

# NEWS AS A CONVERSATION. A COMPARATIVE ANALYSIS OF THE LANGUAGE OF ONLINE AND PRINTED NEWSPAPERS IN ITALY

Stefania Spina<sup>1</sup>

Previous studies on newspaper language have already pointed out how the language of written news has undergone significant changes in the last two or three decades. The aim of this study is to move a step forward and to compare the use of selected linguistic features in online and in printed Italian news. The main hypothesis is that Internet and communication technologies have introduced important transformations in the way news is written, organized and delivered, and that the consequences of these transformations are observable at all levels of linguistic analysis.

## **News as a conversation. A comparative analysis of the language of online and printed newspapers in Italy**

Previous studies on newspaper language (Bonomi, 2002; Conboy, 2010; Cotter, 2003; Dardano, 1986; Gualdo, 2007; Spina, forthcoming; Westin, 2002, among others) have already pointed out how the language of written news has undergone significant changes in the last two or three decades. The growing frequency of sentence-initial connectives (Cotter, 2003), the “reduction of hypotaxis” (Bonomi, 2002, p. 242), the increase in interactional features such as questions and imperatives, and the reduced use of passives (Westin, 2010), are all indicative of new styles adopted by written news, with the purpose of approaching the

---

1 Stefania Spina is Researcher at the Università per Stranieri di Perugia.

immediacy of oral styles and of reproducing forms of communication which are typical of television.

The aim of this study is to move a step forward and to compare the use of selected linguistic features in online and in printed Italian news. The main hypothesis is that Internet and communication technologies have introduced important transformations in the way news is written, organized and delivered, and that, from a linguistic point of view, the consequences of these transformations are observable at all levels of linguistic analysis (Lewis, 2003).

## **1. Sources of change**

The main sources of these changes may be linked to three important features of online communications: technology, information organization and interactivity. I will briefly describe the ways in which each of these features can influence the discourse of online news and produce linguistic changes.

### *Technology*

In the transition from paper to bits, digital texts are delivered through different devices, with specific technological features, such as screens of different sizes. Reading from a computer, a tablet or other mobile device has become a more demanding activity that requires new and differentiated abilities. Online news has to adapt to the technological constraints of these different devices, as well as to the deeply different nature of digital texts, which are dynamic, hypertextual and multimedia.

Digital and hypertextual texts tend to be organized in small chunks of contents, which are layered at different levels of detail. News is broken down into more fine-grained multimedia elements, and each one of them is autonomous and at the same time is a part of a cluster of dynamic, content-related and hierarchically-structured text. Lewis notices that *“these news clusters reveal an emergent new genre differing from both print and broadcast: a theme-based group of news objects held together graphically, overlapping with other such groups, and undergoing progressive updating”* (Lewis, 2003, p. 97).

Hence, online newspapers tend more and more to adopt a real-time organization, and to present news as streams of texts, where the notion of flow replaces the one of accomplished news story. The challenge, from a linguistic point of view, is the development of an online reader-

friendly style in the process of delivering news, in search of a more effective form of *readability*.

### *Information organization*

The web contains a great number of news sources, updated and accessible to the readers. This is the strength, but also, as it is well known, one of the weaknesses of the web, as it can cause problems of disorientation and information overload. Online news has to face these problems and prevent readers from losing direction. In the digital age, news is not rare anymore, it is redundant. In addition, it does not belong to professionals, but it belongs to everyone. Readers can access news directly, without the need of depending on publishing companies.

In this context, journalists are bound to assume new roles, which are connected with two main aspects.

On the one hand, journalists have not only to provide news, that readers can easily obtain elsewhere, but they also have to show them the best way to find it. Acting as intelligent aggregators, online newspapers might be expected to offer connections to news, showing the best route to find the most affordable sources. At present, the key task for journalists is more the selection and organization of connections than the original creation of news stories itself. Following Matheson (2004, p. 458):

One way of characterizing such emergent practices is as a journalism of linking rather than pinning things down, that is situated within a model of knowledge-as-process rather than knowledge-as-product. Readers [...] are set along paths of exploration rather than given nuggets of information, and the status of that information therefore changes.

On the other hand, the new role of journalists is connected to the quality of news, as they have to deal with their contextualization. It has been highlighted that we are facing the “*fading of the traditional ‘news story’*” (Lewis, 2003, p. 98); in online news, narratives are part of larger and more articulate stories. They are in fact a mix of original journalistic work and a wide range of other materials, published and unpublished, taken from heterogeneous sources, ranging from government official documents, legal papers, to statistical data, maps, timelines, etc. As Pavlik (2001, p. xi) puts it, “*This represents a new form of journalism*

*that places stories in a much more historical, political and cultural context”.*

From what has been said, it is clear that in online news the whole process of reporting and organizing news is redefined. This form of “contextualized journalism” (Pavlik, 2001) provides the readers with a new kind of enriched stories; online news has to face these new challenges, and thus become more manageable.

### *Interactivity*

Web 2.0 (O’Reilly, 2005) has transformed the old, unidirectional and static Internet sites into dynamic, interactive and collaborative environments, where a network of people share social relationships and content. As for other domains, users become producers of content; the news produced and shared by ordinary citizens is delivered alongside the news produced by professionals and publishing companies. The result is a mix of content that comes from different sources and media; the boundaries between journalist and citizen, published and unpublished, tend to fade out into a stream of cross-media information that blends across different channels.

In this respect a key role is played by social networks, which are becoming a sort of social and collaborative source for the dissemination of news. This new role requires, even for traditional publishers, a deep change in the classical model of unidirectional dissemination and a radical shift toward interactivity. As Curley (2004) puts it, “*consumers will want to use the two-way nature of the Internet to become active participants themselves in the exchange of news and ideas. The news, as ‘lecture’, is giving way to the news as a ‘conversation’.*”

Interactivity, on the other hand, is leading to new styles of language. The network of social relationships that arise between people who share a common interest encourages the use of informal and conversational styles (Fairclough, 1994); even the printed news tends to favor styles that range between public and private discourse (Spina, forthcoming). As a consequence, online news is also required to be strongly interactive.

These three characters – readability, manageability and interactivity – entail a deep change in the way news is reported; in this sense, they represent a great opportunity for journalists and news professionals to abandon conventional communicative styles, often connected to the printed medium, in search of a new linguistic usability. The main research question that this paper tries to answer is: do these

transformations entail different linguistic choices and specific changes in the discourse of online news?

## 2. Methodology

### 2.1 *An integrated approach*

In an attempt to answer this question, this study adopts an integrated approach. The starting point is the quantitative analysis of data gathered from a printed and an online newspaper; the data, stored and organized in two corpora, is queried in search of significant patterns.

The basic approach derives therefore from Corpus Linguistics and its methodologies of collection, annotation and interrogation of linguistic data (Sinclair, 1991; Spina, 2001). Once significant patterns are obtained, they are analyzed on the basis of a functional approach, in an attempt to interpret each pattern in the light of the function that it conveys at a discourse level. The perspective on the new functions of online news adopted in this study relies upon a functional approach to language, and is informed by Systemic Functional Linguistics (SFL). According to SFL (Halliday and Matthiessen, 2004), language simultaneously enacts three meta-functions: an ideational function, for construing human experience; an interpersonal function, for negotiating relationship, and a textual function, devoted to the organization of information. My assumption is that the interpersonal function is acquiring an increasingly more prominent role in the context of news, a genre where the ideational functional is traditionally dominant. The aim of this study is thus to reveal some of the instantiations of the interpersonal function in online Italian news, based on a data-oriented approach.

### 2.2 *Data*

The data used for this study was extracted from two Italian newspapers, one in its printed version and the other natively online. The printed one is the *Corriere della Sera*, the first national Italian daily newspaper for daily average diffusion (482.800 copies in December 2011<sup>1</sup>). The *Corriere della Sera* was founded in 1876 and is headquartered in Milan.

---

1 Source: *Prima Comunicazione*, <http://www.primaonline.it/2012/03/19/103569/quotidiani-dicembre-2011/>).

The second newspaper is *Il Post* (<http://www.ilpost.it>), an Italian online newspaper that started its activity in April 2010. *Il Post* has 90.000 daily unique visitors and it is updated continuously, from 7 in the morning until midnight<sup>1</sup>. The choice of a newspaper originally created for an online distribution is not casual: online newspapers that are derived from printed versions tend to retain similar styles and often the same texts of their original printed form, while native online news is designed and organized as an online resource from scratch.

Table 1 shows the composition of the two corpora extracted from the two newspapers; their time-frame partly overlaps, as the data from the *Corriere della Sera* covers the period January-April 2012, while the data from *Il Post* refers to the period October 2011-June 2012. All the articles from *Il Post* in the 9 months time-frame were automatically extracted through an RSS feed, while only a selection of the *Corriere della Sera* articles was included in the corpus; more specifically, articles were divided into 8 categories, depending on their type and topic: editorials, news stories, politics, foreign news, business, sports, culture and arts.

	time-frame	n. of articles	n. of words
<i>Corriere della Sera</i>	January-April 2012	5.465	4.054.157
<i>Il Post</i>	October 2011-June 2012	7.717	3.055.948

**Table 1.** The composition of the two newspaper corpora

After having been extracted and collected, the selected texts were annotated at two different levels; at the first level, an Xml annotation was used to add some metadata (the articles date and type), and to describe the articles structure, which is basically composed by headlines, subheadings and the text itself. The second level of annotation (post-tagging and lemmatization) added explicit linguistic information to the raw texts. This multilevel annotation allows the query and extraction of specific linguistic patterns in all the texts or only in selected sections.

---

1 Source: personal conversation with the Director, Luca Sofri (May 2012).

### 2.3 Selected linguistic features

The selection of linguistic features is based on previous studies on linguistic changes in newspapers (Bonomi, 2002; Convoy, 2010; Cotter, 2003; Dardano, 1986; Gualdo, 2007; Lewis, 2003; Ljung, 1997; Spina, forthcoming; Westin, 2002); although not specifically focused on the genre of news, it also relies upon some seminal works by Biber (1988, 1995) on register variation.

Given the previous analysis of the changes occurring in the organization of news, the choice fell on linguistic features that belong to two different dimensions: one which is traditionally considered as prototypical of the language of news, information density; the other related to the new style that news is holding in its online form: interactivity.

Some brief remarks are needed on the meaning of *interactivity*: it is not only a matter of informality: language of news tends more and more toward informality since the last decades (Westin, 2002), much earlier than the rise of the digital era. It is neither only a matter of personal involvement, as it was proposed by Biber (1988). By *interactivity* I mean the conscious attitude of the writer of addressing directly to his readers, as if they could reply to his/her address. This strong attitude to online written conversation, introduced by social networks, where users interact with each other and create interpersonal bonds (Zappavigna, 2012), has an influence on the attitude of online journalists, who have a conversational approach with their readers. Interactivity is thus a search for informality, an attitude of personal involvement, but most of all a marker of conversationality (Fairclough, 1994).

The linguistic features included in this study are a small part of those that Biber (1988) proved to be markers of the two functions of interactivity and information density. The selection was made upon a fine-tuning to the specificity of Italian, that, for example, has no contractions. To this selection I added two new features that, in my opinion, can provide supporting evidence: sentence length in number of words and noun-noun combinations with preposition-deletion, as in *pausa pranzo* instead of *pausa per il pranzo* (*lunch pause*), both of them as markers of information density.

The resulting selected features are listed in table 2.

Interactivity	Information density
discourse markers-	nominalizations
second person pronouns	noun-noun combinations with preposition deletion
amplifiers	sentence length-

**Table 2.** The list of the selected linguistic features in the news articles

The frequencies of all these features were searched in the texts of the two newspapers articles. Headlines, instead, deserve a separate discussion, because of their relevance, their structural and textual differentiation (De Benedetti, 2004) and their function of catching the reader's attention, fitting at the same time in a very limited space.

Accordingly, a separate analysis was carried out on the two newspapers headlines, using the same two dimensions selected for the articles, but using linguistic features that apply expressly to the genre of headlines. The resulting linguistic features analyzed in headlines are shown in table 3.

Interactivity	Information density
direct questions	nominalizations
amplifiers	sentence length
non-phrasal coordination	lexical density
words belonging to fundamental vocabulary	

**Table 3.** The list of the selected linguistic features in headlines

In the next section I will briefly describe the functions associated to each of these features and I will present the results of the frequency counts. Then I will discuss the results and propose their interpretation.

### 3. Results

#### *3.1 Features marking interactivity in headlines*

The measures and frequency counts of the selected features for interactivity are schematized in table 4; what follows is a brief overview of each of the features.



### *Direct questions*

As they presuppose someone answering the questions and thus they imply a concern with interpersonal functions, direct questions have been studied as markers of interactivity (Biber, 1988; Westin, 2002).

In a first phase I have considered only direct questions terminated with a question mark; however, a closer examination has revealed that, especially in *Il Post*, questions not ending with the question mark are also very common, like the one in the example (1):

(1) *Cosa è successo oggi in Libia* (What happened in Lybia today)

The frequency count, thus, includes all the direct questions that end with a question mark and that are introduced by an interrogative pronoun or adverb like *chi, che cosa, dove, perché quanto, come, quale* (the Italian equivalent forms of *wh*-questions).

The results show a clear prevalence of direct questions in online headlines compared to printed ones (I will discuss the differences between question types later).

### *Amplifiers*

As Biber (1988) puts it, amplifiers mark *heightened feeling*, and their use reveals interpersonal interaction and the expression of personal feelings. This function is shown in the example (2):

(2) *Salon ha cambiato (molto molto) la sua grafica.* (Salon has (much much) changed its graphics)

The adverbs marking this boosting effect in the degree of intensification that were included in the query are some of the most frequent in Italian: *estremamente, assolutamente, molto, del tutto, enormemente, interamente, totalmente, perfettamente* (*extremely, absolutely, very, altogether, enormously, entirely, totally and perfectly*).

The data shows that the function of personal involvement connected with amplifiers is more frequently used in the headlines of *Il Post*.

### *Non-phrasal coordination.*

Sentence-initial connectives have been proved to be increasingly used in newspapers (Cotter, 2003; Spina, forthcoming). In recent years their use in headlines of Italian newspapers has become a crystallized

form of opening, that simulates continuity with some preceding statement (Antonelli, 2007), like in the following example (3):

(3) *E ora tutti contro Gingrich.* (And now all against Gingrich)

In the data from the two newspapers, non-phrasal coordination of the two connectives *e* and *ma* (*and*, *but*) is much more frequent in the *Corriere della Sera*.

#### *Words belonging to fundamental vocabulary*

Headlines have to catch the reader's attention; to that aim, they have to be interactive and attractive, but also understandable. The use of simple and general words, similar to those commonly used in informal conversation, is a feature that brings headlines closer to an interaction. In order to obtain a measure of the distribution of words in ranges of difficulty, I have matched the list of all the headlines lemmata to the lemmata included in the fundamental vocabulary of Italian (De Mauro, 1980); this list includes the 7,000 words that can be understood by any Italian native speaker, with any degree of schooling. The results show a more widespread use of general and simple words in the headlines of the online newspaper.

Table 4 provides a summary of measures and frequency counts for the features marking interactivity in the two newspapers headlines<sup>1</sup>.

	<i>Corriere della Sera</i>		<i>Il Post</i>	
	raw frequency	per million words	raw frequency	per million words
direct questions	165	2667,2	548	5428,4
amplifiers	5	80,8	44	435,9
non-phrasal coordination	109	1762	48	475,5
fundamental vocabulary	32563	526381,3	58049	575021,5

**Table 4.** Measures and absolute and normalized frequencies (per million words) for the features marking interactivity in headlines

<sup>1</sup> All the differences between these and the following two series of raw values are highly significant, since chi-squared tests resulted in p-values < 0.0001.

### ***3.2 Features marking information density in headlines***

#### *Nominalizations*

Nominalization is the process of deriving a noun from another part of speech through the addition of affixes. As Biber (1988) and Westin (2002) point out, nominalizations are usually connected with texts of high information density, because they tend to integrate different information and functions in one single noun (Biber, 2003; Ni, 2003). In this study I only have considered nouns derived from verbs, and, among them, only those ending in *-mento*; this class is very productive in Italian and usually includes abstract nouns, like *fallimento* (*failure*). In the frequency count, the printed newspaper uses nominalizations in headlines more often than the online newspaper.

#### *Headline length in words*

The number of words per sentence is traditionally considered an index of sentence complexity (Bonomi, 2002). In the case of headlines, however, since most of them are monoclausal, I have calculated the number of words per headline: the higher the value, the more complex the sentence. The data shows that online news has shorter headlines.

#### *Lexical density*

Lexical density is a measure of the proportion of lexical and non-lexical items in a given text (Ure, 1971). It can provide clues on the weight of content words in comparison with grammatical words, and thus on the degree of information compression in a text. Halliday (1994) highlights that lexical density is a parameter to estimate the level of *activeness* versus *reflectiveness* of a text. Usually, written texts are more reflective and tend to have a higher lexical density, while texts closer to spontaneous speech are more active and have a lower lexical density. The value of lexical density proves to be higher in the printed newspaper corpus.

Table 5 provides a summary of the results of measures and frequency counts for the features marking information density in the two newspapers headlines.

	<i>Corriere della Sera</i>	<i>Il Post</i>
nominalizations	179 (2895,5 per million words)	194 (1921,7 per million words)
headline length in words	11,3	9,3
lexical density	37,95	32,88

**Table 5.** Measures and absolute and normalized frequencies (per million words) for the features marking information density in headlines

### ***3.3 Features marking interactivity in newspaper articles***

#### *Discourse markers*

Discourse markers (Schiffrin, 1987) are lexical items that can assume different and simultaneous functions in a sentence, depending on the context. They can for example connect different elements within sentences, or mark specific interpersonal relationships among participants. As such, they are mainly used in spoken language; however, they are not rare in informative written texts either, where they can be considered a good clue of the search for interactivity (Biber, 1998). For this study I selected the following markers, since they effectively represent the great variety of functions performed by discourse markers (the function of fillers, or of marking request for attention, topic change, reformulation, discourse planning, stressing, hedging, backchanneling, etc.): *diciamo, beh, insomma, allora, ecco, a proposito, guarda, dicevo, no?, vero?, sai/sapete, se non sbaglio, perfetto, esatto, voglio dire (let's say, well, so, then, here, by the way, look, I said, isn't it? You know, I think, perfect, right, I mean)*. The frequency counts report a substantial equivalence in the use of discourse markers in the two corpora.

#### *Second person pronouns*

They are most frequently used in interactions, such as face-to-face conversations; however, they can be used in written informative discourse with the aim of performing an assumed conversation with readers (Westin, 2002).

The selected pronouns are second person singular and plural personal pronouns (*tu, te, voi, ve*) and possessive pronouns (*tuo, vostro*).

The results show a more frequent use of second person pronouns in online news articles.

### *Amplifiers*

For amplifiers in articles I have adopted the same methodology used for headlines. The results show that online news articles use amplifiers noticeably more than printed news.

Table 6 provides a summary of frequency counts for the features marking interactivity in newspaper articles.

	<i>Corriere della Sera</i>		<i>Il Post</i>	
	raw frequency	<i>per million words</i>	raw frequency	<i>per million words</i>
discourse markers	670	169,4	467	158
second person pronouns	919	232,4	974	329,6
amplifiers	3627	917,3	4938	1671,1

**Table 6.** Absolute and normalized frequencies (per million words) for the features marking interactivity in newspaper articles

### **3.4 Features marking information density in articles**

#### *Nominalizations*

The same considerations discussed for headlines apply to the news articles, where I examined the same nouns in *-mento* derived from verbs. The results show a slightly higher frequency in the printed news. *Noun-noun combinations with preposition deletion.*

These combinations are basically prepositional phrases where the preposition has been deleted. The deletion of preposition indicates an effort in the direction of a more compressed organization of information. Diachronically, many of these combinations have become nominal compounds and have totally lost any trace of the preposition, like *fine settimana*, derived from *la fine della settimana* (*weekend*), that has even undergone a change in genre. The results show a clear prevalence of these constructs in printed news.

### *Sentence length.*

Following Westin (2002), for the calculation of this index I have counted all the occurrences of sentence delimiters (full stop, exclamation mark and question mark), and then I have calculated the average number of words per sentence. Surprisingly, sentences prove to be longer in online news (29,13 words per sentence) than in printed news (25,52 words per sentence).

	<i>Corriere della Sera</i>		<i>Il Post</i>	
	raw frequency	<i>per million words</i>	raw frequency	<i>per million words</i>
nominalizations	14698	3717,2	9212	3117,4
noun-noun combinations with preposition deletion	1507	381,1	315	106,6

**Table 7.** Absolute and normalized frequencies (per million words) for the features marking information density in newspaper articles.

## **4. Discussion**

The picture offered by the analysis of data from headlines is fairly clear: online news employs a double number of direct questions, a five times higher number of amplifiers and a higher percentage of general and common words. In addition, online headlines are shorter and less packed with dense and compressed information.

The strong presence of direct questions is a sign of a marked tendency towards conversationality. As Westin (2002, p. 25) puts it, questions may also

be used in writing to make it possible for the author to enter into a fictive conversation with the reader. Questions appearing in [newspapers] must thus, when they are not part of quoted speech, be either rhetorical questions or a device used to involve the reader in a joint communicative venture.

However, a deeper analysis of the function conveyed by the questions used in online headlines reveals more than a search for dialogue and complicity with readers. A great part of the questions used in online

headlines have a different function if compared to those retrieved in printed headlines. Examples 4 and 5 are taken from the *Corriere della Sera*: the two questions are both part of a reported speech, where the author relates the words pronounced by someone else.

(4) “*Modello francese? Per noi è tardi*”. (“French model? For us it’s late”)

(5) *Conte: “Meglio senza Ibra? Non so, ve lo dico alle 23”*. (Conte: “Better without Ibra? I don’t know, I’ll tell you at 11”)

This is a very common pattern in Italian headlines: the stretch of reported speech is often preceded by a proper name + colon, as in example 5, indicating the person to whom those words are attributed. Almost 20% of titles from the *Corriere della Sera* headlines show this pattern.

From what has been said, it is clear that in printed headlines questions are mostly used in reported speech, where words pronounced by others are quoted. The function of direct questions is not to “involve the reader in a joint communicative venture” (Westin, 2002, p. 25), but to simulate a real-time dialogue between the journalist and someone else.

In contrast to this, things radically change if we consider online headlines: examples 6-8 are not part of a quoted speech, but represent a real address to the readers, through which the author tries to establish a real connection with them. By asking a direct question, the journalist wants to communicate that he/she is going to answer that question, suggesting a solution to that problem. In addition, questions are expressed in very common words, all part of the fundamental vocabulary, thus understandable by everyone. The idea that they suggest is consequently an idea of simplicity and immediacy.

(6) *Che cosa succede il 15 ottobre?* (What is happening on the 15<sup>th</sup> of October?)

(7) *Perché a Roma le scuole sono chiuse?* (Why are schools closed in Rome?)

(8) *Che cosa succede stasera in Grecia.* (What is going to happen in Greece tonight)

The analysis of questions in headlines, therefore, shows a shift in their function: headlines have still the role to catch the readers’ attention,

but the way they do it is dramatically changing. The main quality that readers are supposed to be interested in is not only the fact of reporting some conversation or event happening in real-time, suggested by the quoted speech; with the diffusion of social networks, real-time has simply become obvious. The quality readers mostly seek in online news is orientation through that news.

The occurrence of interactive features in online headlines is an answer to this request for direction: questions are a simple way to propose to readers solutions to their problems of orientation. Faced to a topic that is reported and discussed in several news sources (for example, “What is going to happen in Greece?”), a direct question in the headline winks at the reader, searching for his complicity and, at the same time, promising him: “I will guide you through this complexity”. In this context, the use of non-specialized words and amplifiers goes in the same direction: an offer of guidance has to be clear and expressed with simple words, with a style capable of arousing the reader’s involvement.

This new function of headlines also produces, as a consequence, a radical change in the structure of the traditional printed titles, that is often bound to fixed patterns: the use of non-phrasal coordination, for example, has been described as constantly growing in frequency in the last years (Cotter, 2003; Spina, forthcoming); my frequency counts show that non-phrasal coordination has almost disappeared from native online headlines (I only found 48 instances in *Il Post*). According to Cotter (2003, p. 51), “these devices structure the text as well, facilitating reception of the message to the extent allowed by a medium lacking visibility and synchrony”; the decrease in the use of non-phrasal coordination is precisely connected to the decrease in the need of synchrony, which is now assumed by people used to browse the web, and to the parallel pursuit of new styles.

The data on information density in headlines shows other important differences between online and printed news. Printed headlines prove to be more compressed and lexically denser; online headlines, on the contrary, use fewer words per title and fewer nominalizations.

Shifting next to newspaper articles, frequency counts of the features marking interactivity, with the exception of amplifiers, that in online articles occur almost twice as much as in printed news, reveal, at least apparently, more blurred boundaries. However, a deeper analysis indicates that the way those features are used differs more than their amount. Similarly to what we have observed in headlines, in printed



articles a great part of second person pronouns and discourse markers occurs in reported speech. The next two couples of examples (9-12) illustrate this difference:

*Discourse markers*

*Corriere della Sera*

(9) “Tre valori in cui credi”, gli chiese la Bignardi. E lui: “Beh, l’onestà sicuramente...” (“Three values in which you believe”, Bignardi asked him. And he answered: “Well, surely honesty...”).

*Il Post*

(10) Guardo l’orologio. Trentacinque minuti. Beh, penso, posto carino. (I look at my watch. Thirty-five minutes. Well, I think, nice place).

*Second person pronouns*

*Corriere della Sera*

(11) L’ultimo “motivational speech” è stato tenuto proprio da Simon Brown: “Voi siete fortunati, [...] e guadagnate soldi divertendovi”. (The last motivational speech has been performed precisely by Simon Brown: “You are lucky, [...] and you earn money while having fun.”)

*Il Post*

(12) Credo che tutti voi conosciate il Teorema della scimmia instancabile. (I think that you all know the theorem of the tireless monkey.).

It seems clear that the selected features turn into markers of interactivity only when they achieve some kind of dialogue between author and reader. As it happens in headlines, in quoted speech they only have the function of suggesting that the reported conversation is taking place in real time, and thus of providing the discourse with a sense of “animation” (Antonelli, 2007).

As for discourse markers, if we fine-tune the frequency counts and consider as markers of interactivity only their occurrences outside quoted speech, the figures in table 6 should be replaced by the figures in table 8, and the difference between the two newspapers would appear more clearly defined:

	<i>Corriere della Sera</i>		<i>Il Post</i>	
	raw frequency	<i>per million words</i>	raw frequency	<i>per million words</i>
discourse markers	300	75,9	467	158

**Table 8.** Absolute and normalized frequencies (per million words) of discourse markers outside quoted speech in articles

The features marking information density in articles present similar differences in frequency: the noun-noun combinations with preposition deletion are three times more frequent in printed articles. This feature, as well as the greater number of nominalizations, reveals that printed news is more compressed, and the organization of linguistic elements favors a more dense packaging of information.

The increase in sentence length in online news, on the contrary, is quite unexpected: online articles have sentences on average almost four words longer than printed ones, while the trend of the last decades for newspapers was opposite, with a progressive reduction in the number of words per sentence (Westin, 2002; Spina, forthcoming). Again, the reason for this shift is connected to the way information is organized within the sentence: online news does not need to save space, and the progress of sentences can be looser and less compact, resulting in a higher number of words per sentences.

## 5. Conclusions

In their association with web 2.0 and social resources of the Internet, news has to face deep changes in the way it is written, organized and delivered. These changes regard many different aspects, and among them the way language is organized and used. In an attempt to adapt to the new medium, news tends to develop a different organization and a more reader-oriented style, trying to meet the need of readers to disentangle among many different sources of information.

The purpose of this study was to investigate the impact of these changes on how language is used in online news compared to printed news. To this end, the frequencies of selected linguistic features in two corpora of printed and online news have been compared. The linguistic features refer to the two dimensions of interactivity and information density, which are compared in their transition into the online environment; they

also belong to different levels of analysis: vocabulary (lexical density, words belonging to fundamental vocabulary, nominalizations), syntax (non-phrasal coordination, sentence length, noun-noun combinations with preposition deletion); discourse (questions, discourse markers, amplifiers, second person pronouns).

The results of the analysis have shown a marked tendency of online news towards conversationality (Fairclough, 1994); the occurrence of features such as questions, second person pronouns and amplifiers, is an answer to a need for interaction. Through headlines, journalists propose to readers to guide them through the complexity and the variety of news sources. This offer of orientation is a form of interaction between authors and readers. While in traditional newspapers interactive features respond to a demand of simultaneity, which printed news lacks, in online news they represent the answer to a request for direction.

A similar tendency to interactivity and conversationality is found in articles, where online news also reveal a less compressed organization of lexical items and a looser progress of sentences, which are longer and contains more general vocabulary.

This shift towards interactivity, as well as the new function assumed by online titles, proves that, from a functional point of view, the presence of the interactive meta-function, that traditionally has a minor influence in an informational genre like news, is growing in importance, with the transition from printed to online news.

## *References*

- Antonelli, G. (2007). *L'italiano nella società della comunicazione*. Bologna: Il Mulino.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. (1995). *Dimensions of Register Variation. A Cross-linguistic Comparison*. Cambridge: Cambridge University Press.
- Biber, D. (2003). Compressed Noun-phrase Structures in Newspaper Discourse. The Competing Demands of Popularization vs. Economy. In J. Aitchison & D.M. Lewis (Eds.), *New Media Language*. (pp. 169-181). London: Routledge.
- Bonomi, I. (2002). *L'italiano giornalistico. Dall'inizio del '900 ai quotidiani online*. Firenze: Franco Cesati.
- Conboy, M. (2010). *The Language of Newspapers: Socio-historical Perspectives*. London: Continuum.
- Cotter, C. (2003). Prescription and Practice: Motivations Behind Change in News Discourse. *Journal of Historical Pragmatics*, (4), 45–74.
- Curley, T. (2004). Opening keynote. *The Associated Press - Online News Association Conference*, Nov. 12. Retrieved from <http://conference.journalists.org/2004conference/archives/000079.php>

- Dardano, M. (1986). *Il linguaggio dei giornali italiani*, Roma-Bari: Laterza.
- De Benedetti, A. (2004). *L'informazione liofilizzata. Uno studio sui titoli di giornale (1992-2003)*. Firenze: Franco Cesati.
- De Mauro, T. (1980). *Guida all'uso delle parole*. Roma: Editori Riuniti.
- Fairclough, N. (1994). Conversationalization of Public Discourse and the Authority of the Consumer. In R. Keat, N. Whitely & N. Abercrombie (Eds.), *The Authority of the Consumer*. (pp. 253-268). London: Routledge.
- Gualdo, R. (2007). *L'italiano dei giornali*. Roma: Carocci.
- Halliday, M. A. K. (1994). Spoken and Written Modes of Meaning. In D. Graddol & O. Boyd-Barrett (Eds), *Media Texts, Authors and Readers: A Reader*. (pp. 51-73). Clevedon/Philadelphia: Multilingual Matters.
- Halliday, M.A.K., & Matthiessen, C.M.I.M. (2004). *An Introduction to Functional Grammar*. 3rd ed. London: Arnold.
- Lewis, D. (2003). Online News: A New Genre? In J. Aitchison & D. M. Lewis (Eds.), *New media language*. (pp. 95-104). London: Routledge.
- Ljung, M. (1997). Text Complexity in British and American Newspapers. In J. Aarts, I. de Mönink & H. Wekker (Eds.), *Studies in English Language and Teaching*. (pp.75-83). Amsterdam: Rodopi.
- Matheson, D. (2004). Weblogs and the Epistemology of the News: Some Trends in Online Journalism. *New Media & Society*, 6(4), 443-468.
- Ni, Y. (2003). Noun Phrases in Media Texts. A Quantificational Approach. In J. Aitchison & D.M. Lewis (Eds.), *New Media Language*. (pp.159-168). London: Routledge.
- O'Reilly, T. (2005). *What Is Web 2.0*. O'Reilly Network. Retrieved from <http://oreilly.com/web2/archive/what-is-web-20.html>
- Pavlik, J. V. (2001). *Journalism and New Media*. New York: Columbia University Press.
- Schiffirin, D. (1987) *Discourse Markers*. Cambridge: Cambridge University Press.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Spina, S. (2001). *Fare i conti con le parole. Introduzione alla Linguistica dei corpora*. Perugia: Guerra.
- Spina, S. (forthcoming). Changing Trends in Italian Newspaper Language. A Diachronic, Corpus-based Study. In J. Bamford, S. Cavalieri & G. Diani (Eds.). *Variation and Change in Spoken and Written Discourse: Perspectives from Corpus Linguistics*. Amsterdam-Philadelphia: John Benjamins.
- Ure, J. (1971). Lexical Density and Register Differentiation. In G. Perren & J.L.M. Trim (Eds.), *Applications of Linguistics. Selected Papers of the Second World Congress of Applied Linguistics*. (pp. 443-452). Cambridge: Cambridge University Press.
- Zappavigna, M. (2012). *Discourse of Twitter and Social Media*. London: Continuum.
- Westin, I. (2002). *Language Change in English Newspaper Editorials*. Amsterdam: Rodopi.