL are prototypical of CSR reports in the cosmetics industry? (2) Is the use of VL in CSR reports related to their communicative purposes? If so, how?

The study seeks to make several contributions to the current literature on corporate discourse. First, the findings enrich the current understanding of VL in corporate discourse. The quantitative corpus linguistics methods and the detailed qualitative discourse analysis of VL in the beauty industry CSR reports elucidate the use of VL in the construction of CSR reports. Second, from a methodological perspective, the study explores VL in CSR reports based on a relatively large-scale specialised corpus, enabling the particular language expressions in rich contexts to be identified and categorised. Third, the study provides a detailed account of VL in CSR reports and examines how the reports may strategically communicate with heterogeneous audiences, construct a positive corporate image, and legitimise CSR activities. Given that there is no readily available instructional space for VL, the results address the need for incorporating VL into future business training programs and business English learning materials.

2. Definition of vague language

Considerable discussion in recent years has focused on the definitions of VL. While a philosophical approach has suggested that language is inherently vague (Russell, 1997), this is not practical for rhetorical-linguistic analysis. Research addressing the definition of VL has generated considerable interest among applied linguists; however, the internal configuration of this term remains unclear, and even the terms used to refer to VL are themselves vague (Cotterill, 2007). VL has been interpreted in several ways, and presented here are some often-cited definitions. For instance, Channell (1994: 20) argued that the evaluation of VL can be determined by the replacement of other words or expressions if they

appear to “render the same proposition” and is “purposely and unabashedly vague”. Pinkal (1995:12-15) first differentiates between “communicative indeterminacy” and “semantic indeterminacy”. Then, he establishes that there are two types of communicative indeterminacy: vagueness and ambiguity.¹ A similar criterion of determining VL was proposed, focusing on the possibility of making the utterance precise and the indeterminacy arising from linguistic expressions. Ruzaitė (2007: 28), focusing on VL forms and functions, argues that “vague language is a natural, usually purposeful and multi-functional linguistic phenomenon that involves imprecision and is employed for certain communicative strategies.” Zhang (2018: 7) continues this notion by acknowledging that VL can be negotiated to adapt to specific communicative purposes and further specifies that VL is unspecified and context-sensitive.

Based on the abovementioned studies, it is not surprising that the unclear, indefinite configuration of VL makes it difficult to identify VL in actual practice, as many researchers may propose their own interpretations. Therefore, given the intricate nature of identifying VL, it is not my aim to unravel the nuances and minutiae of the various definitions but to be aware of the vague expressions identified by previous studies. The working definition of VL in this study emphasises its unspecified, elastic, negotiable, and communication-sensitive nature (Channell, 1994; Ruzaitė, 2007; Zhang, 2018), which will guide the subsequent data analysis and interpretation of CSR reports in corporate discourse.

3. Vague language used in discourse

VL is a pervasive phenomenon in everyday language. Increasing attention has been paid to the usages and communicative functions of VL in general conversational contexts (Evison, McCarthy & O’Keefe, 2007; Cheng & O’Keefe, 2015) and more specialised discourse domains, such as political discourse (Gruber, 1993), courtroom discourse (Cotterill, 2007), and legal discourse (Anesa, 2007; Li, 2017). The purposeful use of VL in these contexts has been well investigated and summarised. For instance, Anesa (2007) adopted the term “weasel words” to characterise words such as good, appropriate, and reasonable in contracts, as informed by Mellinkoff (1963: 21). Li (2017: 105) created a VL list based on prior studies and categorised the VL used in large-scale legal documents into time, category, quantity, and

degree, and proposed four pragmatic functions: extending applicability, providing flexibility, maintaining a balance between precision and over-elaboration, and mitigating potential problems. Zhang (2018: 123) identified the functions of *some* under four pragmatic dimensions in natural classroom settings: inclusion of an appropriate amount of information, mitigation, withholding information, and discourse management.

Despite its importance in business contexts, limited research regarding VL has been conducted in this field. Liu (2015) discussed the uses of VL in written business communication across cultures, such as politeness, face-protection, and goal achievement. Malyuga and McCarthy (2018) observed that the pragmatic functions of vague category markers (e.g., *and so on*, *or whatever*, *and that*) in spoken business communication was manifested by its knowledge-sharing practices among different business parties to form group solidarity (also see Handford, 2010), or it was used as a face-protection strategy to forestall unexpected circumstances. While VL usages and communicative functions have been well-investigated in general conversational and specialised domains, there has been scarce discussion about VL in the business discourse, particularly regarding the VL usages and functions in CSR discourse.

4. Previous discourse studies of CSR reporting

There is a growing literature concerning CSR discourse. The CSR report is generally recognised as a hybrid genre combining informative and promotional elements (Bondi, 2016). Existing literature, on the one hand, has focused on the description of generic structures of CSR reports based on Swales's (1990) and Bhatia's (2004) frameworks (e.g., Skulstad, 2008; Bhatia, 2012; Yu & Bondi, 2017). On the other hand, some studies have been committed to examining the lexico-grammatical resources utilised in the CSR report. Findings suggest that the linguistic resources companies mobilise for promoting socially responsible actions are used to construct a positive corporate image. Malavasi (2011) observed the frequency wordlists of CSR texts from two companies and analysed the predominant language patterns revealing pervasive corporate values. Bondi (2016) investigated the keywords and phraseologies related to prediction or commitment statements, which are used to legitimise corporate activities and enhance impression management. Stance resources related to expressions of attitudes

and evaluations have been much discussed in this area of inquiry. Fuoli (2012) addressed the appraisal resources informed by systemic functional linguistics (SFL) in the CSR reports by IKEA and BP. He concluded that they employed interpersonal resources to portray a trustworthy corporate image. Later, Fuoli (2017) investigated the stance expressions proposed by Biber et al. (1999) in European- and US-based CSR reports, focusing on their efforts to establish committed, caring, and honest corporate images. Informed by legitimation theory, Lin (2021) examined the discursive strategies adopted by UK and China CSR reports to legitimise negative news. Socio-cultural factors accounting for linguistic variation were discussed, which contribute to the improvement of crisis management strategies used by large companies.

5. Methodology

5.1. Corpus-assisted discourse analysis

A corpus-assisted discourse analysis approach was adopted for this study. The CADS approach is characterised by its synergy of corpus-aided analysis and detailed qualitative discourse analysis, i.e., the automatic detection of linguistic patterns (e.g., frequency, keywords, and collocations) from the naturally occurring texts aided by corpus-analytic tools, enhanced by elaborated interpretations of quantitative patterns in relation to their wider social contexts (Partington, Duguid & Taylor, 2013). The procedure starts from a quantitative analysis aided by corpus techniques such as word frequency and keywords to identify an overall linguistic profile. As a research project unfolds, the analysis become more qualitative and context-sensitive involving substantial efforts of human interpretation and explanation, with less reliance on corpus tools (Baker & McEnery, 2015).

5.2. Corpus construction

In this study, a self-compiled corpus of annual CSR reports was built and employed, consisting of annual CSR reports selected from 10 well-established international beauty groups (L'ORÉAL, Estee Lauder, PROCTER & GAMBLE, KOSÉ, Johnson & Johnson, Amorepacific Group, LVHM, Henkel, COLGATE, and Beiersdorf). It contained 30 texts (three texts selected from each corporation at random) produced from 2015 to 2019, which amounted to 775,237 words. As the study adopted a

CADS approach, the corpus under construction was not committed to compiling large corpora of various genre types, seeking to “obtain an overview of the greatest quantity and variety of discourse types possible” (Partington et al., 2013: 11-12). Specialised corpora of this kind are designed to explore “specific types of genres within specific contexts” (Connor & Upton, 2004: 7-8) that could not be achieved with corpora built for general purposes. Working on a manageably sized corpus also allows researchers to familiarise themselves with the context of the texts and thus enhances the interpretation of corpus findings beyond simply reading concordance lines, countering the criticism that corpus-based research is restricted to somewhat bottom-up lexico-grammatical analysis (L. Flowerdew, 2005).

Corpus	Number of texts	Number of tokens	Number of types	Average words of CSR
CSR	30	775237 essay	20031	25841

Table 1. Description of the CSR corpus.

5.3. Analytical procedure

Antconc 3.5.8 (Anthony, 2019) was employed in processing the corpus data. It is a corpus analysis toolkit for text analysis, including major functions such as wordlists, keywords, and concordances. The software can demonstrate the concordance lines of a searched word or multi-word sequences with the contexts provided. The analysis was conducted in three steps. In the first step, ten texts were randomly selected and read from the total of 30 texts in the corpus. The expressions related to VL were marked, considering the VL items that have been identified by prior studies (e.g., Channell, 1994; Anesa, 2007; Cutting, 2007; Cheng & O’Keeffe, 2015; Li, 2017; Malyuga & McCarthy, 2018). The point of conducting a close reading is that some vague expressions cannot be automatically identified by the text analysis software. Through a detailed reading of the sampled CSR texts, the researcher must generate an initial list of the vague items in the CSR reports. The VL generated in the list generally fell into the categories of approximate quantities, boosters, downtoners, and shields (see Zhang, 2011: 574).

In the second step, I used AntConc to generate the top 50 words from the wordlist and the keyword list (see Appendix) of the self-built CSR cosmetics

corpus to examine whether words or related phrases associated with vagueness in the corpus could be added to the identified VL list. This step aimed to supplement the vague patterns in step with the vague expressions that may not have appeared in the 10 sampled texts. After checking, expressions such as ‘*more*’, ‘*or**’ and ‘*all*’ were added to the list obtained from the Step 1.²

In the third step, AntConc was used to analyse all the VL items identified in Steps 1 and 2 by searching them throughout the corpus. The analyses address several issues: the frequency and semantic categories of vague expressions, the contexts where VL occurs, and the communicative functions for using them in CSR texts. Meanwhile, a further investigation with a research assistant specialised in applied linguistics was conducted to exclude lexical items that were obviously not intended for vague use. Examples were provided to illustrate how VL is strategically used in CSR reports and to analyse the communicative purposes that the corporations intend to achieve.

The quantitative corpus analysis of the specialised corpus categorises the VL in CSR reports into four main types: *quantity*, *degree*, *time* and *softening stance-taking*. Overall, these four categories show the strategic role performed by VL expressions in promoting a trustworthy corporate image. Companies employ these linguistic resources as discursive tactics to “address shareholders’ information needs, but also project readiness and competence; they show that they are in control of the situation and prepared for any contingency” (Fuoli, 2017: 21). In the following sections, VL representing each category is discussed in detail.

6. Results and discussion

6.1. Vague items associated with “quantity”

Compared with other conversational discourse, language in corporate discourse is supposed to be more precise (McCarthy, 2019). It is somewhat surprising to see that the vague items associated with quantity, such as *more than*, *many*, *various*, and *some* used in CSR texts, are at the forefront of frequency level (5.2 per thousand).³ The recurrence of vague quantity avoids revealing a specific numerical value to define the exact quantity level. A possible explanation for this phenomenon is that companies are trying to balance against accuracy and over-explicitness regarding reported corporate

activities in case of alternative interpretations. Examples 1–5 in the following were identified for this communicative purpose.

Example 1. *Additionally, in the past three years alone, this group has helped more than 400,000 American families receive a new recycling cart, totaling more than \$27 million worth of new infrastructure and has forged several state-wide partnerships to drive best practice adoption. (PROCTER & GAMBLE, 2017)*

Example 2. *While the world is making tremendous progress to curb the course of the global HIV epidemic, many challenges remain in the areas of prevention, treatment, access and public education. (COLGATE, 2015)*

Example 3. *This has allowed us to build a portfolio of successful brands in various global markets and to become the leading cosmetic company representing Asian beauty and its wisdom. (Amorepacific Group, 2016)*

Example 4. *Through this ongoing effort, we no longer use or are on our way to eliminating some of the ingredients currently raising consumer questions. (COLGATE, 2015)*

Example 5. *In addition, we have been implementing several programs to ensure that our employees can pursue both parenting and professional growth. (Amorepacific Group, 2016)*

In Example 1, the drafter used “*more than + an approximate number*” instead of the specific number, which is an estimation of quantity. The VL in this example enables the drafter to avoid providing accurate data but does not affect the dissemination of sufficient CSR information at the same time, as it is unnecessary for the target audience to know the specific figures. Another possible explanation for the use of this structure is that the Procter & Gamble (P&G) company does not have specific numbers either. The intentional vagueness may be used to make up for a lack of information and data.

In Examples 2 and 3, several instances such as *many challenges* and *in various global markets* were found to make exact challenges and markets unclear. While these usages of VL decrease the precision of language, it can be inferred that COLGATE and the Amorepacific Group tend to emphasise the efforts they have invested for the prevention of AIDS. The strategic employment of these linguistic devices allows drafters to summarise the current situation of AIDS prevention and the company’s globalisation

process and maintain an optimum degree of elaboration on the corporate action.

Example 4 involves the vague word *some* (e.g., *some* of the ingredients). In the COLGATE report, relevant ingredients of cosmetic products and the specific amount of them were not listed. In this context, the drafter intentionally used VL to withhold ingredient information closely related to quality problems in their cosmetics products. Instead of mentioning specific figures, VL was employed as a strategic linguistic tool to avoid disclosing internal negative information that may threaten its legitimacy in socially responsible actions and cause image crises.

In Example 5, VL provides a generality of meaning. *Several programs* conveys a general meaning instead of a specific one. Amorepacific Group tends to convey that they are committed to helping their employees achieve a work–family balance. “Several” suggests that they invested consistent efforts in establishing programmes, regardless of the exact number.

Furthermore, an emerging frequency of the pattern “*a + singular quantifier + of*” should be noted, such as *a variety of*, *a range of*, *a series of*, and *a number of*, which were also noted in Arinas’s (2012) study that similar expressions can provide a flexible interpretation of the property defined in patents. Here are some examples of the use of *a range of* and *a series of*:

Example 6. *Throughout the year, Amorepacific Group undertakes a range of activities to effectively respond to various risks identified through our risk mitigation and management process. (Amorepacific Group, 2016)*

Example 7. *At our Vision Care facility in Limerick, Ireland, a 3.0 MW wind turbine completed in early 2016 is the latest project in a series of energy use and CO2 emission reduction efforts ongoing since 2006, which through the end of 2015 have reduced the energy cost and CO2 emissions at that location by nearly 30 percent. (Johnson & Johnson, 2016)*

In these two examples, the vague references *a series of* and *a range of* do not over-elaborate on the exact figures accounting for the company’s CSR performance. Internal and external audiences are presented with fuzzy accounts, with only an appropriate amount of information disseminated, indicating the company’s operational and strategic decisions.

Lexical item	Freq	Freq per thousand
more than	760	1.09
many	434	0.62
various	392	0.56
some	294	0.42
several	210	0.30
approximately	205	0.29
nearly	139	0.20
amount of	137	0.20
a variety of	122	0.17
millions of	107	0.15
at least	103	0.15
a range of	84	0.12
numerous	78	0.11
much	62	0.09
almost	62	0.09
a series of	55	0.08
thousands of	53	0.08
or more	38	0.05
most of	33	0.05
a number of	30	0.04

Table 2. Vague items associated with “quantity”.

6.2. Vague items associated with “degree”

The second group of vague expressions is associated with *degree* such as *important, well, significant, better, appropriate, effective, and good*, which follows “quantity” in terms of frequency. The high occurrence of these positive evaluative items echoes previous research showing that companies seek to gain trust from stakeholders (Fuoli, 2017). The company attempts to exemplify organisational competence and rational justifications for its decisions or prospects through adjectives and adverbs such as *important, well, and significant*, and it demonstrates readiness to carefully assess impacts of planning and future scenarios through adjectives such as *relevant and appropriate*.

Example 8. *Through this work, Colgate will be an important catalyst to increase the transparency of supplier companies in the oleochemical industry. (COLGATE, 2017)*

Example 9. *Although we missed our Healthy Future 2015 goal to achieve a 10 percent absolute reduction in water use by 2015, with our 2010 consumption as a baseline, compared to our performance in 2013 and 2014 we made **significant** advancement in water reduction in 2015. (Johnson & Johnson, 2015)*

Example 10. *L'Oréal convened a panel of 12 international experts, who contributed to developing the methodology – in line with European Product Environmental Footprint standards – in order to ensure that the final methodology would lead to a relevant and robust assessment of impacts. (L'Oréal, 2018)*

In Example 8, the drafter used *an important catalyst* to describe Colgate's role in transparency improvement. However, the use of *important* is vague here because there are no specific criteria to delimit whether corporate efforts are important. Likewise, the use of *significant* in Example 9 also makes the meaning of the sentence fuzzy. *Significant* is used to describe the large number of products shipped by the company, demonstrating its overall power in production.

Example 10 uses *very* and *well* to cover a wide range of staff health check-up issues informed by. In the absence of facts and data, *very* and *well* can be seen as a strategic use to praise their social responsibility behaviours, aiming to construct a positive and responsible image.

Moreover, in Table 2, it is apparent that most of these vague expressions (e.g., *important*, *well*, *significant*, *good*, *appropriate*, *effective*, and *great*) are degree markers. The companies made more discursive efforts to demonstrate the current situation of socially responsible behaviours by addressing stakeholders' concerns and needs through concrete language expressions.

Lexical item	Freq	Freq per thousand
important	377	0.54
well	356	0.51
significant	328	0.47
better	283	0.40
relevant	217	0.31
appropriate	173	0.25
effective	161	0.23
very	158	0.23
main	134	0.19
good	106	0.15
necessary	102	0.15
closely	95	0.14
essential	92	0.13
general	92	0.13
significantly	88	0.13
great	84	0.12
vital	83	0.12
fundamental	72	0.10

Table 3. Vague items associated with “degree”.

6.3. Vague items associated with “time”

Another major group of vague expressions is associated with “time” or “frequency,” as demonstrated by the use of expressions in this group, including *early*, *recent*, *often*, *by the end of*, and *always*. Compared to the other three categories, it seemed that they were rarely observed. This might be explained by the fact that CSR reports are often overtly promotional with the frequent use of positive evaluative expressions (Livesey & Kearins, 2002). Vague time references have been found to be prevalent in legal texts for other reasons (Li, 2017). here, they appear to play a specific part related to management of the corporate image. Given CSR’s role in both reporting the regular information over the year and predicting the future, they provide leeway for the company to deal with strategic decision making or fulfil its commitment.

Example 11. *Efforts in 2015 include two fuel cell projects in California, and the construction of a new windmill project in Ireland, which was completed in early 2016.* (PROCTER & GAMBLE, 2015)

Example 12. *In Japan's cosmetic market, which is now mature, new entrants have entered from other industries in recent years, and, as a result, competition has intensified.* (KOŠÉ, 2018)

Example 13. *Many employees cycle to work, which is good for the environment and their fitness; however, they often ride without helmets, although these are vital for their personal safety in rush-hour traffic.* (Beiersdorf, 2016)

In Example 11, the vague item *early* is related to the time range. The start and end dates of *early 2016* are not clarified. *Early 2016* in a general sense may refer to the first two or three months of 2016. It is unnecessary to specify the exact date of completion of the new windmill project because *early 2016* as a time reference would not prohibit the smooth understanding of target readers. Besides, it is very likely that the drafter would not know the exact date.

Similarly, Example 12 demonstrates the vague term *recent*, which can refer to several years ago or last year. A possible explanation for the use of *recent* is that it is difficult to determine which year new entrants from other industries enter Japan's market, which is a continuous process.

In Example 13, *often* suggests that the frequency of employees who ride bicycles without helmets is relatively high, even though the specific rate of those who ride without helmets may not be officially counted by the corporation. In the absence of specific data and information, it is wise to use vague items like *often* to express the same meaning. Moreover, the company uses *often* to emphasise the seriousness of current safety problems, highlighting the role of the safe cycling project launched by the company.

Lexical item	Freq	Freq per thousand
early	155	0.22
recent	100	0.14
often	98	0.14
by the end of	89	0.13
always	87	0.12
recently	47	0.07
at the end of	30	0.04
usually	16	0.02
sometimes	16	0.02
frequently	11	0.02

Table 4. Vague items associated with “time”.

6.4. Vague items associated with “softening stance-taking”

VL in this category is manifested by the use of perspective-taking verbs (*suggest, think*) and possibility indicators such as *possible* and *may*. The emergence of these vague items may reflect a general tendency that CSR reports are subjective. These linguistic manifestations project reasoned justification for company decision and engagement in CSR activities, while also leaving adequate space for negotiations in the case of unforeseen circumstances or negative evaluations, thus revealing the “dialogic nature” of corporate discourse (Fuoli, 2017). The examples below may be interpreted as a responsive discursive strategy aiming at displaying their careful engagement with social and environmental performances.

Example 14. *Although the quality of Kosé’s products is world-class, I think it is crucial to carry out branding that achieves universal global acceptance. Kosé will do this not by giving up its unique qualities but by enhancing them as it flexibly responds to changing social conditions. (Kosé,2019)*

Example 15. *In 2018, OHSAS 18001 will probably be replaced by the international standard ISO 45001, and we already integrated the expanded requirements associated with this into our internal checklist for our ESMAS audit program in 2016... (LVHM, 2017)*

In Example 14, compared to the likelihood verb *believe* expressing a strong sense of obligation and accountability, *think* conveys a weaker degree of

engagement. Thus, the statement signals an emerging ambition on the part of *Kosé* to promote its branding globally, but also leaves the possibility open if *Kosé* fails to fulfil this expectation.

In Example 15, the use of *probably* implies the drafter's unwillingness to commit him/herself concerning whether the statement is true or false. However, VL here gives stakeholders a clear message that companies are actively adjusting themselves to cope with changes given the relevant standards. Meanwhile, LVHM leaves some room for itself in case what they predict does not happen.

Lexical item	Freq	Freq per thousand
probably	228	0.33
possible	201	0.29
may	134	0.19
assume	109	0.16
estimate	94	0.13
think	38	0.05
might	29	0.04
suggest	21	0.03

Table 5. Vague items associated with "softening stance-taking".

6.5. Communicative functions of vague language

From the above discussion, it can be seen that the drafters of CSR reports can benefit from the use of VL. Overall, five major situations can be summarised as to how drafters tend to strategically use VL, with adaptations based on Li (2017) and Zhang's (2018) work: to give an appropriate amount of information, demonstrate solidarity, enhance persuasion, and self-protection.

- a) To give an appropriate amount of information. In drafting a CSR text, writers are expected to provide relevant information on CSR-related development as an annual highlight to stakeholders for future investment, on the one hand. On the other hand, there will always be unexpected circumstances or future contingencies so that companies

are unable to inject precise details and thus resort to VL for self-protection. There may even be a genuine lack of precise information on the part of corporate management. The vague expressions *some* or *many* in Examples 2 and 4, on the one hand, suggest that companies pay attention to protecting the information that is crucial to cover themselves. While precise information can be obtained if corporations spend money and time collecting missing information, the information-seeking process may result in an imbalance between investment and economic return. It is thus likely that unclear referencing is an effective cost-saving strategy. A similar finding can be seen in Li's (2017) study on VL in legal texts, in which she argued that draftsmen may not have time to check and access exact information, which thus results in the emergence of VL.

- b) To enhance persuasion. In this corpus, we can see that drafters often use vague items associated with stance markers (e.g., *important*, *very*, *essential*). VL in this category comprises positive evaluative language items mainly in the form of adjectives to discursively build up a committed image, establish a positive relationship with stakeholders, and underscore their annual contribution to social and environmental issues (Bondi, 2016; Fuoli, 2017). For instance, Colgate seeks to demonstrate the active role they lead by signalling a confident attitude, for example by using *important* as a stance presentation to persuade stakeholders (Example 8).
- c) To protect one's image. To build a favorable corporate image, companies may downplay unfavourable information or soften their position through strategic use of VL. As shown in Examples 14 and 15, concerning prospects and technological advancement, companies use VL to soften their stances regarding the blueprint of the product. Rather than elaborate on the advancement of production standards, the company presents itself as a rational decision-maker with cautious attitudes, leaving negotiable space for not fulfilling the objectives. Also, VL cushions the negative effects of potential criticisms. For instance, as evidenced by the vague term *some* in "experienced *some* difficulties from excessive store openings due to business expansion", *some* can be construed as a face-saving discursive strategy to maintain a positive corporate image

7. Summary and conclusions

The sections above address the following research questions: What are the prototypical types of VL in CSR reports of the cosmetics industry? Is the use of VL in CSR reports related to their communicative purposes? If so, how? Four types and four communicative functions were illustrated. The linguistic representations of VL are associated with the semantic categories “quantity,” “degree,” “time,” and “softening stance-taking.” The communicative functions underlying the linguistic findings are 1) Give an appropriate amount of information; 2) Enhance persuasion; and 3) Self-protection. I arrived at these findings through a CADS approach by analysing CSR through corpus methods; at the same time, the contextual analysis surrounding VL use in their contexts sensitised me to the salience of some of the communicative strategies. The approach helps to uncover how cosmetics companies communicate socially responsible initiatives through the discursive constructions of VL in CSRs. The semantic types and pragmatic functions of VL align with the findings proposed in prior research (Li, 2017; Zhang, 2018) and confirm some observations made in previous research on the promotional and subjective nature of CSR reports, as evidenced by the positive evaluative expressions and softening stance expressions.

The study also contributes to an enhanced understanding of VL on the part of business practitioners and business communication scholars, enabling them to critically understand and assess the VL features in corporate discourse. Such findings of these categorisations and pragmatic functions can be transferred into actual drafting practices, such as when a company drafting CSR reports seeks effective and image-enhancing methods. The strategic employment of these linguistic resources may manage readers’ impressions and navigate them toward a favorable direction. Beyond the textual level, given the strategic nature of VL use, the study may bridge the gap between applied linguists and business practitioners, prompting them to be aware of the value and feasibility of collaboration.

Some avenues for future research on VL in corporate discourse can be pursued. Notably, in the current study, VL was identified based on prior studies and sample readings. Several other linguistic manifestations of VL are not considered in this paper (e.g., conditional sentences, passive sentences, and epistemic modality). The paucity of these lexico-grammatical features provides a potential source of inspiration for applied linguists to

uncover more semantic types and pragmatic functions of VL. Further, while I focused on the investigation of VL mainly through frequency as conducted in prior research (Li, 2017; Malyuga & McCarthy, 2018), future research could focus on a more sophisticated quantitative approach (e.g., collocational patterns), which expands the scope of the data analysis. In addition, only English versions of CSRs are included in this research. It is hoped that future research would focus on the VL in both English and non-English CSRs and then identify the potential social, cultural, and organisational factors that may account for the linguistic variation between these two types of reports.

Article history:

Received 03 October 2020

Received in revised 15 February 2022

Accepted 25 March 2022

References

- Anesa, P. (2007). Vagueness and precision in contracts: A close relationship. *Linguistica e Filologia*, 24, 7-38.
- Arinas, I. (2012). How vague can your patent be? Vagueness strategies in U.S. patents. *Hermes – Journal of Language and Communication in Business*, 25(48), 55-74.
- Baker, P., & McEnery, T. (2015). Introduction. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating discourse and corpora* (pp. 244-265). Palgrave Macmillan.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Bondi, M. (2016). "The future in reports: Prediction, commitment and legitimization in CSR". *Pragmatics and Society*, 7, 57-81.
- Bondi, M., & Yu, D. (2017). The generic structure of CSR reports in Italian, Chinese, and English: A corpus-based analysis. *IEEE Transactions on Professional Communication*, 99, 1-19.
- Bhatia, V. K. (2004). *Worlds of written discourse: A genre-based view*. Continuum.
- Bhatia, B. (2012). The corporate social responsibility report: the hybridization of a "confused" genre (2007-2011). *IEEE Transactions on Professional Communication*, 55, 221-238.
- Channell, J. (1994). *Vague language*. Oxford University Press.
- Cheng W., & O'Keeffe, A. (2015). Vagueness. In K. Aijmer & C. Rühlemann (Eds.), *Corpus pragmatics: A handbook* (pp. 360-378). Cambridge University Press.
- Choi, A., & Triantis, G. (2010). Strategic vagueness in contract design: The case of corporate acquisitions. *The Yale Law Journal*, 119, 848-924.
- Cotterill, J. (2007). 'I Think He Was Kind of Shouting or Something': Uses and abuses of vagueness in the British courtroom. In J. Cutting (Ed.), *Vague language explored* (pp. 97-114). Palgrave Macmillan.
- Cutting, J. (2007). *Vague Language Explored*. London: Palgrave Macmillan.
- Connor, U., & Upton, T.A. (2004). *Discourse in the professions: Perspectives from corpus linguistics*. John Benjamins.
- Ditlevsen, M. G. (2012). Revealing corporate identities in annual reports. *Corporate Communications: An International Journal*, 17(3), 379-403.
- Ditlevsen, M. G., Nielsen, A. E., & Thomsen, C. (2013). Corporate reporting: An integrated approach to legitimacy claims. *Journal of Modern Accounting and Auditing*, 9(12), 1637-1643.
- Ecovia Intelligence. (2010). CSR & sustainability: How the beauty industry is cleaning up. <<https://www.ecovaint.com/r1805/>>
- Evison, J., McCarthy, M., & O'Keeffe, A. (2007).

- 'Looking out for love and all the rest of it': vague category markers as shared social space. In J. Cutting (Ed.), *Vague language explored* (pp. 138-157). Palgrave Macmillan.
- European Commission (2011). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A renewed EU strategy 2011-14 for corporate social responsibility* (COM (2011) 681). <<https://publications.europa.eu/en/publication-detail/-/publication/ae5ada03-0dc3-48f8-9a32-0460e65ba7ed/language-en>>
- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24, 321-332.
- Fordham, A. E., & Robinson, G. M. (2018). Mapping meanings of corporate social responsibility – an Australian case study. *International Journal of Corporate Social Responsibility*, 3(1), 1-20.
- Fuoli, M. (2012). Assessing social responsibility: A quantitative analysis of appraisal in BP's and IKEA's social reports. *Discourse & Communication*, 6(1), 55-81.
- Fuoli, M. (2017). Building a trustworthy corporate identity: A corpus-based analysis of stance in annual and corporate social responsibility reports. *Applied Linguistics*, 39(6), 846-885.
- Fuoli, M., & Hart, C. (2018). Trust-building strategies in corporate discourse: An experimental study. *Discourse & Society*, 29(1), 514-552.
- Gruber, H. (1993). Political language and textual vagueness. *Pragmatics*, 3, 1-28.
- Handford, M. (2010). *The language of business meetings*. Cambridge University Press.
- Kushal, S. J. (2011). *Business communication*. V.K. Global.
- L'Oreal. (2018). Annual report. <<https://www.loreal-finance.com/en/annual-report-2018/>>
- Li, S. (2017). A corpus-based study of vague language in legislative texts: Strategic use of vague terms. *English for Specific Purposes*, 45, 98-109.
- Liu, Y. (2015). *A paradigm for business communication across cultures: Theoretical highlights for practice*. Tsinghua University Press.
- Lin, Y. (2021). Legitimation strategies in corporate discourse: A comparison of UK and Chinese corporate social responsibility reports. *Journal of Pragmatics*, 177(2), 157-169.
- Livesey, S., & Kearins, K. (2002). Transparent and caring corporations? A study of sustainability reports by the body shop and royal Dutch/Shell. *Organization and Environment*, 15, 233-258.
- Maignan, I., Ferrell, O. C., & Ferrell, L. (2005). A stakeholder model for implementing social responsibility in marketing. *European Journal of Marketing*, 39(9/10), 956-977.
- Malavasi, D. (2011). 'Doing Well by Doing Good': A Comparative Analysis of Nokia's and Ericsson's Corporate Social Responsibility Reports. In G. Garzone & M. Gotti (Eds.), *Discourse, communication and the enterprise: Genres and trends* (pp. 193-212). Peter Lang.
- Malyuga, E., & McCarthy, M. (2018). English and Russian vague category markers in business discourse: linguistic identity aspects. *Journal of Pragmatics*, 135, 39-52.
- McCarthy, M. (2019). Vague language in business and academic contexts. *Language Teaching*, 53(2), 203-214.
- Mellinkoff, D. (1963). *The language of the law*. Little Brown & Co.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. John Benjamins.
- Pinkal, M. (1995). *Logic and lexicon*. Kluwer.
- Russell, B. (1997). Vagueness: A reader. In R. Keefe & P. Smith (Eds.), *Vagueness: A reader* (pp. 61-68). MIT Press. *Australasian Journal of Philosophy and Psychology*, 1, 84-92. (Original work published 1923).
- Ruzaitė, J. (2007). *Vague language in educational settings: Quantifiers and approximators in British and American English*. Peter Lang.
- Scotto di Carlo, G. (2015). 'Weasel Words' in legal and diplomatic discourse: Vague nouns and phrases in UN resolutions relating to the Second Gulf War. *International Journal for the Semiotics of Law*, 28, 559-576.
- Skulstad, A. (2008). Creating a 'green' image in the public sphere: Corporate environmental reports in a genre perspective. In R. Wodak & V. Koller (Eds.), *Handbook of communication in the public sphere* (pp. 181-201). De Gruyter.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Language. Cambridge University Press.
- van Riel, C. B. M., & C. Fombrun (2007).

Essentials of corporate communication: Implementing practices for effective reputation management. Routledge.

Widdowson, H. G. (2004). *Text, context, pretext. Critical issues in Critical Discourse Analysis.* Blackwell.

Zhang, G. (2011). Elasticity of vague language. *Intercultural Pragmatics*, 8(4), 571-599.

Zhang, G. (2018). *Vague language, elasticity theory and the use of "some": A comparative study of L1 and L2 speakers in educational settings: linguistic analysis of some.* Bloomsbury Publishing Plc.

Bixi Jin received her doctoral degree in English language Education at the University of Hong Kong. She is now affiliated with the School of Foreign Languages, Beihang University. Her research interests include corpus linguistics, discourse analysis, and computational social science. Her recent work appeared in *IEEE Transactions on Professional Communication* and *Ibérica*.

NOTES

¹ The relationship between vagueness and ambiguity is far more complex and the distinction between the two linguistic phenomena may be blurred, but this is beyond the scope of the current study.

² *Or other, or more, and or any* were identified as the VL clusters through N-grams in Sketch Engine.

³ The frequency of each vague item (normalised to per thousand word) obtained from the current study may not be directly comparable to the frequencies in previous studies (normalised to per million word).

Appendix

Rank	Word	Freq
1	the	33476
2	and	31387
3	of	23892
4	to	21888
5	in	18958
6	our	13426
7	a	10643
8	for	9223
9	we	8087
10	with	7434
11	on	5799
12	is	5256
13	by	4816
14	as	4758
15	are	4449
16	button	4421
17	that	4024
18	this	3285
19	sustainability	3280
20	from	3199
21	at	3125
22	health	2855
23	employees	2744
23	products	2650
25	have	2489
26	an	2437
27	management	2430
28	their	2383
29	business	2373
30	all	2310
31	more	2234
32	its	2156

33	water	2145
34	global	2102
35	people	2049
36	Johnson	2029
37	or	1936
38	also	1920
39	new	1913
40	has	1906
41	group	1809
42	through	1748
43	environmental	1733
44	safety	1659
45	product	1642
46	which	1619
47	Colgate	1580
48	social	1567
49	care	1563
50	program	1541

Table A1. Top 50 words from the wordlist of the CSR (cosmeics industry).

Rank	Word	Keyness
1	Colgate	1258.82
2	Henkel	1072.857
3	GRI	709.54
4	Kosé	694.74
5	Amorepacific	694.304
6	LVMH	591.087
7	Beiersdorf	490.015
8	tesa	303.321
9	Gorsky	207.69
10	sustainability	148.45
11	Maisons	134.657
12	CDP	125.034
13	RSPO	120.58
14	Innisfree	117.819
15	EHS	116.228
16	UNGC	110.451
17	Janssen	98.977
18	Sekkisei	98.587
19	Lauder	97.139
20	KRW	96.398
21	GHG	95.904
22	Estée	93.972
23	Dior	85.806
23	Hennessy	79.975
25	Decorté	79.943
26	handwashing	78.545
27	supplier	76.195
28	earthwards	74.601
29	button	73.426
30	BSBF	72.252
31	recyclability	71.146
32	stewardship	70.529
33	parfums	70.351
34	CSR	69.849

35	Sephora	69.783
36	G4-DMA	69.035
37	Nivea	68.187
38	Guerlain	66.518
39	FY16	64.991
40	Credo	64.936
41	Bertuti	61.269
42	foreword	61.239
43	ELC	60.923
44	Co2e	60.141
45	packaging	60
46	sustainably	56.277
47	Ecovadis	55.83
48	Terracycle	53.645
49	deforestation	53.129
50	Johnson	52.889

Table A2. Top 50 words from the keyword list of the CSR (using the English Web 2020 in Sketch Engine as the reference corpus).