

Kim, Kyoungmi (2016) Not so 'innocent' after all? Exploring corporate identity construction online. *Discourse & Communication*, 10 (3). pp. 291-313.

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/5353/>

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version:

<http://dx.doi.org/10.1177/1750481315623902>

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. [Institutional Repositories Policy Statement](#)

# RaY

Research at the University of York St John

For more information please contact RaY at  
[ray@yorks.ac.uk](mailto:ray@yorks.ac.uk)

## PAPER FOR DISCOURSE & COMMUNICATION

**Title:** Not so '*innocent*' after all? Exploring corporate identity construction online

**Authors:**

The PAD research group

The following members of the Professional and Academic Discourse (PAD) research group at *The University of Warwick* have contributed to the research reported in this paper: Stephanie Schnurr & Sue Wharton (co-ordinators), Suha Alansari, Jo Angouri, Rachel Chimbwete Phiri, Teresa Chiriatti, Carolin Debray, Christina Efthymiadou, Sixian Hah, Tilly Harrison, Ana Kedveš, Attapol Khamkhien, Kyoung-mi Kim, Nor Azikin Mohd Omar, Thi Hong Nhung Nguyen, and Sophie Reissner-Roubicek.

**Version number and date:** revised version (first revision) 05 October 2015

**Corresponding Author:**

Stephanie Schnurr  
Centre for Applied Linguistics  
University of Warwick  
Social Science Building  
CV4 7AL, UK.  
Email: [S.Schnurr@warwick.ac.uk](mailto:S.Schnurr@warwick.ac.uk)

**Short title:** Corporate identity construction online

## **Abstract**

While a lot of attention has been paid to online branding and the construction and communication of a company's identity via its website (e.g. Bravo et al., 2013; Halliburton and Ziegfeld, 2009; Pollach, 2005), there is only very little research that looks at the processes involved in these activities from a discourse analytical perspective. This paper aims to address this gap by conducting a case study of *innocent*, a UK producer of fruit juices. Combining corpus analytical tools with discourse analytical techniques and considering both text and multimodal features, we explore some of the strategies through which *innocent* creates a set of inter-related and closely intertwined identities on its website thereby constructing the company's brand image. However, our findings also reveal that some of the company's identity claims (especially in relation to being an inclusive and welcoming 'family') are relativised and to some extent contradicted by the discursive processes through which these claims are articulated.

## **Keywords**

Identity construction, website, discourse analysis, corpus analysis

**Bio note**

The Professional and Academic Discourse research group (PAD) is a group of academic staff and doctoral students based at the Centre for Applied Linguistics, University of Warwick. We are all interested in discourse analysis, from a range of different perspectives. We meet regularly for talks and to support each other in our research, as well as to work on projects together. The paper published here stems from collaborative work between January and July 2015. To find out more about our activities and the publications of various members please go to <http://www2.warwick.ac.uk/fac/soc/al/research/groups/pad/>

# Not so ‘*innocent*’ after all? Exploring corporate identity construction online

The PAD research group<sup>1</sup>

## Abstract

While a lot of attention has been paid to online branding and the construction and communication of a company’s identity via its website (e.g. Bravo et al., 2013; Halliburton and Ziegfeld, 2009; Pollach, 2005), there is only very little research that looks at the processes involved in these activities from a discourse analytical perspective. This paper aims to address this gap by conducting a case study of *innocent*, a UK producer of fruit juices. Combining corpus analytical tools with discourse analytical techniques and considering both text and multimodal features, we explore some of the strategies through which *innocent* creates a set of inter-related and closely intertwined identities on its website thereby constructing the company’s brand image. However, our findings also reveal that some of the company’s identity claims (especially in relation to being an inclusive and welcoming ‘family’) are relativised and to some extent contradicted by the discursive processes through which these claims are articulated.

## Keywords

Identity construction, website, discourse analysis, corpus analysis

---

### Corresponding Author:

Stephanie Schnurr, Centre for Applied Linguistics, University of Warwick, Social Science Building, CV4 7AL, UK.

Email: s.schnurr@warwick.ac.uk

---

<sup>1</sup> The following members of the Professional and Academic Discourse (PAD) research group at *The University of Warwick* have contributed to the research reported in this paper: Stephanie Schnurr & Sue Wharton (co-ordinators), Suha Alansari, Jo Angouri, Rachel Chimbwete Phiri, Teresa Chiriatti, Carolin Debray, Christina Efthymiadou, Sixian Hah, Tilly Harrison, Ana Kedveš, Attapol Khamkhien, Kyoung-mi Kim, Nor Azikin Mohd Omar, Thi Hong Nhung Nguyen, and Sophie Reissner-Roubicek.

## **1. Introduction**

This paper aims to contribute to research on identity construction by exploring some of the discursive processes involved in the online identity construction of *innocent*, a UK company which caters for the European market. While a lot of scholarly attention is currently paid to identity construction, especially as an aspect of professional communication (Schnurr, 2013), most of this research focuses on spoken interactions (e.g. the contributions in Van De Mierop and Clifton, 2012) or written texts (e.g. Koller, 2007), corporate websites which combine spoken, written and interactive elements, tend to be largely ignored. This paper aims to address this gap and hopes to make a contribution not only in relation to the discursive processes involved in identity construction but also with regards to the methodological procedures facilitating such an undertaking – especially in online environments.

We have chosen *innocent* as the case study for our research not only because it is a very successful company (being the biggest smoothie maker in the UK (Anderson, 2014) with an interesting business story), but also because questions relating to its brand authenticity are frequently discussed in the media. Thus, what makes *innocent* so interesting from the angle of identity construction is the ways in which the company has successfully managed its brand over the years and has maintained the positive and

health-conscious image that it propagates for itself – even in the face of several severe challenges.

### *1.1 innocent – a brief background*

Having been established in 1999, *innocent* fruit smoothies were first introduced to the UK market in 2000. Based on immediate success, expansions to Ireland and other European countries followed soon after, and the range of products was expanded. In 2004 the company set up an independent registered charity called the Innocent Foundation, supported by the donations of the company itself (Innocent Foundation, n.d.).

One of the unique selling points of the company is its philosophy of '*doing good*' and taking corporate social responsibility seriously. This high emphasis that *innocent*, like many other multinational corporations (e.g. Fuoli, 2012), puts on ethics is reflected, for example, in the values of the company which include not only '*be natural*', but also '*be responsible*' and '*be generous*' – with reference to the planet, the environment, the people and the community. Through various sections on its website, *innocent* portrays itself as an ethical, sustainable and even charitable company. These claims for an ethical and responsible identity are, of course, already reflected in the name of the company

itself where the adjective 'innocent' immediately evokes associations of 'pure', 'natural', and even 'good' as opposed to 'artificial' and 'guilty'. This impression is further strengthened by the company's logo, which is a halo over a smiley face. Taken together, the name and the logo thus already provide strong indications for the kind of brand image or identity the company wants to portray for itself. This identity construction is clearly strategically motivated and targeted specifically at affluent customers – a move which has paid off for *innocent* (Johnston et al., 2011; Starr, 2009).

However, the road to success has not always been without bumps for *innocent* and the company's image as ethical and innocent, was, in fact, challenged by customers at various points. Some of the more serious challenges to the identity claims that *innocent* make about themselves, occurred when the company announced its partnership with *McDonald's* (in 2007) and *Coca-Cola* (in 2009). Although joining forces with *McDonald's* seemed like a lucrative move, it was perceived very negatively by *innocent's* customers who saw it as a loss of independence and threat to the ethical values that *innocent* is so actively promoting on its website. The excerpts below are representative of the way customers rejected the company's identity claims at that moment in time:

*"This is a lesson in how to destroy a fantastic brand profile with one act of greed.*

*Partnering with McDonald's - a firm with zero moral or ethical beliefs - has sullied*

*your company's reputation. [...] I certainly won't be buying 'Innocent' products ever again."*

In the aftermath of having joined forces with *McDonald's*, *innocent* was even described as having lost its soul (Sweeney, 2007). Similar or even more extreme sentiments were expressed when *innocent* allowed *Coca-Cola* to buy considerable shares of the company:

*"Some of us who have been here throughout the 10 years to date were buying your products precisely because there was no link to grandiose, greedy corporations such as Coke. Your strategy is foolish and self-destructive. You will regret it, as many of us do now. Adieu!"*

Interestingly, in both instances *innocent* immediately 'confessed' their actions with detailed explanations of their reasons and allowed open debate on their blog, cultivating the impression of transparency as part of their ethical stance. We return to this point later in our qualitative analysis.

## *1.2 Creating a brand and constructing identities – online*

Constructing and maintaining a unique brand and communicating it successfully to various audiences (including stakeholders, as well as (current and potential) clients and

staff) is an important, and potentially income-generating, aspect of any company's strategy. In this context, identity is often understood as a set of meanings that allow consumers, stakeholders and others to associate, describe, relate to and establish and maintain links with a company (Balmer 2001; Balmer and Soenen, 1999). As Cheney and Christensen (1999) argue, identity is a pressing issue for many institutions and questions of identity, or of what the organisation is or stands for, cut across and bring together many different organisational goals and concerns (see also Myers, 1994).

Considerable attention has been paid in the previous literature on the internet as a marketing and branding tool. For example Simmons (2007), who understands branding as the process of creating value through the provision of a compelling and consistent offer and customer experience, maintains that as customers develop trust in a brand, companies actively invest in building relationships with their customers thereby further strengthening their own brand. However, in spite of this interest in the internet as a branding tool (see also: Park et al., 2005; Pollach, 2005); and an increasing recognition of the benefits of discursive approaches to identity construction (e.g. Bucholtz and Hall, 2005), there are only very few studies that take a discourse analytical approach to exploring some of the strategies through which identity construction is accomplished in these online environments. One of these studies is Koller (2007) who analysed how the *HSBC* banking group represent themselves in their externally oriented discourse as a

'glocal' brand successfully combining discourses of 'the global' and 'the local'. Drawing on various data sources, including websites, brochures, and advertisements, Koller identifies several visual and linguistic strategies through which the bank's image is successfully constructed and communicated to the audience. For example, by strategically including historical imagery and modern photos with easily recognisable historical symbols, as well as the frequent use of specific colours throughout the various documents, and making repeated explicit and implicit references to 'the global' and 'the local' a dichotomy is created between global/modern on the one hand and local/historical on the other. Through these strategies *HSBC* portray themselves as a 'glocal bank', a corporation successfully combining 'the global' and 'the local'. In another study on corporate self-representations of 20 websites, Pollach (2005) analysed the use of language and hypertext resources used to construct favourable images of the company. Among the strategies frequently used on the websites to construct a positive image of the companies were presenting claims as facts, humanising the organisation, removing agents (e.g. by using a passive voice to suggest trustworthiness), use of pronouns and use of interactivity.

In this paper we aim to contribute to this research and take a discourse analytical approach – focusing on text, visual representations and interactive features of the websites (e.g. Kress and Van Leeuwen, 1996) – in order to gain a better understanding

of some of the strategies and processes through which identity construction is performed in these online environments. Moreover, by combining quantitative and qualitative methods, as we outline in more detail below, we aim to discuss some of the methodological challenges that arise when doing research on discourse and communication in online environments, and we illustrate how these challenges might be addressed and overcome by strategically drawing on the tools and procedures of each method.

## **2. Identity construction – a discourse perspective**

Moving away from early work in the field that treated identity as a static and fixed attribute of speakers, we take a social constructionist stance and conceive identity in more dynamic terms. In line with more recent research, we view identity as socially constructed in and through discourse and communication, and as constantly negotiated between participants (e.g. Benwell and Stokoe, 2006; De Fina, 2010). Identity is understood as a process, as something people ‘do’ or ‘perform’, rather than a static attribute that they ‘have’ (Bamberg et al., 2006, 2011; Bucholtz and Hall, 2005; Butler, 1990; Widdicombe, 1998); it is perceived as a context bound and relational phenomenon (Bucholtz and Hall, 2005; Schnurr, 2013).

These re-conceptualisations of identity as a dynamic process or performance lend themselves very nicely to discourse analytical approaches. Indeed, in the past decade or so, discourse analytical research on identity in professional (and other) contexts has increased dramatically. Taking discourse analytical approaches and/or drawing on discourse analytical tools, these studies were able to identify and describe some of the specific processes through which identities are constructed, negotiated, reinforced, as well as sometimes challenged and resisted by interlocutors in a range of different contexts and across different texts and talks (for comprehensive discussions of identity construction in the professional domain see for example the contributions in Angouri & Marra (2011) and van de Mierop & Schnurr (fc)).

One of the most comprehensive frameworks to study identity construction is summarized in Bucholtz and Hall (2005). In their paper the authors outline five principles of identity construction which are based on previous work conducted in several linguistic and anthropological traditions. In this paper we strategically and selectively draw on the relationality and indexicality principles.

The relationality principle, which lies at the heart of the framework (Bucholtz and Hall, 2005: 587), is in line with our social constructionist understanding of identity, and

maintains that identities are relational phenomena which are not constructed in isolation but are always created in relation to other identities. They are

intersubjectively constructed through several, often overlapping, complementary relations, including similarity/difference, genuineness/artifice, and authority/delegitimacy. (Bucholtz and Hall, 2005: 598)

The first pair of relations, similarity/difference, captures the processes through which identities are created by emphasising similarities with or differences from other individuals or groups and by positioning oneself in relation to these others (see also Locher, 2008: 513). The second set of relations, genuineness/artifice, describes the perception of specific identity claims by an audience – i.e. whether the claimed identities are perceived as genuine or artificial and even fake. The third pair of identity relations, authority/delegitimacy, refers to whether institutionalized notions of power and ideology either affirm and impose identities (in the case of authority), or (in the case of delegitimacy) dismiss and reject claims for identities.

The indexicality principle is not so much concerned with the relational nature of identities but rather describes some of the specific processes through which identities are created. It rests on the assumption that when people interact with each other they do identity work. In other words, during their interaction they constantly claim identities

for themselves, project identities onto their interlocutors, and at the same time they accept, reject, and challenge certain types of identities that are projected on them. This identity work takes place not only by explicit referral to specific identity categories (such as ‘family’ or ‘friend’ in the case of *innocent*) but also more implicitly by mobilising signs that are normatively associated with these categories. The notion of indexicality captures some of the processes through which particular stances are evoked through the use of particular linguistic forms and styles, and how those stances, in turn, are associated with (and hence create) certain identities (Ochs, 1992: 341). Indexicality is thus the process of relating certain linguistic forms to “complex systems of meaning, such as ideologies, social representations about group membership, social roles and attributes” etc. (De Fina, 2010: 215). Some of the indexical processes mentioned in Bucholtz and Hall (2005) include overt mentioning of identity categories and labels, evaluative and epistemic orientations, interactional footings and participant roles, as well as the use of linguistic structures and systems that are ideologically associated with specific personas and groups (Bucholtz and Hall, 2005: 594). However, none of these linguistic forms and styles directly indexes a specific identity – rather, they evoke specific interactional stances, which in turn are associated with particular identities (Ochs, 1992, 1993); and any utterance or linguistic form may contribute to the construction of more than one identity at the same time (Holmes et al., 1999: 353).

Although Bucholtz and Hall's (2005) framework primarily describes the processes involved in identity construction that takes place in (spoken) interactions, it has successfully been used in the past to capture the processes involved in constructing an organisation's identity (e.g. McEntee-Atalianis, 2013). We use the two principles described above and employ discourse analytical and corpus analytical methods to analyse some of the processes involved in the online identity construction of *innocent* on its website.

We also selectively draw on Membership Categorisation Analysis (henceforth MCA) to capture the discursive processes through which identity categories are created and membership is assigned to particular agents. MCA has its roots in the early work of Sacks (1992) who "saw the construction of social identities as closely associated to the process of relating categories to specific relationships and activities" (De Fina 2010: 209). These categories are relevant for identity construction as they carry a set of specific activities and attributes associated with them, which are then projected onto their members (Benwell & Stokoe 2006: 39).

### **3. Data and methodology**

In order to explore some of the ways in which identities are constructed online and to address the research aims and objectives outlined above, this paper reports on a case study of the websites of *innocent*, and combines quantitative and qualitative methods (see e.g. Baker & Levon, 2015 for the advantages of such an approach). Some of the benefits of using such a mixed methods approach for identity research were pointed out by van de Mieroop (2007: 1122) who maintained that a quantitative analysis provides “a bird’s eye view of the presence of identity” thus offering initial insights into the data – which is particularly useful when dealing with relatively large data sets, as in our case – which can then help identifying potentially interesting sites for further scrutiny. A qualitative analysis, on the other hand, provides more detailed insights into the specific ways in which identities are constructed in a specific interactional context. We take a similar approach in this paper by combining tools and processes from corpus analysis with discourse analytical techniques.

### *3.1 Building and working with a corpus*

In order to be able to examine patterns of language choice that were repeated over the website as a whole, we decided to extract as much as possible of the plain text and store it as a corpus in which each subpage formed an individual file. However, the notion that

a corpus can be developed from a single website, or can be seen as instantiating the textual content of a website, brings with it some methodological challenges. One of the issues that we had to deal with was the question of how to set boundaries to the target website. The *innocent* website is not, of course, a self-contained entity but includes many links to external sites. It is therefore necessary firstly to articulate a principled definition of the target website for the purposes of this research, and secondly to find practical ways to harvest text from those pages which come under this definition.

For this research, we decided to work with the website as contextualised within a specific socio-cultural context (the UK) and as representative of a coherent company. We therefore chose to include only those pages which come under the *innocent.co.uk* domain, excluding versions of the site for other countries and excluding websites about activities with which *innocent* may be linked, e.g. charity appeals or festivals. By excluding these external sites, we retained the notion of *innocent's* links with the activity (as expressed on an *innocent.co.uk* page) but were able to form a principled boundary to our corpus.

A second issue that we faced concerns the fact that the *innocent* website regularly changes. Since it was not within the scope of the current research to attempt to capture

and discuss those changes, we chose a specific, albeit random, date for our corpus compilation: 05 May 2015.

A third methodological challenge that we encountered was related to the sheer size of the data. A preliminary look at the *innocent* website indicated that it consisted of several thousand pages and subpages, and so it was necessary to use a webcrawler to identify the urls whose text we wished to include in our corpus. Pre-programmed crawlers are available, but these would typically open every link which could be found on the target website and so may not respect the corpus boundaries which we had attempted to define. We therefore chose to programme our own crawler to open only links leading to other pages on the *innocentdrinks.co.uk* website, excluding links to external sites or to versions of the website for other countries. Programming a custom crawler gave us much more control over our data (Suchomel and Pomikálek, 2012; Baroni and Ueyama, 2006) and gave us confidence that all of the selected urls contained content which was relevant to our corpus design principle.

Having arrived at a list of 4870 urls, we then used the programme SketchEngine to create a corpus (Kilgarriff et al., 2004). SketchEngine includes a relatively simple interface enabling researchers to create a corpus from urls, but it does require a number of decisions to be made: the programme offers options to use whitelist or blacklist

keywords, to exclude files above or below a certain size, and to exclude duplicate content. For our research we wished to be as inclusive as possible within the principled boundaries which we had set – we therefore did not use any of these options, leaving the software to harvest as many of our target urls as possible and allow any repetition of large chunks of text to be represented in our corpus statistics. The resulting corpus consists of 2889 files, representing those urls from which SketchEngine was able to harvest sufficient text, and containing 2,052,019 words.

### *3.2 Discourse analysis*

In the second step of our analysis we used discourse analytical techniques to analyse in more detail some of the trends and patterns observed in the first step. While discourse analysis is a broad umbrella term that signifies a whole range of theoretical traditions in different disciplinary fields (linguistics, politics, sociology, education to name but few), and several different schools and epistemologies have developed as a consequence, this paper is firmly situated in applied and socio- linguistic enquiry and positions itself under the relatively new but established field of Workplace Discourse (Angouri and Marra, 2011). In the analysis to follow, our main site of investigation is the written text on the *innocent* websites but in line with recent work in the field (e.g the various contributions in Scollon and Levine, 2004) we take multimodal features of the interaction into account

and pay specific attention to the ways in which the visual and interactive features of the pages, including pictures, fonts, colours, and weblinks, contribute to meaning construction and, more specifically, to the creation of *innocent's* identity (e.g. Knox, 2009; Kress and van Leeuwen, 1996; van Leeuwen, 2011). These procedures are in line with sociolinguistic research which, although often privileging the analysis of language, acknowledges that identity construction is not only linguistic but that different channels of communication create a semiotic aggregate (Jones, 2014; Scollon and Scollon, 2003). Such a mixed methods approach, combining techniques and procedures from corpus analysis and discourse analysis, we believe, can capture more holistically the processes that contribute to constructing identities on websites.

#### **4. Analysis**

We focus here particularly on indexicality and relationality, as outlined above, and discuss some of the ways in which *innocent* articulates specific identity claims by indexing belonging to specific categories and by constructing the company's identities in relation to others. In particular, through the construction of the *innocent* 'community' a claim of belonging is projected onto the user who is invited to participate in marketing activities that are implemented in different communities on- and off-line. The consumer

of the *innocent* products is thus set up as a co-constructer of the company's identities which goes beyond the purchase of products to participation in (on- and off-line) events, festivals or activities.

#### ***4.1 Findings from the corpus analysis***

In the corpus analysis we employed standard techniques of searching for the most frequent words and n-grams, and searching for keywords. Frequency counts give 'raw' information as to frequency; a keyword analysis, in contrast, compares a focus corpus to a reference corpus to find out which words are unusually frequent in the focus corpus, thus indicating salient aspects of its content or linguistic repertoire.

##### ***4.1.1 Most frequent words.***

Depending on the purpose of the analysis, one way 'in' to a corpus is to make a word list to order the words by frequency. Although the top 20 words of any large corpus are likely to be similar, interesting comparisons can be made based on these insights. In Appendix 1 we compare the top 25 words in the *innocent* corpus with the British National Corpus (BNC) and UK Web as Corpus (ukWaC) in terms of frequency. The two corpora were chosen as examples of general corpora of British English, which although

similar in some ways differ in size and content. The BNC was carefully compiled to be a balanced general corpus of British English published or spoken between 1960s and 1990s (Aston & Burnard 1998; Lew, 2009). ukWaC on the other hand was compiled from websites with the domain name .uk, allowing largely random initial selection but very principled post hoc sorting to maximise useful, non-repeated natural language (Ferraresi et al., 2008).

In the comparison of these corpora we see firstly the similarities between the BNC and ukWaC in terms of the very frequent function words. However after the top five words we start to see that the *innocent* corpus privileges first and second person pronouns, the most frequent being 'you', then 'I' (also found in the BNC top 10). We can also see that the percentage of the corpus these words take up is relatively large compared to the other two corpora. Of the 25 most frequent words in the corpus, personal pronouns constitute 4.88% of the corpus as compared to 1.98% in the BNC and 1.01% in ukWaC. Interestingly for the purposes of this analysis, the *innocent* corpus gives 'we' and 'our' as its 11th and 19th most frequent words. This suggests that in this corpus the company makes frequent reference to itself (exclusive 'we') and possibly to its community (inclusive 'we') – these aspects, and their meaning in terms of identity construction, are discussed in more detail in the qualitative analysis in the next section.

#### *4.1.2 Keywords against ukWaC*

The raw frequency comparison above can be further refined by checking which words are 'key' in our corpus compared to a reference corpus. For this we chose ukWaC, available in Sketchengine. ukWaC is suitable as a reference corpus for our investigation because it is large, over 1.3 billion words, is contemporary, and comes from a similar domain, namely .uk websites (for a full account of corpus creation and comparison with the BNC, see e.g. Ferraresi, Zanchetta, Baroni & Bernardini (2008)). The keyword function excludes words common to both corpora and focuses on the unusually frequent words in the corpus under investigation. Words identified as having the highest keyness (using default parameters in SketchEngine) were as follows:

word	Innocent 5 May		ukWaC		Score
	Freq	Freq/mill <sup>?</sup>	Freq	Freq/mill	
nosy	<a href="#">2,957</a>	1164.2	<a href="#">287</a>	0.2	984.1
DOB	<a href="#">2,882</a>	1134.7	<a href="#">593</a>	0.4	822.8
smoothies	<a href="#">2,694</a>	1060.6	<a href="#">588</a>	0.4	771.0
smoothie	<a href="#">2,150</a>	846.5	<a href="#">519</a>	0.3	635.9
Innocent	<a href="#">1,972</a>	776.4	<a href="#">1,747</a>	1.1	366.7
n't	<a href="#">872</a>	343.3	<a href="#">13</a>	0.0	341.5
Share	<a href="#">6,287</a>	2475.2	<a href="#">10,107</a>	6.5	331.0
Sapodilla	<a href="#">593</a>	233.5	<a href="#">23</a>	0.0	231.1
guardian	<a href="#">2,976</a>	1171.7	<a href="#">7,043</a>	4.5	212.6
Knit	<a href="#">564</a>	222.0	<a href="#">758</a>	0.5	150.1
innocent	<a href="#">4,603</a>	1812.2	<a href="#">18,682</a>	12.0	139.7
veg	<a href="#">1,155</a>	454.7	<a href="#">3,844</a>	2.5	131.5
hoodie	<a href="#">381</a>	150.0	<a href="#">239</a>	0.2	130.9
Sign	<a href="#">2,895</a>	1139.8	<a href="#">13,032</a>	8.4	121.9
cartons	<a href="#">559</a>	220.1	<a href="#">1,301</a>	0.8	120.5
fete	<a href="#">535</a>	210.6	<a href="#">1,261</a>	0.8	117.0
sap	<a href="#">639</a>	251.6	<a href="#">1,889</a>	1.2	114.2
friendship	<a href="#">2,897</a>	1140.6	<a href="#">16,871</a>	10.8	96.6
mangoes	<a href="#">289</a>	113.8	<a href="#">377</a>	0.2	92.4
Sorry	<a href="#">3,097</a>	1219.3	<a href="#">19,675</a>	12.6	89.6

**Figure 1.** Screenshot of SketchEngine’s output of words that are ‘key’ in the *innocent - 05 May 2015* corpus compared to the ukWaC corpus.

As could be expected, many of the strongest keywords are to do with the subject matter of the website, such as ‘*innocent*’, ‘*smoothie*’, ‘*mangoes*’. What is more of interest is the appearance on the list of lexical items which one would not necessarily associate with a fruit drinks website – such as ‘*nosy*’, ‘*DOB*’, ‘*hoodie*’, ‘*friendship*’, and ‘*sorry*’. As we explore in more detail in the next section, all of these items can be seen as forming part of a bigger picture which contributes to constructing *innocent*’s identity.

### 4.1.3 Most frequent n-grams

Taking the investigation of unusual frequency one step further, we generated a frequency list of 6 word n-grams. In previous research, frequent n-grams have tended to be studied as possible indicators of phraseological characteristics of certain genres or language varieties (e.g. Liu, 2012; Simpson-Vlach and Ellis, 2010). In our case, looking at a specific website, we were interested in n-grams as a way of locating larger chunks of text which were repeated throughout the site. The list below shows that there are 17 6-word n-grams which all occur the same number of times:

**Table 2.** The top 6-grams for the *innocent May 05 2015* corpus.

<b>your date of birth to check</b>	<a href="#"><u>2,955</u></a>
<b>we have to ask for your</b>	<a href="#"><u>2,955</u></a>
<b>to do this bit for you</b>	<a href="#"><u>2,955</u></a>
<b>to ask for your date of</b>	<a href="#"><u>2,955</u></a>
<b>please ask a parent or guardian</b>	<a href="#"><u>2,955</u></a>
<b>parent or guardian to do this</b>	<a href="#"><u>2,955</u></a>
<b>or guardian to do this bit</b>	<a href="#"><u>2,955</u></a>
<b>of birth to check that you</b>	<a href="#"><u>2,955</u></a>

<b>if this seems a bit nosy</b>	<a href="#"><u>2,955</u></a>
<b>have to ask for your date</b>	<a href="#"><u>2,955</u></a>
<b>guardian to do this bit for</b>	<a href="#"><u>2,955</u></a>
<b>for your date of birth to</b>	<a href="#"><u>2,955</u></a>
<b>date of birth to check that</b>	<a href="#"><u>2,955</u></a>
<b>ask for your date of birth</b>	<a href="#"><u>2,955</u></a>
<b>ask a parent or guardian to</b>	<a href="#"><u>2,955</u></a>
<b>a parent or guardian to do</b>	<a href="#"><u>2,955</u></a>
<b>Sorry if this seems a bit</b>	<a href="#"><u>2,955</u></a>

By checking the concordance context of these phrases we find that they all come from the same pop up on the website, in which the company apologises, in what seem to be over-polite and clearly humorous ways, for needing to check the date of birth (DOB) of the user:

*sorry if this seems a bit nosy. by law, we have to ask for your date of birth to check that you're over 12 years old. if you're 12 or under, please ask a parent or guardian to do this bit for you. thanks very much.<sup>2</sup>*

---

<sup>2</sup> The exclusive use of lower case is not a coincidence (or mistake) here – a point which we come back to in our qualitative analysis below.

The website obliges any user who wishes to sign up to a service or post blog comments to enter their date of birth; the pop-up appears when the website users click a question mark next to the 'date of birth' field, or when they omit to fill in this field. It explains the unusual frequency of the words '*nosy*', '*guardian*' and '*sorry*' pointed out in the keyword analysis. This n-gram search has thus indicated a textual feature which may – on the grounds of its ubiquity within the website design (it can be accessed from 2955 pages), its interactive nature and its placing – be seen as an important aspect of the company's communication strategy. Rather than simply 'demand' a user's date of birth, the company offers an explanation – in humorous, informal, second person address – for doing so. The pop-up thus aligns with the construction of a coherent and appealing brand identity which ultimately contributes to the customers buying the company's products.

After having identified and briefly described some of the salient features of the language used on the *innocent* website, the next section explores some of these observations in more detail with the aim of understanding how some of these discursive strategies and linguistic features may contribute to portraying *innocent* in a particular and idiosyncratic way thereby constructing its unique and distinctive, as well as authentic and legitimate identity.

## *4.2 In-depth analysis of selected aspects of the website*

Based on previous research (e.g. Pollach, 2005) we have decided to focus in the qualitative analysis on the 'us' section of the website, where *innocent*, like other companies, explicitly describe who they are, how the company was founded and how it has developed since. These so-called 'about us' sections, provide important sites for identity construction as they are generally used – often very strategically – by the company to construct a particular self-image which they convey to their (potential and actual) customers and which is closely related to their brand. We first describe some of the discursive processes through which *innocent* make specific identity claims thereby attempting to construct the company's identity – largely through drawing on the processes of indexicality and relationality described above – before taking a more critical look at how these processes and related identity claims are sometimes complementary and sometimes contradictory.

### *4.2.1 'hello we're innocent' – making identity claims*

The 'us' section of *innocent* is a good example of the interplay between linguistic, visual, spatial and time-based elements described in Pauwels (2012) which constructs

meanings (and identities). The 'us' page comprises two sections: one entitled '*hello, we're innocent*' (all in lower case), and the other called '*the innocent timeline*' which consists of a combination of pictures accompanied by short texts capturing the company's development since 1998 up to 2014. Combining images and text, this webpage caters for at least two potential user groups: those who look for specific information and those who 'stumble' across the page while looking for something else. In both cases, well-constructed 'about us' pages, which strike a balance between providing quality information (while at the same time avoiding self-congratulatory statements) and using images and relevant weblinks can contribute to image formation and hence identity construction (Pollach, 2005: 298). The company consistently uses these pages to index belonging to an imagined 'ethical' company category, which, according to MCA, projects certain positive attributes and behaviours often associated with such a company onto *innocent*. At the same time the website user is constructed as someone who follows a 'healthy' lifestyle (defined by the consumption of *innocent* products amongst others). This is achieved through the imagery evoked on the various webpages, and drawing on the different processes captured in the relationality and indexicality principles described above. In the 'us' section *innocent* portrays itself as a fun company that is driven by a strong interest in being (and being perceived as being) an ethical, sustainable, and health-conscious brand.

A good example of an attempt to construct these identities can be found in the timeline entry for 2014 entitled '*tastes good, does good*' (in lower case), which is accompanied by a photograph of what appears to be a farmer (who, with a big smile on his face, is holding a papaya fruit towards the camera; see Picture 1).



**Picture1.** '*tastes good, does good*' (<http://www.innocentdrinks.co.uk/us/our-story>)

Here, identity construction takes place via setting the company in close relation to the workers it employs (as per relationality principle). By positioning themselves in relation to these others (who are here represented by the smiling man in the picture), the stances of responsibility and ethicality are evoked, which, following the indexicality principle, contribute to constructing the company's identity as 'a good employer' who cares about its workers. These identity claims are also supported by the frequent use of primary colours on the website, which carry a strong link to purity (van Leeuwen 2011), which is used here to help creating the company's identity as ethical. These trends are

further complemented by the following text which accompanies the picture of the farmer, described above, on the right hand side:

*'After a lengthy brainstorm we conclude that the things we make taste good, do you good, and do others good too. And thus 'Tastes Good, Does Good' is born.'*

This strategic combination of headline, picture, colours and text, and the content which makes reference to *innocent's* foundation (*'Tastes Good, Does Good'*) thus work together and complement each other to create this 'ethical' and 'caring' image of the company. This professional identity of the company is complemented by another identity that the company consistently constructs for itself on its website, and that is closely related to the stances of playfulness, friendliness and fun.

This other, perhaps more social, identity is constructed through various modes and channels, including colour (e.g. the use of the colours of the rainbow for the years on the timeline, and the use of mainly saturated colours to indicate "positive, exuberant, and adventurous" values (van Leeuwen 2011)), fonts (e.g. the exclusive use of lower case in the headings but, interestingly, not in the texts underneath), and text (e.g. the overall very informal and often amusing and humorous interactional style and tone of the text on the website). Of particular interest is the overall humorous style which is characteristic of the company's websites and which is not only constructed verbally but

also visually and interactively. Verbally, for example, the choice of words is often less serious than could be expected for a corporate website – especially when delivering information, such as in the timeline where the company’s beginnings are described (*‘so we resigned from our jobs the next day and got cracking’*), as well as in the various headlines used to describe memorable moments in the company’s history (e.g. *‘world domination continues’*, *‘our quest for more natural, delicious, healthy foods that help people live well and die old’*), the regular use of unusual lexical items (such as *nosy’*, *‘hoodie’* and *‘fete’* as identified in the corpus analysis above<sup>3</sup>), and the frequent play with words which is of course central to the name of the company itself and which is also reflected in the heading on the ‘us’ webpage: *‘hello, we’re innocent’*. Visually, this humorous tone is accomplished, for example, by evoking the stances of playfulness, friendliness and fun in the various pictures in the timeline by surrounding them with what appear to be hand-painted frames. Interactively, this humorous style is achieved by providing links to numerous social media (e.g. the blogs) and the recurring integration

---

<sup>3</sup> The unusual use of ‘hoodie’ is related to one of the competitions that *innocent* regularly organise for their consumers and that contribute to the theme of community building in a self-deprecating way as described above. The same function is performed by the lexical item ‘fete’ which is frequently used to describe an event they held in Regents Park. Using the term ‘fete’ rather than ‘party’ or ‘event’ emulates the traditional ‘village fete’ of English rural life and evokes associations of community and family thereby further adding to the company’s attempts to set up these identity categories.

of the pop-up mentioned in the corpus analysis, as well as an invitation to '*sign up for love, friendship, a weekly newsletter*' (at the bottom of almost every page)). All these features contribute to constructing the company's image as someone who does not take themselves too seriously thereby enhancing both the image of the company as well as its products (e.g. Guillory and Sundar, 2014; Sundar and Kim, 2005).

Through a combination of various discursive strategies, then, including informality, use of humour and play with words, the company evokes the stances of being likeable, friendly, playful, fun, non-serious, and by mobilising connotations associated with children/youth, these identity claims are substantiated. All these adjectives, and the stances they evoke, are also often associated with a friend, someone fun to be with. This finding is perhaps not surprising as it ties in nicely with the observations made in our corpus analysis where we identified '*friendship*' as one of the unusually frequently used keywords on the *innocent* website. These processes are further supported by claiming membership in specific kinds of categories (such as that of an 'ethical' and 'fun/friendly' company) which enables the company to benefit from the positive attributes and behaviours associated with them. Through these processes, then, *innocent* constructs an identity which appears to be primarily social rather than professional. However, these two identities are not contradictory but are inter-related and interwoven in complex ways. Thus, rather than claiming to be either one or the other, *innocent* skilfully

combines them into a set of identities which ultimately contribute to creating the company's highly successful brand image as ethical, healthy and professional, yet friendly, humorous and playful.

In the next section we focus on a different aspect of identity construction and look in more detail at those instances where the company's identity claims are not simply accepted by the website users but where they are questioned and challenged – for example (drawing on Bucholtz & Hall 2005) for not being authentic and legitimate. We explore some of the strategies that *innocent* employs to deal with and respond to these challenges thereby legitimating their identity claims.

#### *4.2.2 Dealing with challenges to identity claims from customers*

A good example to illustrate some of the ways in which *innocent* deals with and responds to these challenges to their identity claims is the representation of the controversial take-over from *Coca-Cola* on the timeline (2009) on the 'us' section. One of the strategies that *innocent* successfully uses in this context is humour, or, more specifically, self-denigrating humour and self-irony. These kinds of humour are renowned for assisting the user in coping with a difficult situation and protecting them from criticism by others (e.g. Hay, 2001), as well as in admitting own mistakes (Zajdman, 1995) thereby

creating a positive self-image and portraying oneself as being in control of the situation (Campbell, 2000).

Under the heading *'hello coca cola'* the take-over by *Coca-Cola* is described as follows:

*'Coca-Cola invests in innocent. A nation rejoices/sends an angry letter/doesn't really notice'*.

While a lot could be said about the humorous wording of this text with its multiple contrasting meanings and interpretations, what is particularly noteworthy about this item on the company's timeline is the image accompanying this text. The picture shows a coke can with a juice pack squeezed out by the coke can's wide mouth with sharp teeth (see Picture 2).



**Picture 2.** *'hello coca cola'* (<http://www.innocentdrinks.co.uk/us/our-story>)

While the coke can (and by implication the *Coca-Cola* company it represents) is portrayed as a predator gulping down its prey, the *innocent* juice carton is depicted as the prey, the victim and as, quite literally, 'innocent'. This impression is further reinforced by what appear to be red dots (of blood?) on the white floor around the violent coke can. Through this visual representation, *innocent's* identity is thus constructed by positioning the company (as embodied by a smoothie carton) in relation to another company, *Coca-Cola* (represented by a coke can). Drawing on MCA, this relationship between the two companies is presented here as the standardised relational pair (Sacks 1992) of culprit and victim, which actively and, one could argue, strategically, contributes to identity construction by portraying *innocent* as the ('good', 'innocent', and perhaps even helpless) victim in relation to the ('nasty' and 'dangerous') *Coca-Cola*. *Innocent* thereby skilfully rebukes the criticisms of its customers (see, e.g. the comments at the beginning of the paper) and, using Bucholtz and Hall's terminology, tries to counter the customers' perception of the company's identity claims as artificial and fake, and perhaps even as delegitimate. Through these strategies *innocent* at the same time reinforces its own identity claims as 'good', 'ethical' and 'innocent'. Similar strategic self-representations were also observed in a study by Buttny (2009) who analysed how *Wal-Mart* was represented to the community during a public hearing in light of the community's criticism of the company's plans. Like *innocent*, *Wal-Mart's*

representatives tried to place their company in a positive light by rebuking their customers' points of criticism. However, unlike the juice company, they did not use humour to help them achieve this but rather drew on the discourse of facts and information to support their own arguments.

So far our analyses have described some of the diverse processes through which *innocent* attempts to construct a set of closely intertwined (professional and social) identities on the company's website. The processes described above and the various stances they evoke and identities they create have been relatively complementary and have contributed to a more or less homogenous, even if multifaceted, image of the company as 'ethical' and yet 'fun'. In the next section we critically re-examine some of these identity claims and take a closer look at the discursive processes through which these claims are being articulated and sometimes undermined by focusing on the choice of pronouns the company uses to portray itself – especially in relation to its customers.

#### *4.2.3. How 'innocent' are we really? Making contradictory identity claims*

Choice of pronouns is another discursive strategy that is often associated with identity construction (e.g. Pollach, 2005; van De Mieroop, 2007). It is, in fact, one of the strategies through which relationship claims are uttered and thereby constructed between the company and its website users (Fuertes-Olivera et al., 2001; Pollach, 2005;

Serrano and Aijón Oliva, 2013). In this section we focus on the use of pronouns in relation to *innocent's* claims of a 'family' identity.

As we have elaborated in our quantitative analysis above in more detail, pronouns feature prominently on the *innocent* webpages. As shown in Table 1, the most frequently used pronouns on the website are 'you' and 'I' followed by 'we' and the corresponding possessive pronoun 'our'. According to Callow (1998), first person pronouns, like 'we', 'us' and 'our' are particularly effective in establishing a relationship between the author of a text and its readers, while second person pronouns, such as 'you' and 'your', are characteristic for mass communication in spite of the fact that the authors clearly do not know their audience. These generic uses of second person pronouns "draw the audiences into the discourse, thereby remedying the impersonality of mass communication (Fairclough, 1989)" (Pollach, 2005: 296).

The first person plural pronoun 'we' is particularly complex as it can be used as an audience inclusive, audience exclusive or impersonal pronoun (e.g. Serrano and Aijón Oliva, 2013), and its reference can shift, sometimes within the same sentence. The ways it is used and the meanings it conveys on the 'us' webpage are particularly interesting as they challenge some of the identity claims the company makes on this page.

Under the link 'family' the main heading on this webpage reads '*welcome to the innocent family*', which together with a framed black and white photograph of what appears to be a family sounds like an invitation to customers to become a member of the *innocent* family (see Picture 3).



**Picture 3.** 'family' (<https://www.innocentdrinks.co.uk/us/family>)

Through the image and the text on this webpage and especially by explicitly mentioning the identity category '*family*', *innocent* portrays itself as a collective (rather than an individual) which, according to MCA, is welcoming and inclusive (as indexed by '*welcome to*' and '*join*') and which carries all the positive attributes and behaviours associated with a family. This reference to family is in line with the frequent use of '*friendship*' throughout the websites as described above, thus making further identity claims for the company by evoking stances that index a collective, supportive, friendly and inclusive identity for *innocent*.

However, upon closer reading of the text below this inviting headline next to the 'family' photograph, a rather different picture emerges. Of all the pronouns used in this text, half refer to the collective identity of *'the family'* while the other half refer to what appears to be a singular 'you' – i.e. the user of the website. Particularly noteworthy in this context are the ways in which the 'we' is often set in opposition to the 'you' – thereby creating distinct subject positions rather than an inviting and welcoming family. A good example of this is the opening sentence: *'We were wondering if you'd like to join the innocent family'* in which the 'you' is clearly put in opposition to the 'we'. This 'us' versus 'you' dichotomy (Van Dijk, 2001) is maintained throughout the entire paragraph and at no point do the 'we' and the 'you' merge. This is further shown in the promise that once 'you' have joined *'our family'* *'[w]e'll email you our news once a week'*. It is thus very clear that the 'we', although it appears inclusive, is, in fact exclusive, and means 'us at *innocent'*, while the 'you' refers to *'the people who drink our drinks, i.e. you'* as the website explicitly states. This setting up of rather distinct subject positions by explicit mentioning specific (and in this case exclusive) identity categories which are set in opposition to each other (i.e. *'the innocent family'* versus *'the consumers'*), is rather surprising and appears to undermine and challenge the company's attempts to portray itself as an inclusive and welcoming 'family'.

These observations on the ‘family’ webpage are further supported by a quantitative approach to the referents of ‘we’ across the *innocent* website. Looking at concordance lines, as was noted in Table 1 above, there are over 20,000 instances of ‘we’ on the websites, making it unfeasible to look qualitatively at each instance. But a qualitative analysis of a random sample of 250 concordance lines generated only *one* instance of ‘we’ in its inclusive use – thereby providing further support for our observations that this pronoun is predominantly used in its exclusive meaning, which, in turn, challenges some of the company’s identity claims as being an open and welcoming ‘family’.

## **5. Discussion and conclusion**

Through a strategic combination of tools and techniques from corpus analytical and discourse analytical approaches, and using Bucholtz and Hall’s (2005) framework and MCA as a guiding principle, we have identified and described several processes through which identity claims are articulated (and sometimes justified and reinforced, as well as strategically played down and contradicted) on the corporate website of *innocent*. This identity work takes place in different modes and channels and involves discursive, visual, and interactive processes. Through these various modes and largely drawing on the various processes captured by the relationality and indexicality principles (Bucholtz &

Hall 2005) and MCA, the company makes specific identity claims for itself and creates a set of interrelated and intertwined professional and social identities. This set of identities enables the company to strategically foreground and exploit the different functionalities and meanings associated with each identity, such as professionalism, ethicality and responsibility, or friendliness and informality, which is particularly useful when dealing with customers' criticism and attempting to uphold specific identity claims, and to construct an authentic and legitimate brand image for the company.

Like Koller (2007) in her study of *HSBC* we have also observed that one of the strategies through which this identity construction takes place is the creation of a dichotomy which enables the company to draw on and utilise competing discourses thereby setting up specific, and sometimes opposing, subject positions for itself and its customers. For example, through their use of certain images (e.g. the black and white picture of what appears to be a family in Picture 3), choice of pronouns (e.g. 'we') together with the explicit mentioning of specific identity categories (e.g. '*family*') *innocent* indexes the stances of collectivity and togetherness, as well as friendliness thereby claiming membership in this category and benefiting from the various positive attributes associated with it. Through these processes *innocent* portrays itself as an open, inclusive and inviting family (who asks its website users to '*join us*'). However, these claims for a 'family' identity are at the same time relativised and perhaps undermined to a certain

extent by using the pronoun 'we' predominantly in its exclusive (rather than inclusive) meaning thereby creating an 'us' versus 'you' dichotomy in which '*the innocent family*' and '*the consumers*' are positioned at opposing sides.

But rather than interpreting these contradictory practices negatively and assuming that they are inconsistencies which render the identity claims of *innocent* on its website fake or artificial (Bucholtz & Hall 2005), we would argue that they are a reflection of the recipient design of websites which needs to consider and accommodate various different potential audiences (e.g. consumers, but also stakeholders and staff) who may all visit this website and expect to feel addressed appropriately in their needs and expectations. Using the identity category 'family' could thus be seen as a strategic move which enables *innocent* to achieve this by drawing on different kinds of family at different points, corresponding for example to the nuclear family (including staff), the extended family (including stakeholders) and the metaphorical family (including consumers). In order to achieve this, the referential vagueness of the we-form as well as the generic you-form are strategically drawn upon to design the website for the different recipient groups. This choice of pronouns together with the use of the category 'family' thus allows for different degrees of proximity to the 'family core', making every potential website user feel 'at home'.

In addition to generating these insights into the complexity of corporate identity construction online, this paper also makes several methodological contributions. Not only do our analyses provide further support for the benefits of combining quantitative and qualitative procedures when researching corporate or professional identities (see also van de Mierop 2007; Koller 2007), but we have also discussed possible ways of dealing with the methodological challenges of conducting research on websites. And while we do not want to claim too much here, we hope that future research on (identity construction in) online environments which deals with relatively large quantities of data, may find some of the procedures that we have outlined here useful and may want to follow (or challenge) some of the procedural decisions that we had to take. A particularly interesting avenue for future research could be to take a more diachronic approach and examine possible changes in online identity construction over time – especially given the rapidly changing nature of this environment.

Clearly, identity construction is a complex yet fascinating topic of inquiry which takes place beyond the borders of the written text and is accomplished creatively and conjointly across the different modes and channels. All these modes and channels constitute important sites for meaning making in identity creation – and more emphasis should be put on systematically incorporating visual and interactive elements in

quantitative and qualitative analyses of identity construction – both methodologically as well as theoretically and conceptually.

### **Acknowledgments**

We are very grateful to *innocent* for giving us permission to reproduce the pictures from their website in this paper, and to Ben Heuer for his generous support in programming the webcrawler to harvest the data from the *innocent* websites. We would also like to thank *The Centre for Applied Linguistics* at Warwick University for providing financial support to finalise this paper.

### **References**

- Anderson E (2014) Richard Reed: From innocent smoothies to changing the world. *The Telegraph*, 29 October. Available at: <http://www.telegraph.co.uk/finance/businessclub/11154017/Richard-Reed-from-Innocent-Smoothies-to-changing-the-world.html> (accessed 07 July 2015)
- Angouri J (2015/fc) Identity. In Zhu Hua (ed.) *Research Methods in Intercultural Communication*. Blackwell.
- Angouri J and Marra M (2011) *Constructing Identities at Work*. Houndmills: Palgrave Macmillan.
- Baker P and Levon E (2015) Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication* 9: 221-236.
- Balmer JM and Soenen GB (1999) The acid test of corporate identity management™. *Journal of Marketing Management* 15: 69-92.
- Balmer JM (2001) Corporate identity, corporate branding and corporate marketing - Seeing through the fog. *European Journal of Marketing* 35: 248-291.
- Bamberg M, De Fina A and Schiffrin D (2006) *Discourse and Identity*. Cambridge: Cambridge University Press.
- Bamberg M, De Fina A and Schiffrin D (2011) Discourse and identity construction. In: Schwartz SJ, Luyckx K and Vignoles VL (eds) *Handbook of Identity Theory and Research*. Berlin: Springer, 177-199.

- Baroni M and Ueyama M (2006) Building general-and special-purpose corpora by web crawling. *13th NIJL international symposium, language corpora: Their compilation and application*, pp. 31-40.
- Benwell B and Stokoe E (2006) *Discourse and Identity*. Edinburgh: Edinburgh University Press.
- Bravo R, de Chernatony L, Matute J, et al. (2013) Projecting banks' identities through corporate websites: A comparative analysis of Spain and the United Kingdom. *Journal of Brand Management* 20: 533-557.
- Bucholtz M and Hall K (2005) Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7: 585-614.
- Butler J (1990) *Gender Trouble: Feminism and the Subversion of Gender*. New York: Routledge.
- Buttny R (2009) Wal-Mart's presentation to the community: discursive practices in mitigating risk, limiting public discussion, and developing a relationship. *Discourse & Communication* 3: 235-254.
- Callow K (1998) *Man and Message: A Guide to meaning-based Text Analysis*. Boston, MA: University Press of America.
- Campbell K (2000) *'Chuckle while you work': Subordinates' perceptions of superiors' humour orientation, assertiveness, responsiveness, approach/avoidance strategies, and satisfaction*. Unpublished MA Dissertation, West Virginia University, USA.
- Cheney G and Christensen TL (1999) Identity at issue: linkages between 'internal' and 'external' organisational communication. In: Jablin FM and Putnam LL (eds) *New Handbook of Organisational Communication*. Newbury Park, CA: Sage Publications, pp. 231-269.
- De Fina A (2010) Discourse and identity. In: Van Dijk TA (ed) *Discourse Studies. A multidisciplinary Introduction*. London: Sage, pp. 263-282.
- De Fina A (2010) The negotiating of identities. In: M Locher and S Graham (eds) *Handbook of Interpersonal Pragmatics*. Berlin: deGruyter, pp. 205-224.
- Fairclough N (1989) *Language and Power*, London: Pearson.
- Ferraresi A, Zanchetta E, Baroni M, et al. (2008) Introducing and evaluating ukWaC, a very large web-derived corpus of English. *4th Web as Corpus Workshop (WAC-4) Can we beat Google*, pp. 47-54.
- Fuertes-Olivera PA, Velasco-Sacristán M, Arribas-Baño A, et al. (2001) Persuasion and advertising English: Metadiscourse in slogans and headlines. *Journal of Pragmatics* 33: 1291-1307.
- Fuoli M (2012) Assessing social responsibility: A quantitative analysis of appraisal in BP's and IKEA's social reports. *Discourse & Communication* 6: 55-81.

- Guillory JE and Sundar SS (2014) How does website interactivity affect our perceptions of an organization? *Journal of Public Relations Research* 26: 44-61.
- Halliburton C and Ziegfeld A (2009) How do major European companies communicate their corporate identity across countries? An empirical investigation of corporate internet communications. *Journal of Marketing Management* 25: 909-925.
- Hay J (2001) The pragmatics of humor support. *International Journal of Humor Research* 14(1), 55–82.
- Holmes J, Stubbe M and Vine B (1999) Constructing professional identity: “Doing power” in policy units. In: Sarango S and Roberts C (eds) *Talk, Work and Institutional Order: Discourse in Medical, Mediation and Management Settings*. Berlin and New York: Mouton de Gruyter, pp. 351-385.
- Innocent Foundation (n.d.) How we began. Available at: <http://www.innocentfoundation.org/about-us/how-we-began> (accessed 07 July 2015).
- Johnston J, Szabo M and Rodney A (2011) Good food, good people: Understanding the cultural repertoire of ethical eating. *Journal of Consumer Culture* 11: 293-318.
- Jones RH (2014) The multimodal dimension of claims in food packaging. *Multimodal Communication* 3: 1-11.
- Kilgarriff A, Pavel R, Pavel S, et al. (2004) The Sketch Engine. Conference Proceedings of EURALEX, Lorient, France.
- Knox JS (2009) Punctuating the home page: Image as language in an online newspaper. *Discourse & Communication* 3: 145-172.
- Koller V (2007) “The World's Local Bank”: Glocalisation as a strategy in corporate branding discourse. *Social Semiotics* 17: 111-131.
- Kress GR and Van Leeuwen T (1996) *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Liu D (2012) The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes* 31: 25-35.
- Locher M (2008) Relational work, politeness, and identity construction. In: Antos G and Ventola E (eds) in cooperation with Weber T *Handbook of Interpersonal Communication*. Berlin: De Gruyter, pp. 509–540.
- McEntee-Atalianis LJ (2013) Stance and metaphor: Mapping changing representations of (organizational) identity. *Discourse & Communication* 7: 319-340.
- Myers G (1994) *Words in ads*. London: Arnold.
- Ochs E (1992) Indexing gender. In: Duranti A and Goodwin C (eds) *Rethinking Context. Language as an Interactive Phenomena*. Cambridge: Cambridge University Press, pp. 335-358.

- Ochs E (1993) Constructing social identity: A language socialization perspective. *Research on Language and Social Interaction* 26: 287-306.
- Park S-e, Choi D and Kim J (2005) Visualizing e-brand personality: Exploratory studies on visual attributes and e-brand personalities in Korea. *International Journal of Human-Computer Interaction* 19: 7-34.
- Pauwels L (2012) A multimodal framework for analyzing websites as cultural expressions. *Journal of Computer-Mediated Communication* 17: 247-265.
- Pollach I (2005) Corporate self-presentation on the WWW: Strategies for enhancing usability, credibility and utility. *Corporate Communications: An International Journal* 10: 285-301.
- Sacks, H (1992) *Lectures of Conversation*, vol. 1, G Jefferson (eds). Oxford: Blackwell.
- Schnurr S (2013) *Exploring Professional Communication: Language in Action*. Oxon: Routledge.
- Scollon R and Levine P (2004) Multimodal discourse analysis as the confluence of discourse and technology. In: Scollon R and Levine P (eds) *Discourse and Technology: Multimodal Discourse Analysis*. Washington DC: Georgetown University Press, pp. 1-6.
- Scollon R and Scollon SW (2003) *Discourses in Place: Language in the Material World*. London: Routledge.
- Serrano MJ and Aijón Oliva MÁ (2013) Seguimos con la actualidad... The first-person plural nosotros 'we' across Spanish media genres. *Discourse & Communication* 7: 409-433.
- Simmons GJ (2007) "i-Branding": Developing the internet as a branding tool. *Marketing Intelligence & Planning* 25: 544-562.
- Simpson-Vlach R and Ellis NC (2010) An academic formulas list: New methods in phraseology research. *Applied Linguistics* 31(4): 487-512.
- Starr MA (2009) The social economics of ethical consumption: Theoretical considerations and empirical evidence. *The Journal of Socio-Economics* 38: 916-925.
- Suchomel V and Pomikálek J (2012) Efficient web crawling for large text corpora. *Proceedings of the seventh Web as Corpus Workshop (WAC7)*: 39-43.
- Sundar SS and Kim J (2005) Interactivity and persuasion: Influencing attitudes with information and involvement. *Journal of Interactive Advertising* 5: 5-18.
- Sweney M (2007) Has Innocent smoothies sold its soul by hooking up with McDonald's? *The Guardian*, 02 May. Available at: <http://www.theguardian.com/media/organgrinder/2007/may/02/hasinnocentsmoothiessoldit> (accessed 07 July 2015).

- Van De Mieroop D (2007) The complementarity of two identities and two approaches: Quantitative and qualitative analysis of institutional and professional identity. *Journal of Pragmatics* 39: 1120-1142.
- Van De Mieroop D and Clifton J (2012) The interplay between professional identities and age, gender and ethnicity: Introduction. *Pragmatics* 22: 193-201.
- Van de Mieroop D and Schnurr S (fc) *Identity Struggles. Evidence from Workplaces Around the World*. Amsterdam: Benjamins.
- Van Dijk TA (2001) Critical Discourse Analysis. In: Schiffrin D, Tannen D and Hamilton H (eds) *The Handbook of Discourse Analysis*. Malden, MA: Blackwell, pp. 352-371.
- Van Leeuwen T (2011) *The Language of Colour: An Introduction*. London: Routledge.
- Widdicombe S (1998) 'But you don't Class Yourself': The interactional management of category membership and non-membership. In: Antaki C and Widdicombe S (eds) *Identities in Talk*. London: Sage, pp. 52-70.
- Zajdman A (1995) Humorous face-threatening acts: Humor as strategy. *Journal of Pragmatics* 23: 325-339.