Patrizia Pierini
(Università di "Roma Tre")

'Fully refurbished end terraced house'.
A Corpus Study of Online Classified Ads

Advertising is a versatile and dynamic genre incorporating a constellation of subgenres, most of which are short messages subject to spatial or temporal restriction. One of them is the classified ad, a form of small, local advertising that, like other forms of written communication, has found its way on to the Web.
The present article, which investigates online classified ads offering properties for sale, is based on a qualitative and quantitative analysis of a small corpus of texts downloaded from UK-based websites. The aim is to explore their main textual, discursive and linguistic features by combining corpus linguistics methodology with discourse analysis and genre analysis approaches. The study first outlines the evolution of the genre from print to the Web; then, it profiles the general features and structure of online classified ads, discusses variation across the texts in the corpus, and identifies the linguistic features by analysing frequencies and collocations of function words, verbs, nouns, adjectives, abbreviations and acronyms. The analysis also sheds some light on the phraseology of classified property advertising.

1. Introduction

In everyday life, we are users, mostly as receivers, of an array of small texts – public notices, text messaging, weather forecasts, horoscopes, back-cover blurbs, book and film reviews, the labels on food products, epigraphs, graffiti and banners. But we also find them in other contexts, such as literary communication (e.g. aphorisms, epigrams) and specialised communication, where we find a number of small texts in any domain: e.g. medicine (prescriptions, case histories, patient information leaflets); law (indictments, verdicts, wills, warranties); business (tenders, job adverts, invoices). These genres, that might be considered as minor in written communication, are not at all marginal. Usually, they do not attract the attention of researchers, even if there are some studies devoted to them, e.g. Suter (1993) on wedding reports, Mansfield (2001) on job adverts and applications, Brutiaux (2005) on promotional catalogues, Gesuato (2007) on back-cover blurbs.

Advertising, so pervasive in contemporary society, is a versatile and dynamic genre incorporating a constellation of subgenres (Bhatia 2002: 57-65), most of which are short messages subject to spatial or temporal restriction. They are similar in terms of their communicative purpose, i.e. persuading an intended audience to purchase goods or services. They are also different in terms of the product promoted, the type and size of the audience targeted, the strategy used, and the medium exploited to disseminate the message. They range from the print advert in newspapers and magazines, junk mail, the handbill, the audio advert on radio, the audiovisual TV commercial, the poster on hoardings to the advert on displays in rail and metro stations. More recently, multimodal forms, such as the promotional/commercial website, the banner ad pasted across web pages and the pop-up window, have emerged on the Web (Janoschka 2004).

Classified ads as a form of small, local advertising have been briefly examined by Bex (1996:158-162) and extensively analysed by Brutiaux (1996), who has explored the linguistic features of print ads in American newspapers. This paper, which investigates online classified ads offering properties for sale, is based on a qualitative and quantitative analysis of a small corpus of texts downloaded from UK-based websites. The aim is to explore the main linguistic, textual and discursive features of online ads by combining corpus linguistics methodology (Tognini Bonelli 2001; Bowker / Pearson 2002) with specialised discourse analysis (Gotti 2005) and genre analysis (Bhatia 2004) approaches. The paper is organised as follows: Section 2 describes corpus and methodology; Section 3 presents an overview of the evolution of the classified ad as genre; Section 4 profiles the general features and structure of online ads; Section 5 discusses variation across the texts in the corpus, while 6 analyses linguistic features; Section 7 contains some concluding remarks.

2. Corpus and methodology

The Web is a multi-billion repository of documents offering an enormous quantity of samples of language use that are fresh, authentic and representative, since any domain, content area, discourse type or genre are present. But we do not know anything about the proportions of the different genres. For example, a search on Google of 'free classified ads UK'

(18.12.2009) produced a list with an indeterminate number of classified websites. Most publish ads on a wide range of products, others on just one product (e.g. www.thehorseexchange.com; www.ukpetmart.com), so the solution is to extract a random sample and analyse it.

We selected four classified websites: one is associated with a traditional British ads paper (www.loot.com); the others are purely digital publications (www.freeads.co.uk; www.ukclassifieds.co.uk; www.adflyer.co.uk). They are both local and national: they publish ads from any part of the UK, each ad targeting its local market. Individuals – and small businesses – can place their ads freely, composing a message where they can also upload photos.

The content of online ads is varied: some offer employment, others seek romantic partners, but most promote the selling of a wide variety of products grouped under categories classifying them. Property ads constitute a relevant subgenre that allow hundreds of thousands of anonymous sellers and buyers to make commercial transactions. Their success is determined by its many advantages: an inexpensive way for sellers to reach a potentially global audience; the timeliness of the message; the ease of access for potential buyers; the capability for online search.

The corpus was compiled from November to December 2009 by downloading 25 ads of flats and houses for sale from each website for a total of 100 texts, and saving them both in MHT (web page archive format) and text format. They were selected randomly, but taking into consideration one aspect: in the real estate market, there are also intermediaries (real estate agencies) that publish their print advertising materials, often publish their own websites, and also take advantage of a further form of marketing by publishing ads in free classified websites. So the corpus contains both private ads and trade ads offering properties in the UK. The corpus was post-edited, by removing the paratextual elements external to the messages. For this reason, it is small, amounting to circa 18,000 running words; but, as pointed out by Bowker / Pearson (2002: 48), even small corpora can be reliable aids when investigating domain-specific discourse.

As for methodology, the investigation used basic corpus linguistics methodologies, namely, word frequency counts, concordancing and the extraction of collocational information. Two aspects were considered: the high frequency of words as well as an infrequency that can be equally meaningful (Tognini Bonelli 2001: 84). The data were analysed with the help of *Wordsmith Tools 5.0* (Scott 2007).

3. The classified ad: the evolution of a genre

The classified ad is a traditional genre which appeared first in newspapers. Prior to 1850 (Leech 1966: 166-67), most British advertising was constituted by small, local adverts on the lines of today classified ads – their readership was small and their senders were small merchants and private individuals. Later, when large-scale consumer advertising developed, they kept on being published, and today, they also appear in magazines and free ads papers.

Jacqmain (1973: 2-3) identifies three general features of print classified ads, when compared to large-scale advertising: transience (messages are short-lived), individuality (messages are produced by private individuals to other individuals) and inexpensiveness (messages are cheap or free). We add another feature, namely, brevity: economy is one of the criteria governing specialised discourse (Sager *et al*. 1980; Gotti 2005: 28-31) and becomes dominant in specialised genres subject to spatial constraints. It can be achieved through information compacting (e.g. nominalisation, premodification) and conciseness (e.g. compounding, acronymy) (Gotti 2005: 69-81, 40-41). We shall now discuss which linguistic devices are employed in print and later in online classified ads.

Employing the parameters of genre analysis (Swales 1990:58 fll), print classified ads can be characterised as follows: the senders of messages are mostly private individuals, their communicative purpose is to persuade receivers to purchase the product presented and the intended audience is a restricted local market segment. Hence, the setting of the communicative event is a limited geographical area shared by sender and receiver, the interaction between sellers and potential buyers is (usually) direct because there are no intermediaries, and the products promoted are single items.

Print classified ads are small texts typically consisting of verbal text only and just a few column lines in length. They open with the name of the product, followed by a description of its primary characteristics, rarely offering an enticing presentation. In his analysis of four categories of ads (second-hand cars, flats for rent, job offers, personals), Brutiaux (1996) sees them as a case of 'economy register' – as the one observed in telegrams – characterised by two linguistic features. The first is a low degree of syntactic complexity, realised through parataxis, ellipsis and a limited use of function words – articles, auxiliaries, modals, negatives,

personal pronouns and conjunctions. The second feature concerns lexis, which displays prefabricated sequences of words and a set of abbreviations and acronyms to save space and money, when the ad involves payment. These features combine to render the classified ad a highly conventionalised genre. Consider the ad of a flat for rent in (1), taken from a free ads paper, exemplifying the language of these messages:

> (1) WALSALL 2 bedroom furnished flat, close to town centre and schools, DG, CH, spacious rooms, no pets, no DSS, £400 pcm [phone number]

In particular, we note premodification, juxtaposition (a sequence of unlinked syntactic groups separated by commas), and some acronyms, which are genre markers and may be cryptic to receivers not familiar with this type of text: *DG* is for *double glazed*, *CH* for *central heating* and *pcm* for *per current month*. *DSS* (Department of Social Security) is used to refer to the housing benefit provided by the British government to people who need financial help for rented accommodation. The text is an example of a very short print ad, where persuasion seems to be expected to flow from the description, rather than from any explicitly positive evaluation of the item.

To sum up, the brevity of classified ads, achieved by means of syntactic and morpho-lexical devices, reflects the discursive practices of commercial transactions on one hand, and the more intimate kinds of exchange between interlocutors with common interests (selling/buying), on the other. Their interpretation requires some specialist background knowledge.

Like other forms of written communication, the classified ad has found its way on to the Internet. The rise of Web texts has given rise to research that has highlighted issues in the identification, analysis and classification of Web genres (e.g. Kwasnik / Crowston 2005; Santini 2006; Garzone 2007; Giltrow / Stein 2009). Provided that any genre placed in a computer-mediated environment is somehow affected, we identify three categories of genres. The first includes existing genres reproduced on the Web with little adaptation, such as the research article in the electronic edition of print journals or in e-journals. The second category includes traditional genres that have been re-mediated (Bolter / Grusin 2000: 44-50, 196-210): they have been adapted to the new medium and enriched by exploiting

the functionalities of hypertexts – hyperlinking, multimediality, interactivity, searchability – such as the electronic newspaper (Lewis 2003). The third category includes novel genres: the rise of the macrogenre 'website' has produced a number of new genres (e.g. the blog, the e-commerce site, the wiki, the search engine site, the portal) and microgenres (e.g. the homepage, the FAQ, the About us). As we shall see in the next section, the online classified ad has developed from the print ad reshaped for the new medium.

4. *General features and structure of online classified property ads*

The online classified ad belongs to the class of Web genres whose analysis calls for a theoretical framework that integrates the traditional approaches of discourse and genre analysis with tools capable of accounting for the particular features of Web texts. Garzone (2007) suggests adopting a semiotically stratified approach taking into consideration basic factors such as communicative purpose, content, context, structure and style, plus the notion of 'medium' (the Web) that adds unique properties to texts and also determines how they are used by receivers. Within this framework, we define a Web genre as a class of communicative events characterised in terms of sender, content, communicative purpose, intended audience, a set of linguistic and non-linguistic features, plus the functionalities afforded by the Web.

The classified property ad can be considered a specialised genre realising the discourse of business. It talks about a specific product (property); it fulfils a specific purpose (to persuade the receiver to purchase a property); it is addressed to an intended audience constituted by people who are already inclined to purchase a property in a given location. Specialised texts are, by definition, produced by senders having a wider knowledge of the topic than receivers (see the 'knowledge condition' in Sager *et al*. 1980: 52). In the case in point, estate agents are experts in properties, while private individuals can be seen as 'experts' on their own property. The context in which the ad appears is the highly competitive marketing situation on the Web.

Online ads retain the communicative purpose and content of print ads, exhibiting old and new features. Old features are transience,

inexpensiveness and brevity, even if they are longer than print ads because of the larger space available on the Web (see section 5). Individuality becomes from a constitutive to an accessory feature because of the strong presence of real estate agencies. New features are a high degree of interactivity and multimodality. The interactivity with the computer allows users to search for the type of residence they need (flat/house) by means of search boxes and menus, choosing parameters such as area, price range and number of bedrooms, and to move within the site by means of links. The interactivity through the computer allows users to contact the seller by email. Exploiting a plurality of semiotic resources (words, images, graphology), online ads are multimodal texts resulting from the particular combination and interplay of verbal and visual elements (Kress / van Leeuwen 2001: 20).
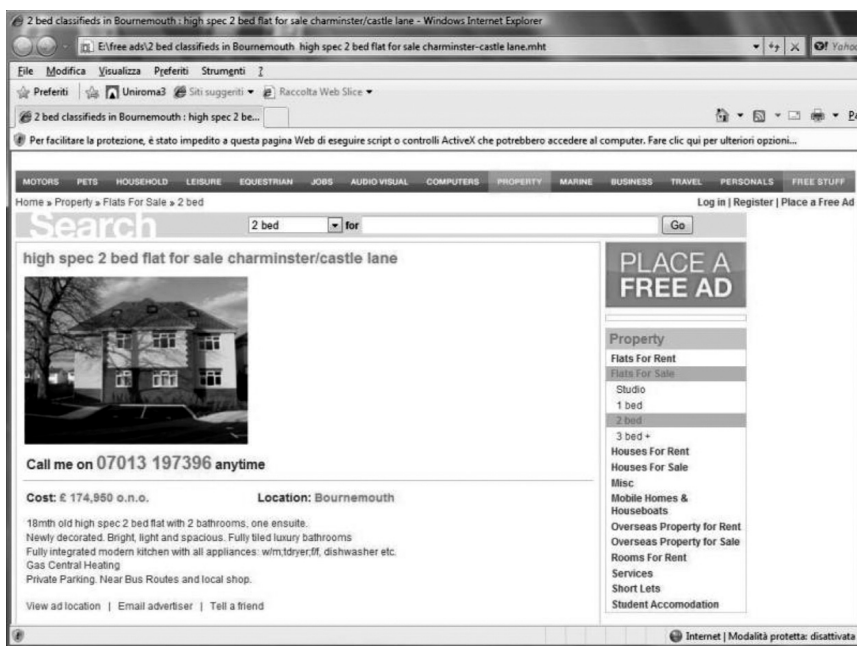

Figure 1. *Screenshot of private ad from www.freeads.co.uk (27.11.2009)*

Each ad is presented in one web page, within a semiotic space visually separated from the other parts of the page, and placed on the left side

(Figure 1) or in the centre of the page (Figure 2). Within the message, the property offered for sale is thematised both verbally and visually. It is thematised by being announced in the headline where it is given prominence by using graphological resources such as bold type, colour or capital letters, and by adding photos of the property (Figures 1 and 2). The graphological resources exploited serve various functions: the white background provides cohesive unity; empty spaces serve to distinguish one structural component from the others; colour serves to highlight, by contrast, portions of text otherwise written in black, such as links.
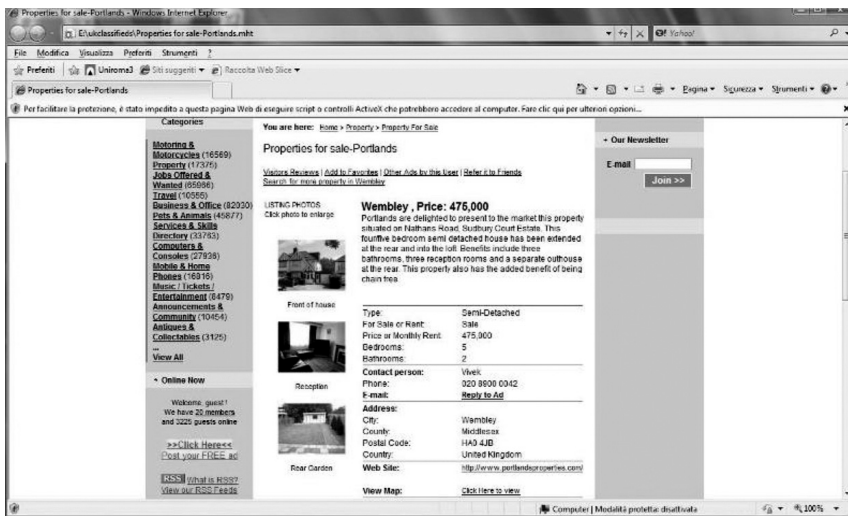

Figure 2. *Screenshot of trade ad from www.ukclassifieds.co.uk (14.12.2009)*

Instead of starting from moves, as in traditional genre analysis (Bhatia 2004: 63-65), we start from the textual structure of ads to go to their rhetorical patterning: they are characterised by a verbal-visual rhetoric in which a move can be realised by more than one element, also of different semiotic nature, and one element can realise more than one move. Below, the seven structural components identified, and their rhetorical functions:
– headline: it is an attention-seeking device (Goddard 2002: 10-19) that presents essential information about the property by exploiting linguistic and graphological resources, in particular colour, large font size, bold. Headlines mention the number of bedrooms (e.g. *2 bed*

ground floor flat), location (*Bushey Hill Road Camberwell SE5*), or both (*3 bed cottage Wimborne, Dorset)*. Very rare is the presence of evaluative words (e.g. *immaculate 2 bedroom flat*). This component exhibits the typical nominal style, with complex noun phrases, premodification and juxtaposition.

– phone number, to contact the advertiser. Its function is to solicit response, just like *ad hoc* links (see below).

– price. This element, missing in standard adverts, characterises the genre. It is clearly stated by the advertiser in order to restrict the vast audience of potential purchasers and target a specific market segment. This and the following three components constitute the four steps of the move 'detailing the product'.

– location: name of town/city, optionally road and postal code.

– description of the property: a verbal presentation of its features and parts.

– image(s): one or more colour photos showing the exterior and/or the interior of the property. Their function is multiple: to provide a visual presentation of the property, to highlight details of the property not described verbally and to attract the receiver's attention. Optionally, they are accompanied by the plan of the property, and a map of the area where it is located.

– (a variable number of) links on the bottom and/or the top of the message, exhorting the receiver to interact with and through the computer: there are links to contact the advertiser, hence to solicit response (e.g. *Email seller, Reply to ad*), to access other information (e.g. *View ad location*; *Visitors reviews*), to do various actions (e.g. *Tell a friend*, *Save ad, Add to favourites, Search for more property in […]*). Their feature is the use of the imperative having the pragmatic value of instruction and/or suggestion. Many trade ads also provide a link to the website of the estate agent to promote their commercial activity.

5. *Variation across ads*

There are some obvious differences among the ads in our corpus due to the different format established by each website. But following the established format does not lead to uniform texts within a website. Variation across the 100 ads examined concerns length and style.

Considering the 'description' components, we noted that they range from 16 to 667 words: for example, the description in Figure 1 consists of 46 words, the one in Figure 2 of 106 words. About the 75% of them amount to no more than 200 words (Table 1).

No. of words	No. of descriptions
up to 100	55
101 to 200	21
201 to 300	11
more than 300	13

Table 1. *Length of descriptions out of 100 classifieds*

With regard to promotional catalogues, Brutiaux (2005) observes that, when they deal with 'glamorous' products, descriptions tend to be more detailed, hence longer, and messages show a higher degree of syntactic complexity. Conversely, the descriptions of 'utilitarian' (i.e. not expensive) products tend to be shorter and exhibit syntactic 'minimalism'. By examining the descriptions in our corpus, the differentiating factors at play seem to be the sender and the price of the property. On the whole, private ads are shorter than trade ads; the shortest descriptions are produced by individuals, who reproduce the paratactic style of print ads, as in (2), or add the message an informal and emotive tone, as in (3):

> (2) 1 bedroom top floor conversion apartment, located close to Thornton Health High Street, vacant possession.

> (3) By Owner. Desperate. I'm over extended and can't make house payments. Low Down. Nice looking ranch. 3 bed/ 1 bath Ashford area

But when private ads promote more expensive properties, descriptions are longer than 100 words and show a certain degree of syntactic complexity, as in (4):

> (4) Set within this pleasant cul-de-sac of the detached properties, a tastefully presented three bedroom detached residence providing a

gas central heating system together with sealed unit double glazing, gardens to the front and rear and single garage. The property would be of interest to a variety of potential purchasers and offers accommodation briefly comprising entrance hall, extensive sunny lounge and dining room, kitchen with built-in oven, hob and dishwasher, three generous bedrooms with two having fitted wardrobes and master with en-suite shower room together with a family bathroom/wc. The property is nicely situated within the estate of Branklyn Gardens and we would strongly recommend early viewing.

Such descriptions usually consist of a combination of full finite clauses, elliptical clauses where the subject or the subject/copula cluster is omitted (*Suitable for 1st-time buyers or buy-to-let),* non-finite clauses (*Offering off road parking, 2 reception rooms* [...]), and juxtaposition of complex phrases. Evaluative or positively connoted elements are sometimes used, such as *pleasant, tastefully (presented), sunny (lounge), generous (bedroom), nicely (situated)* in (4).

Most of the longer descriptions are trade ads (not always clearly identifiable): estate agents tend to provide more detailed descriptions of properties. They can be 'technical', as in (5), or exhibit the persuasive language typical of print large scale advertising relying on creativity, the use of evaluative elements and the interpersonal involvement with the receiver, as in (6)[1]:

(5)   Through Lounge
        25' 1" x 15' 4" narrowing to 8'10" (7.65m x 4.67m) Laminate flooring, double radiator, power points, gas fire with surround, double glazed bay window to front, door to bathroom.
        Bathroom
        Tiled floor and walls, chrome towel radiator, semi pedestal wash basin with mixer taps, close coupled wc, panelled bath with mixer taps and shower attachment, cupboard understairs.

(6)   Fancy a bright, spacious and stylish pad within easy reach of the imminent East London Line extension? Look no further! This funky one bed ticks all the boxes with a cracking interior to boot!

---

[1] In cases like these, print material may have been uploaded on the website without any modification.

Now some comments on what has emerged from the data. Writing a classified ad to post in a website is a challenge for the text producer: the aim of the message is to be noticed and selected in a proliferation of similar competing offers, and then to persuade the receiver to purchase the product. The production of long descriptions displaying linguistic elaboration seems to be a strategy used to impress the receiver: on one hand, estate agents are interested in constructing their image as 'professionals'; on the other hand, individuals selling expensive properties are interested in presenting their property as something special, and/or making affluent or status-conscious receivers feel different if they will purchase that property.

But there is another aspect to consider, i.e. communicative efficiency, which is vital for promotional and commercial Web messages. The elements contributing to it can be traced back to two types of factors: a) technical factors, related to the usability of websites, e.g. ease of navigation; b) linguistic-communicative factors, related to the information architecture of websites, i.e. the organisation and labelling of content (Morville / Rosenfeld 2006), and the rhetoric of the Web (Ursini 2005) by which efficiency means to achieve the intended communicative purpose by adopting a style suitable for the new medium.

The classical study of Morkes / Nielsen (1997) has shown that screen reading is slower than print reading, and that users do not actually read, rather, they scan the web page. Since people read in a different way, text producers should write in a different way; hence the need for a style suitable for the Web that favours readability, both perceptively and linguistically. The numerous works devoted to Web style (e.g. Lynch / Horton 2004) have stressed some points: light style, characterised by short and simple sentences; short paragraphs; typographic resources enhancing scanning (e.g. bulleted lists); each web page should not exceed the screenview. As observed by Janoschka (2004: 135), "for reasons of time, space and perception, Web ad messages must be short and to the point".

On the basis of these considerations, several ads exhibit length and style unsuitable for the new medium: some are too short, reproducing the extremely concise style of print ads, such as (2), while more space is available on the Web; others appear to be very long, not readable because of syntactic complexity, such as (4), or technicality, such as (5), or not to the point because exceedingly persuasive in tone, reproducing the style of

print large-scale advertising, such as (6). But in our corpus there is a number of communicatively efficient messages: they are no longer than 100 words, exhibit a light style, providing a rather brief, but sufficiently detailed description of properties, make use of bulleted lists of the parts and features of properties, and avoid an explicitly evaluative language.

6. Linguistic features

In this section, we shall attempt to identify the linguistic features of classified ads by analysing frequencies and collocations of function words, verbs, nouns, adjectives, abbreviations and acronyms.

6.1 Function words

Table 2 shows the most frequent function words; the top eight words are also the top items in the frequency list.

word	frequency	percentage	word	frequency	percentage
and	587	3.20	from	50	0.27
the	525	2.86	as	49	0.27
to	468	2.55	here	48	0.26
a	415	2.26	it	45	0.25
with	293	1.60	through	45	0.25
for	219	1.19	more	44	0.24
of	218	1.19	you	41	0.22
in	205	1.12	there	40	0.22
this	120	0.65	well	35	0.19
on	102	0.56	within	33	0.18
or	102	0.56	off	32	0.17
by	74	0.40	fully	30	0.16
an	62	0.34			

Table 2. *Function words occurring at least 30 times in the corpus*

The data in Table 2 suggest some observations. We note the wide range of prepositions and the high frequency of the conjunction *and* operating

within phrases, as results from an analysis of its concordances. That reveals that syntactic complexity resides at the phrase level, rather than at the clause level. Further evidence is provided by the absence of conjunctions introducing dependent clauses, except for *as*, which functions as conjunction in 35 occurrences out of 49, very frequently occurring in two patterns that do not depend on full clauses, as illustrated below:

> 11ft 3ins (3.60m x 3.44m) Carpet **as** laid 3 double power points
> Street Borough SE1 £210,000.00 **as** stated Tel:

As observed by Bruthiaux (2005), prepositions are the most semantically highly contentful among function words, and usually, they are not left out in the discourse of classified advertising where they explicit essential relationships.

Then, we find the articles (*the, a, an*), commonly omitted in economy registers, and three deictic items (*this, here, there*). The article *the* signals anaphoric reference, but also implies the uniqueness of the property or the elements related to it. Mostly occurring in the patterns *this property/ house/ flat,* we find the 'introductory *this*' (Biber *et al*. 2000:274), used to introduce new information. Only co-occurring with *click, here* is used to refer to the co-text, while *there* very frequently occurs in the existential *there* construction (see 6.2).

One item worth noting is the personal pronoun *you* occurring 41 times, not found in print ads. *You, you'll* (Table 3) and imperatives (see 6.2) are typical forms of direct user addressing that constitutes a general discursive feature in Web advertising (Janoschka 2004: 139-143).

Another interesting group is constituted by the manner adverbs *well,* clearly evaluative and *fully,* taking on a positive connotation, as in the following concordances that show the typical patterns in which they occur:

> For Sale, this **well maintained** and presented Detached Bungalow
> Liverpool, Price: 85,000 **Well appointed** three bedroomed town House
> an exceptionally spacious and **fully refurbished** 5 bedroom, 3 bathroom, 4
> counter and cabinet space. A cool, **fully tiled** shower room with heated

Finally, we would like to say something about modals, infrequent in print ads, but found in our corpus (Table 3).

word	frequency	percentage
you'll	16	0.09
will	15	0.08
can	12	0.07
would	9	0.05
must	5	0.03
could	4	0.02

Table 3. *Modal auxiliaries*

The modals in our corpus have a very low frequency, with *will* and *can* more frequent than the others. This is not surprising since they are the most frequent modals in absolute terms, varying in frequency and values across registers (Biber *et al*. 2000: 483-497). They are the typical modals of print advertising (Leech 1966:125-126) as well as Web advertising: Janoschka (2004: 140-141) has discussed instances of *will* in promotional/commercial websites, and Pierini (in press) has found *will* and *can* in hotel websites with a higher frequency (both with a percentage of 0.24). In our corpus, *will* and *can* express, respectively, the semantic values of 'prediction' and 'possibility', taking on the pragmatic values illustrated in the following concordances:

> storage, loads of natural light and **you'll** enjoy a peaceful, mature location
> to build fourth bedroom upstairs which **will** add value. Externally there is
> main bedrooms. Each two-bedroom flat **can** sleep 5 people. Ideally located in
> area with adjoining pine doors that **can** be left open or closed

Will expresses the value of a promise (line 1), or the infallibility of the claim (line 2). *Can*, which in most cases takes as its subject a noun denoting the property or one of its parts, expresses the idea that the possibilities offered by the property will satisfy the receiver's needs or benefit him/her.

Because of their low frequency, *will* and *can* do not represent a discursive feature of online ads. A possible explanation for their presence is the influence of other forms of Web advertising.

6.2 Verbs

word	frequency	percentage
is	190	1.03
fitted	78	0.42
are	77	0.42
email	70	0.38
detached	66	0.36
click	65	0.35
has	65	0.35
call	56	0.31
glazed	40	0.22
contact	35	0.19
built	34	0.19
presented	31	0.17

Table 4. *Verb forms occurring at least 30 times*

From Table 4 it emerges that there are various finite verb forms indicating the use of finite clauses. Verbs are also indicators of discourse type, and the ones found point to two types of discourse: a) descriptive/informational discourse, realised by the present tense of stative verbs. *Is, are* and *has* are used as main verbs, respectively, 142 out of 190, 39 out of 77, 49 out of 65. *Is* and *are* occur in the existential *there* construction for a total of 37 occurrences; when auxiliaries, they occur in the passive voice with a stative value:

> Bank House **is** a four bedroomed "black and white" style 15th century
> washing machines. Outside **there is** a stone built coach-house at the entrance
> garden to the front of the house **is enclosed** by a picket
> Local shops, post office and school **are** all within walking distance.
> kitchen. To the first floor **there are** three bedrooms and family bathroom.
> and the rear entrance to the property **are accessed** via Isaacs Road.
> bath. Every window in the house **has** new blinds (fitted 2005).
> floors and new carpets. Property also **has** an enclosed garden and

b) directive discourse, realised by activity verbs (*email*[2], *click, call, contact*), only used as imperatives, having the pragmatic value of instruction or invitation (Janoschka 2004: 135-138) and, except for *call*, characterising Web discourse.

The two groups of verbs have also been found in the promotional discourse of tourism (Pierini in press). What is characterising the discourse under examination is the use of past participles denoting features of properties (*fitted, detached, glazed, built, presented*). By scanning their concordances, we note that they very frequently function as premodifiers, each occurring in particular patterns. For example, *detached* co-occurs with *property/house/cottage/bungalow*; *glazed* mainly with *window(s)*; *presented* is modified by evaluative adverbs. Here are some typical patterns:

> comprises: cloakroom, lounge, dining room, **fitted kitchen**, utility and a gym
> Location: Poole Lovely modern **purpose built flat** in quiet cul-de-sac within
> bedroom down stairs with **built in wardrobe** and patio windows leading
> Price: 174950 An exceptionally **well presented** and spacious three bedroom semi-detached
> the detached properties, a **tastefully presented** three bedroom detached residence providing

6.3 Nouns

The vocabulary of any domain-specific discourse consists of general language words, technical and semi-technical terms (Dudley-Evans & St John 1998:80-83). Since messages are addressed to a wide audience, the most frequent nouns in the corpus are general words or semi-technical terms, which can grouped under two main fields, commerce and house building (Table 5). Adverts talk about the product promoted by using concrete words referring directly to the product (Leech 1966: 57); so the richest area is that including nouns referring to the parts and features of properties.

---

[2] By analysing the concordances, we could distinguish between *email* verb (70 occurrences) and *email* noun (14).

commerce/marketing		house building		
		types of home	parts of home	fixtures & fittings
property	132	*house* 123	*bedroom* 126	*floor* 75
sale	108	*flat* 40	*kitchen* 107	*heating* 44
price	64		*bedrooms* 102	*gas* 43
seller	63		*bathroom* 97	*door* 38
rent	52		*room* 86	*window* 36
ad	50		*garden* 79	*shower* 32
			front 67	
			reception 64	
			rear 62	
			bed 61	
			lounge 49	
			bathrooms 46	
			entrance 34	
			rooms 33	
			garage 30	

Table 5. *Nouns occurring at least 30 times*

In the first column, we find the most frequent lexical word in the corpus, i.e. *property*, a semi-technical term establishing the topic of the texts and functioning as hypernym of the words denoting the types of home (*house, flat*). The presence of *rent* is determined by the fact that about 50% of the properties are *for sale or rent*. In the column 'types of home', we find *house* and *flat*, heavily pre- and postmodified by elements detailing the property:

> An impressive 1 bedroom spacious **flat** within walking distance to Aldershot is spacious newly converted one bedroom **flat** on the first floor of
> This is a nice 3 bed semi detached **house** in a quiet neighbourhood
> 3 bathroom, 4 reception room detached **house** of considerable quality and character.

The most 'crowded' area is that referring to the parts of properties. The high frequency of *bedroom* and *bedrooms* is due to two different uses: the

word denotes a room for sleeping in, as in general language; or, as typical of property ads, the number of bedrooms is used as a conventional measure of the size of the property[3]. Being a basic information, it is mentioned both in the heading and the description. *Bed* is placed in this column on the basis of its concordances: it results to be the shortening of *bedroom*, only occurring as pre-modifier in patterns informing about the size of the property, such as *1/2/3 bed property/house/flat*. Another point is the use of *reception* and *lounge*: they are typical of British property ads, preferred to *living room* and *dining room* occurring, respectively, 21 times and 13 times.

In the column 'fixtures & fittings', we find nouns that co-occur with premodifiers having a designative or categorising function (e.g. *gas heating)*, thus going into a 'technical' description of the item. Consider the following concordances:

> Morning room and kitchen to **ground floor**. Three bedrooms and bathroom to
> hallway leads you up to the **flat's first floor** entrance. A super
> shower/toilet with **new porcelain tiled floor**. Two downstairs double bedrooms

From an analysis of the concordances it has emerged that premodification and postmodification, with a preference for the former, are syntactic devices largely employed to compact information.

6.4 Adjectives

word	frequency	percentage
double	124	0.68
large	93	0.51
central	43	0.23
good	37	0.20
new	37	0.20
private	35	0.19
modern	34	0.19
spacious	32	0.17

Table 6. *Adjectives occurring at least 30 times*

[3] It is a culture-bound measure, typical of Anglo-Saxon culture; in Italian, for example, the size of a home is indicated by means of the number of *stanze* (rooms), including the reception.

Unexpectedly, we do not find many adjectives in the corpus: the items in Table 6 belong to a familiar vocabulary, have a not high frequency, and all occur in attributive position. From a semantic point of view, adjectives can be placed on a cline from the most objective to the most subjective (Pierini 2006: 10-11): at one extreme, there are denotational adjectives denoting inherent stable qualities; at the other extreme, adjectives expressing evaluation; in between, adjectives describing sensory perceptions or qualities whose value is dependent on some implied norm or their co-text.

Out of eight adjectives, three are denotational (*double, central, private*) and co-occur with a very limited set of nouns in recurrent patterns, as illustrated by the following concordances:

> First floor, family bathroom, two **double bedrooms** and a child's bedroom
> family bathroom. Benefiting from mostly **double** glazed **windows**
> Gas Central Heating **Private parking**. Near Bus Routes and local shop
> lounge & Kit, bathroom & w/c, **private garden**, easy access to local
> The property has gas **central heating system** and the windows are

One is evaluative (*good*) and proves to be an all-purpose adjective occurring with a number of nouns (e.g. *size, investment, decoration, condition, standard*). The others are adjectives whose value depends on norms or the noun they modify; they co-occur with nouns denoting the property or its parts (*large, spacious)*, or its features as well (*new, modern*):

> No chain. A **modern** detached **property**, with approximately 150 square
> luxury bathrooms. Fully integrated **modern kitchen** with all appliances
> separate recessed storage points and **modern** wood effect **flooring**. Along
> the hall

The data suggest that explicitly positive evaluation, commonly expressed by means of adjectives, is not a discursive feature of classified advertising.

6.5 Abbreviations and acronyms

We have found two types of formations contributing to conciseness (Table 7); some were extracted electronically, others manually because

abbreviations		acronyms	
approx	approximately	CH	central heating
bedrms	bedrooms	DG	double glazed
combi (central heating)	combination/	EPC	Energy Performance
	combined		Certificate (part of HIP)
dbl	double	GC	gas central
dble	double	GCH	gas central heating
det	detached (house)	HIP	Home Information Pack[4]
f/f	fully furnished	HMO	house in multiple occupation[5]
fridge	refrigerator	LPG	liquefied petroleum gas
gym	gymnasium	NHBC	National House Building
kit	kitchen		Council[6]
mins	minutes	o.n.o/ono	or near offer
mod	modern	OSP (for cars)	open space preserve
p/built (block)	purpose-built	PVC	poly vinyl chloride
p/x	part exchange[7]	PVCu / uPVC	poly vinyl chloride
rad	radiator		unplasticised
recep	reception	TRV	thermostatic radiator valve[8]
semi	semi-detached		
	(house)		
sep (wc)	separate		
shw	shower		
(high) spec	specification		
tdryer	tumble dryer		
wc	water closet		
w/m	washing machine		
w/basin	wash basin		
wh (basin)	wall-hung		

Table 7. *Abbreviations and acronyms and their full forms*

[4] The HIP is a set of documents providing important information about a property, such as its energy efficiency, legal title and planning permission (see http://www.houseinformationpacks. gov.uk).

[5] It refers to residential property where 'common areas' exist and are shared by more than one household.

[6] The NHBC is the standard setting body and warranty provider for new or newly-converted homes (see http://www.nhbc.co.uk).

[7] The occurrence found is *£ 279,000 p/x taken*. The expression refers to a way of buying a good by giving a similar object as part of the payment for a more expensive one.

[8] The occurrence is *radiators with TRV regulators*.

not captured by the software programme. Acronyms are a vital part of specialised vocabularies, used in speaking as well as writing. Abbreviations are a writing-based device, which consists in the shortening of words by leaving out some letters; their degree of shortening is variable and their source form can be one word or one phrase (López Rúa 2002). Abbreviations differ from clippings generally described as a speaking-based process, prototypically concerning one word.

Both elements have a low frequency that may be explained by less tight spatial constraints than in print. The most frequent are: *approx* (21 occurrences), *ono* (19), *wc* (17), *uPVC/PVCu* (8), *PVC* (3) and *GCH* (6). The others occur one or two times.

As for abbreviations, three are better described as clippings (*combi, fridge, gym*). Among the others, some are widely used in writing (e.g. *approx*); most of them are typical of property ads, can only be interpreted on the basis of their co-text and appear to be not motivated (e.g. *dbl, kit, mod, shw*) because of the wider space available on the Web. They differ from *bed* as shortening of *bedroom* (see 6.3) that is genre-specific in print as well as on the Web.

With regard to acronyms, except for one currently used in general language (*PVC*), the elements found are technical: one belongs to the field of commerce (*ono*), the others to house building (e.g. EPC, HIP, NHBC). One can only be interpreted on the basis of its co-text: OSP, usually referred to 'green' areas of protected or conserved land or water, here refers to space reserved for cars.

7. *Conclusions*

Technology affords ways of doing things that people can then discover and exploit. So the development and enacting of genres is fundamentally a social process: genres emerge from communicative practices, and they, in turn, shape those practices. The online classified ad appears to be a popular Web genre that is solidifying its identity: it has a stable format but length and style vary across the texts in our corpus. Such variation reflects individual preferences and writing skills of the text producer on one side, and extralinguistic factors, on the other side. We have seen that many texts seem to be in search of a scannable style suitable for the new

medium, as required by the rhetoric of Web communication. Brevity, which is motivated by tight spatial constraints in print, is less crucial on the Web and is motivated by the need for readability, that is, by factors related to users, i.e. time and perception.

As for discursive features, online ads combine informational-descriptive and directive discourse, involving direct user addressing. As for syntactic aspects, the research has highlighted that the messages exhibit a combination of finite, non-finite and elliptical clauses, juxtaposition of syntactic groups, and syntactic complexity at the phrase level, with a preference for premodification. As for lexical aspects, the analysis has revealed that the vocabulary of property ads is largely made up of general words and semi-technical terms belonging to the fields of house building and commerce, while the use of abbreviations and acronyms is limited. It has also shed some light on the phraseology of classified property advertising.

Patrizia Pierini
Università degli Studi "Roma Tre"
Dipartimento di Linguistica
Via Ostiense 236 - 00146 ROMA
pierini@uniroma3.it

References

Bex, Tony, 1996, *Variety in Written English: Texts in Society, Societies in Text*, London, Routledge.

Bhatia, Vijay, 2004, *Worlds of Written Discourse*, London/New York, Continuum.

Biber, Douglas / Johansson, Stig / Leech, Geoffrey /Conrad, Susan / Finegan, Edward, 2000, *Longman Grammar of Spoken and Written English*, Harlow, Pearson Education.

Bolter, Jay D. / Grusin Richard, 2000, *Remediation: Understanding New Media*, Cambridge (Mass.), the MIT Press.

Bowker, Lynne / Pearson, Jennifer, 2002, *Working with Specialized Language: A Practical Guide to Using Corpora*, London, Routledge.

Bruthiaux, Paul, 1996, *The Discourse of Classified Advertising: Exploring the Nature of Linguistic Simplicity*, New York, Oxford University Press.

Bruthiaux, Paul, 2005, "In a Nutshell: Persuasion in the Spatially Constrained Language of Advertising". In: Halmari, Helena / Virtaten, Tuija (eds.), *Persuasion Across Genres: A Linguistic Approach*, Amsterdam, John Benjamins: 135-151.

Dudley-Evans, Tony / St John, Maggie Jo, 1998, *Developments in ESP: A Multi-disciplinary Approach*, Cambridge, Cambridge University Press.

Garzone, Giuliana, 2007, "Genres, Multimodality and the World Wide Web: Theoretical Issues". In: Garzone, Giuliana / Poncini, Gina / Catenaccio, Paola (eds.), *Multimodality in Corporate Communication*, Milano, Franco Angeli: 15-30.

Gesuato, Sara, 2007, "Evaluation in Back-cover Blurbs". *Textus* XX/1: 83-101.

Giltrow, Janet / Stein, Dieter (eds.), 2009, *Genres in the Internet*, Amsterdam, John Benjamins.

Goddard, Angela, 2002, *The Language of Advertising*, London/New York, Routledge, 2nd edition.

Gotti, Maurizio, 2005, *Investigating Specialized Discourse*, Bern, Peter Lang.

Jacqmain, Monique, 1973, *Il linguaggio della pubblicità. Uno studio sulle inserzioni nella stampa italiana,* Firenze, Sansoni.

Janoschka, Anja, 2004, *Web Advertising: New Forms of Communication on the Internet,* Amsterdam, John Benjamins.

Kress, Gunther / van Leeuwen, Theo, 2001, *Multimodal Discourse: The Modes and the Media of Contemporary Communication*, London, Arnold.

Kwasnik, Barbara H. / Crowston, Kevin, 2005, "Genres of Digital Documents". Introduction to the Special Issue of *Information Technology & People* 18/2: 76-88, available at http://genres.syr.edu/papers/itp2005genreintro.pdf (last accessed 18.02.2010)

Leech, Geoffrey N., 1966, *English in Advertising*, London, Longmans.

Lewis, Diana M., 2003, "Online News: a New Genre?". In: Aitchison, Jean / Lewis, Diana M. (eds.), *New Media Language*, London/New York, Routledge: 95-104.

López Rúa, Paula, 2002, "On the Structure of Acronyms and Neighbouring Categories: a Prototype-based Account". *English Language and Linguistics* 6/1: 31-60.

Lynch, Patrick / Horton, Sarah, 2004, *Web Style Guide*, 2nd edition, available at http://webstyleguide.com/ws2/index.html (last accessed 11.1.2010)

Mansfield, Gillian, 2001, "Job Advertisements and Applications: the Reflection in the Mirror". *Linguistica e Filologia* 13: 185-215.

Morkes, John / Nielsen, Jakob, 1997, "Concise, Scannable and Objective: How to Write for the Web", available at http://www.useit.com/papers/webwriting/writing.html (last accessed 26.02.2010)

Morville, Peter / Rosenfeld, Louis, 2006, *Information Architecture for the World Wide Web*, Sebastopol, O'Reilly, 3rd revised edition.

Pierini, Patrizia, 2006, "Adjectives in English: Description and Applications to EFL Teaching". *Studi Italiani di Linguistica Teorica e Applicata* 35/ 1: 5-24.

Pierini, Patrizia (in press), "Web Corpora for Web Translation: Discovering Translation Equivalents in Comparable Tourism Corpora". In: Di Martino, Gabriella / Lombardo, Linda / Nuccorini, Stefania (eds.), *Challenges for the 21st Century: Dilemmas, Ambiguities, Directions*, volume: Language Studies, 24th AIA Conference (Rome, 1-3 October 2009), Roma, Edizioni Q.

Sager, Juan / Dungworth, David / McDonald, Peter, 1980, *English Special Languages*, Wiesbaden, Oscar Brandstetter.

Santini, Marina, 2006, "Identifying Genres of Web Pages", TALN 2006 (Leuven, 10-13 April 2006), available at http://www.nltg.brighton.ac.uk/home/Marina.Santini/06_04_c_MSantini_TALN_2006.pdf (last accessed 15.02.2010)

Scott, Michael, 2007, *Wordsmith Tools 5.0,* available at http://www.lexically.net/wordsmith/

Suter, Hans J., 1993, *The Wedding Report*, Amsterdam, John Benjamins.

Swales, John, 1990, *Genre Analysis,* Cambridge, Cambridge University Press.

Tognini Bonelli, Elena, 2001, *Corpus Linguistics at Work,* Amsterdam, John Benjamins.

Ursini, Francesca, 2005, "Il testo in Rete e i suoi principi organizzativi retorici". In: Rosati, Luca / Venier, Federica (a cura di), *Rete retorica. Prospettive retoriche della Rete*, Perugia, Guerra Edizioni: 31-64.

Web sites

www.freeads.co.uk (last accessed 22.12.2009)

www.ukclassifieds.co.uk (last accessed 22.12.2009)

www.adflyer.co.uk (last accessed 22.12.2009)

www.loot.com (last accessed 22.12.2009)

www.thehorseexchange.com (last accessed 28.11.2009)

www.ukpetmart.com (last accessed 28.11.2009)

http://www.houseinformationpacks.gov.uk (last accessed 11.06.2010)

http://www.nhbc.co.uk (last accessed 11.06.2010)