

A pragmatic cognitive model for the interpretation of verbal–visual communication in television news programmes

Article (Published Version)

Piazza, Roberta and Haarman, Louann (2016) A pragmatic cognitive model for the interpretation of verbal–visual communication in television news programmes. *Visual Communication*, 15 (4). pp. 461-486. ISSN 1470-3572

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/53167/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

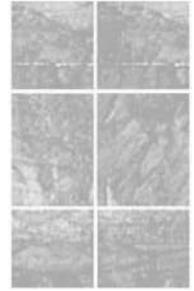
Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A pragmatic cognitive model for the interpretation of verbal-visual communication in television news programmes



ROBERTA PIAZZA

University of Sussex, Brighton, UK

LOUANN HAARMAN

University of Bologna, Italy

ABSTRACT

The combination of the verbal and the visual track in television news discourse poses a considerable analytical challenge. In the viewers' minds the co-habitation of these two semiotic channels triggers a complex network of inferential processes, based on expectations of coherence and relevance, with which they make sense of the representation of the world offered in the news. Through the analysis of a number of news items, this article considers the cognitive processes which viewers may activate when extracting meaning from the multimedial messages contained in television news. The analysis of news items from two British television networks offered by the authors traces the possible meanings that, it is assumed, become available to a potential, 'idealised' or 'implied' viewer, who accesses the information with some social and cultural knowledge of contemporary Britain. Building on existing studies, the article proposes a model for the classification of verbal-visual relations.

KEYWORDS

cognitive pragmatics • image-text relations • implicature in TV news • television news • verbal-visual communication

INTRODUCTION

Television news discourse poses a considerable analytical challenge because by its nature it is so complex. The combination of the verbal track with its various elements of direct speech, dialogue and voice-over, the sound track including signature music and ambient sound, and the visual track with the use of film footage, stills and graphics, make the television news programme a complex semiotic event that necessarily activates various channels of cognition. When watching news programmes, viewers follow the same inferential processes with which they make sense of the world by drawing on information whose bases are primarily social. While the meaning viewers make of news discourse is context-specific and strictly related to the very topic covered in the report, the comprehension process entails an expectation of coherence and relevance which drives viewers' inferential processes to see connections between individual elements even though they are not explicit.

Through the analysis of a number of news items, this article considers the cognitive processes which viewers may activate when extracting meaning from television news. Specifically, while avoiding a reflection on journalistic practice and intention, it investigates how sense emerges from the interaction between the verbal and the visual tracks and how the attribution of meaning is 'the outcome of sometimes complex and cumulative inferential work' (Montgomery, 2007: 97). The close analysis we offer of news items from two British television networks traces the possible range of meanings that we presume become available to a potential, 'idealised' or 'implied' viewer (Dahlgren, 1988: 288) who accesses the information with some social and cultural knowledge of contemporary Britain.

Since it is addressed to a general public with diverse professional expertise and affiliations, cultural background, political and religious creeds, interests and experiences, a news text 'cannot necessarily be treated as one that is finely recipient-designed' (Jayyusi, 1991: 170). Our interest therefore is not in the actual reception of any particular viewers, but rather in identifying the potentialities inherent in the television text for creating meaning, with the understanding, however, that 'the production of meaning can have a dimension which is very private and even idiosyncratic – we all have our own sphere of "personal knowledge"' (Dahlgren, 1988: 287, see also Thompson, 1984, in Higgins, 1991: 152).

Moreover it has been suggested that 'the messages available through television ... never deliver one meaning; they are, rather the site of a plurality of meanings' (Hall et al., 1976, in Wojcieszak, 2009: 462). Our attention, therefore, is directed to the *manner* in which sense is encoded in news stories. At no point do we suggest that the potentiality of the text, the possible meanings negotiated by viewers and the inferential processes emerging are the only ones available to an individual, given the equivocal and polysemic, even contradictory nature of certain combinations of verbal and visual texts (Andrews, 2007; Fiske, 1987; Juluri, 2002; Oguss, 2005; Tolson, 1996).

Our work builds on previous research on images in television news programmes which has addressed the role and function of the connection between words and images in the process of meaning-making. In so doing, we take a pragmatic/cognitive approach that considers discourse in context and focuses on the expectations that viewers may have when watching television news. Our aim in this article is thus to combine the insights of previous work with our own data analysis and to propose a model for the systematic classification of the relations between verbal and visual tracks in television news. To show how our work is framed within studies on the role and function of the verbal–visual synchrony in news reports and more generally on the relationship between these two planes, the following section reviews the most relevant research in the area.

CONTEXT OF THE STUDY

Work on the relationship between words and images in television news belongs to the general line of inquiry into the verbal–visual correspondence in texts of all kinds, inspired by the seminal work of Roland Barthes (1977), and to the research domain of multimodality following on from the work of O’Toole (1994) and above all Kress and Van Leeuwen (1996). Barthes’ study of newspaper photographs was ground-breaking in discerning three types of interaction between text and image: *anchorage*, in which the text supports the image, clarifying its meaning; *illustration*, whereby the image supports the text, again clarifying its meaning; and *relay*, in which the text and the image reciprocally reinforce each other. This characterization has been key in most subsequent research on verbal/visual combinations, particularly that focusing on images in the press (e.g. Andén-Papadopoulos, 2008; Griffin, 2004; Hartley, 1992; Zelizer, 2002, 2004). It has also been fundamental in sparking a large body of research on the image–text relationship and on multimodal texts in general within the framework of Hallidayan Systemic–Functional semiotics, covering fields as various as science, art, advertising, and news, (e.g. Bednarek and Caple, 2012; Caple, 2008, 2010; Lemke, 1998; O’Halloran, 1999; Royce, 1998).¹

Various systems or models for the classification of the verbal–visual relationship have been proposed. Martinec and Salway (2005), working with still images in a range of written genres, offer a system² for image–text relations based on Halliday’s (1985, 1994) concepts of status and logico–semantic relations and Barthes’ *anchorage*, *illustration* and *relay* mentioned above. The system, which aims to explain how images and words alter or complement each other, distinguishes between *unequal* relations if either mode modifies the other, *equal* status if they are both independent and triggering their own comprehension processes in the users, and *complementary* if the verbal and visual tracks are joined and equally modify one another (Martinec and Salway, 2005: 343). The model also deals with the logico–semantic relations of *expansion* and *projection*, i.e. with forms of qualification, specification and meaning addition of one track to the other.³

One of the earliest attempts at classification of moving images and text was Nichols' (1976) work on documentaries, later developed and extended by Van Leeuwen (1991). Based also on Barthes (1977), Halliday and Hasan (1976), Halliday (1985),⁴ and Martin (1983), Van Leeuwen (1991) offers a highly detailed analysis of conjunctive structure in five documentary sequences, setting out a model for categorising the ways in which elements on the visual plane conjoin those on the verbal. His work reveals an apparent diachronic 'shift in the relation between image and text' (p. 113). In the early film documentaries, the image has primacy: 'images function ... as empirical evidence, and text functions as "anchorage"'. In the television documentaries instead, the verbal has primacy: 'the authoritative text of the anchorperson precedes "images of the world", and in the voice over section directionality is reversed: the visual authenticates, particularises and exemplifies the verbal' (p. 113).

Griffin (1992) offers one of the earliest investigations of moving images in television news. He reflects on such issues as viewers' inattentiveness and the very nature of news reports that tends to be 'information thin and image rich' and agrees with Dahlgren (1988) that much of news viewing involves an affective and entertainment orientation rather than a rational and cognitive process (Griffin, 1992: 123). Against those who view words in news reports as determined by and dependent on the access to images (Altheide, 1987), Griffin argues that it is stories that actually determine the meaning of visuals. His close analysis of news stories regarding an Avianca plane crash in Colombia and drug smuggling charges brought against General Noriega illustrates how the visual procedures in television discourse fulfil four goals. They provide: 'symbols of access' to the location of the reported event and the people involved in it; 'signs of information' that verify and confirm journalists' reportage; 'loosely associated imagery' or visuals that are only tangentially related to the verbal report and that viewers, Griffin (1992: 136) claims, 'casually accept without questioning their veracity or relevance'; and 'continuity and transition', which is a main concern in news programmes. It can be seen how Griffin's insightful model points to the different functions the visuals realise at a macro level, specifying the type of relation between verbal and visual – a matter which receives greater attention in Meinhof's (1994, 1998) subsequent work.

Also working with moving images, specifically in television news, Wojcieszak (2009: 459) approaches the problem of how 'images in television news are contextualized, complemented, displaced, explained, or contradicted by the auditory channel and linguistic messages on the screen'. In her analysis of CNN, MSNBC, and FOX news coverage of the 9/11 terrorist attacks and Hurricane Katrina, and based on Barthes' (1977) observations about the role of images and text in film, she carries out a content analysis of her data and proposes a macro interpretative framework of the relation between what she defines as the three semiotic levels of news messages, i.e. the iconic, the linguistic and the audio. Her conclusions indicate that while in televised news coverage:

on-screen messages might be the primary means of directing the reading of images, with the concurrent voice-over providing additional information ... it is often the audio channel that supplies meaning or newsworthiness to iconic messages and determines the interpretation of news items. (Wojcieszak, 2009: 477)

In a more media discourse analytic perspective, the interpretative potential of combinations of visual and verbal texts has been highlighted in work by Thornborrow et al. (2012), a comparative study of the representation of Europe in UK, Italian and French television news, and by Thornborrow and Haarman (2012), an analysis of television coverage of the 2009 G20 summit. Pointing to the often personalised, ironic and stylized nature of such combinations in the discursive practices of UK television news, the authors describe and deconstruct meanings implied or made explicit through the 'artful exploitation of verbal and visual synchrony' (p. 389).

Two further studies have been fundamental to the present paper. Meinhof's (1994) seminal work on the 'double-encoding' of news reporting highlights the most crucial features of news texts and their openness to interpretation in terms of production and, more importantly for us, in terms of reception. She writes: 'no text is more than a meaning potential; a preferred reading is only one of many possible other readings, and no audience is just a receptacle for a given set of messages' (pp. 220–221). She agrees with other scholars that, in TV news, information from the soundtrack and images may be different and highlights that the difference 'between what we hear and what we see is not necessarily a difference between words and pictures ... but a difference in how we receive the information through our eyes or through our ears' (Meinhof, 1998: 25). Meinhof (1994: 217) envisages three different cases:

- *overlap*, if 'the film footage and the text share the same action component', even in the case of indirect and metonymic representations such as a flag suggesting a specific country;
- *displacement* if the verbal text discusses for example the causes of an event and visuals illustrate the consequences, as in the case of the report on an earthquake and the showing of relocated victims in provisional accommodation; and finally,
- *dichotomy* occurring when no coherence exists between the visual and verbal tracks, as in the case of a news item on Soviet troops in Afghanistan associated with a view of a cyclist on a market street in Kabul.

Similar to Griffin (1992), who captures the role of the images 'loosely associated' with words, Meinhof deals with the effect of context on viewers' comprehension and interpretative processes in her discussion of how a canister on a bicycle was understood by viewers to be carrying water or fuel, depending on the information given in the verbal report.

Montgomery (2007: 96–98), noting Meinhof’s three types of inter-relations, takes this line of inquiry further by establishing two Principles of Intelligibility that are seen to guide viewers in comprehending news reports:

- Rule 1, according to which ‘for any referring expression in the verbal track, [viewers will] search for a relevant referent in the image track’; and
- Rule 2, according to which viewers will ‘treat any element depicted in a shot in the visual track as a potential referent for a referring expression in the verbal track’.

Montgomery’s insights are fundamental to the present article, particularly with reference to his attention to the process of inference-drawing and his analysis of deixis in news reports at the verbal–visual conjunction, issues which we shall address and further elaborate here.

Our attempt to classify the verbal–visual ‘co-reference’ (to use Montgomery’s term), or ‘juxtaposition’ (Wiggin and Miller’s, 2003, term) relies on Sperber and Wilson’s Relevance Theory (1995[1986]). Differently from other pragmatic theories (e.g. Grice, 1975), Relevance Theory is also a cognitive theory concerned with the relation between meaning and context, and our adoption of this framework may thus be seen as consistent with Meinhof’s and Montgomery’s work. Like other pragmatic theories, it also conceives of meaning as determined by linguistic and non-linguistic factors (Chapman, 2011: 103). More importantly, Sperber and Wilson’s theory is concerned with the addressees’ mental processes and sees text interpretation as regulated by the Principle of Relevance, a general overarching communicative and cognitive principle that encompasses and subsumes Grice’s cooperative maxims. Furthermore, according to Relevance Theory, communication is based on an inferential process that ‘starts from a set of premises and results in a set of conclusions which follow logically from, or are at least warranted by, the premises’ (Sperber and Wilson, 1995[1986]: 12–13). Any process of interpretation is regulated by pragmatic principles that discriminate between what is said explicitly and what is implicit and hence inferred; therefore ‘pragmatic inference is just as important in determining what is explicitly communicated as in determining what is implicitly communicated’ (Chapman, 2011: 104). In any message, the explicit and the implicit therefore combine and the difference between them is a matter of degrees rather than a clear-cut division (Shiro, 2005: 171). Communication is seen as a process that Sperber and Wilson (1995[1986]: 63) call ‘ostensive-inferential communication’, according to which a ‘communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of the stimulus, to make manifest ... to the audience a set of assumptions’. It may be remembered that we started from the assumption that news is processed like any other kind of communication in which the viewer sees himself or herself as an addressee who can expect to take any act of ostensive–inferential communication as communicating a

'presumption of optimal relevance' (p. 158). Therefore, viewers' implicatures – defined by Sperber and Wilson as 'information that is communicated to the viewer (reader/listener) implicitly' (p. 273) – are generated on the grounds that news reports have some intention to communicate relevant meaning, even if in fact, in the case of news, not every second in a report is of course dictated by a will to convey a specific meaning.

Wiggin and Miller's (2003) study of persuasion in television advertising proposes a pragmatic approach to what they term the verbal–visual 'juxtaposition'. In particular, the authors address the issue of the metaphorical relation between verbal and visual, which encourages abstract thinking in viewers (see also Messaris, 1994). Wiggin and Miller describe verbal–visual metaphoric constructions in terms that are similar to the analysis of metaphors carried out below (in the section 'Style: artful conjunction of verbal and visual tracks') and similarly reference conceptual metaphor theory by Lakoff and Johnson (1980) and subsequent work on pragmatics and metaphor by Forceville (1996). What is most relevant to our study is the pragmatic approach of these authors, who view the meaning of metaphors as having weak or strong 'implicature' (Sperber and Wilson, 1995[1986]) according to the 'transparency' of the terms: TIME IS MONEY produces a strong implicature in its reference to time as a commodity; a weak implicature is generated if the metaphor is expressed as 'the *exchange rate* favors foreigners whose *spending power* lasts for a long or short time, depending on the currency' (Wiggin and Miller, 2003: 274, emphases in the original).

Building also on Montgomery's (2007) work on inference and his Principles of Intelligibility, in this article we propose to operate within the theoretical framework of Relevance Theory according to which viewers assume that any message manifestly conveys a willingness to communicate. In the present study we are concerned with the reception, rather than the production level of news programmes. While avoiding claims as to how news reports are produced (whether, in other terms, the words are made to fit the images, as would appear to be the current perception of good journalistic practice,⁵ or vice versa), we argue that viewers presume the optimal relevance of every message, which ties in necessarily with their expectation about the communicator's intention to convey an optimally relevant meaning. This point, we emphasise, is not equivalent to saying that only one meaning is obtainable from a specific shot, voice-over or a specific verbal–visual combination, nor are we saying that interpreting the verbal and visual messages in news items involves tracing the journalists' or camera operators' intention. Optimal relevance and intentionality are assumed as organising criteria in the viewers' mind and operate only at the cognitive level of viewers' expectations, thus playing an important role in the comprehension process.⁶

In conclusion, this is the relevance theoretic framework at the basis of the proposed classification of the verbal–visual juxtaposition that presumes the communicator's intention to convey a relevant message, and allows for

weak pragmatic implicature in those cases in which the verbal–visual correspondence is quite transparent, and for strong pragmatic implicature in cases when it is needed. Example 1 from an item on new trains to be imported from Japan illustrates both types (of pragmatic implicature). Viewers will probably easily see the co-reference between the reporter’s words and the image of the train (weak pragmatic implicature), but may need to draw a more complex inference to see the metaphorical connection representing the train arrival presumably at a station, which is associated with the concept of a train in motion, and the abstract concept of *arriving* as synonymous with the appearance of technological advancement (strong pragmatic implicature).

Example 1. ITV, 12 February 2009, on new trains to be imported from Japan.

Verbal text	Visual text
<R VO>* At last a train’s arriving. It’s the type that only comes once a generation.	A fast moving train enters the shot from right in an advertising video.

*In all examples, R VO = reporter voice-over.

Having reviewed some of the studies that were most influential to our investigation, in the following section we describe the criteria we followed in the data selection and classification, before presenting our pragmatic cognitive model illustrated with a number of examples.

DATA

The data on which our model is based consist of 22 news items drawn from ITV and BBC evening news programmes broadcast in February and March 2009.⁷ The items, averaging 2.5 minutes in length, were transcribed and segmented by scene, defined as a shot or shots of the same subject, bounded by adjacent scenes of different subjects (Graber, 1990: 90). The verbal text of each scene was displayed in a table beside a description of the accompanying images;⁸ the combination was then analysed and classified according to the model proposed in the next section.

PRAGMATIC COGNITIVE MODEL FOR THE CLASSIFICATION OF TEXT-IMAGE INTERACTION IN TELEVISION NEWS

The following model (a graphic version of which is presented in Figure 1) has been developed to attempt to capture the ongoing reciprocal relation between the verbal and the visual tracks, and the way in which ‘the words seem to be driving the pictures at the same time that the pictures seem to be driving the words’ (Montgomery, 2007: 104). The model offers two perspectives on the verbal-visual interplay: the first focuses on the relationship of the text to the images (Barthes’ anchorage), the second on the function of the

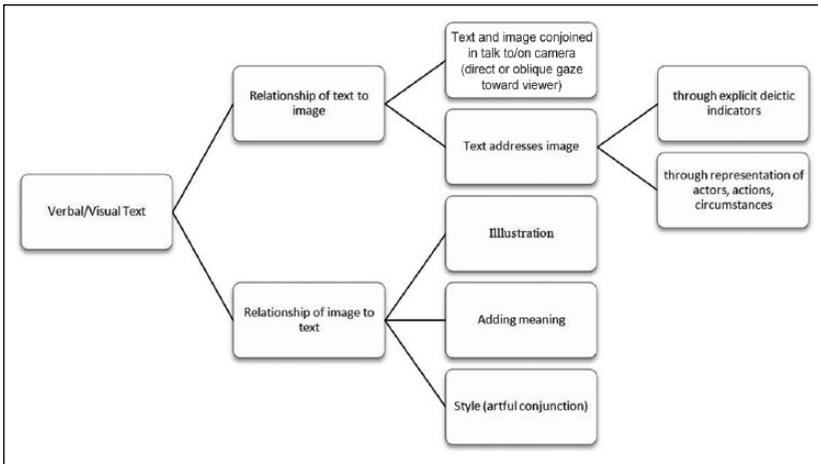


Figure 1. Model for the classification of text-image interaction in television news programmes.

images vis-à-vis the verbal track (Barthes' illustration). Together, these two perspectives allow us to more fully describe and understand the structure of the multimodal television text and to capture the function that the visuals fulfil in relation to the words. In this way, we identify the potentialities inherent in the multimodal text and suggest on what grounds the implied or idealised viewer may draw inferences and construct meaning.

RELATIONSHIP BETWEEN VERBAL TEXT AND IMAGES

Text and image conjoined in newsworker's or news actor's talk to or on camera

This category comprises the news presenter's and reporter's direct speech to camera, delivered with direct gaze to the viewers. It also includes talk by news actors (experts or public figures, but also ordinary people) delivered with an oblique gaze presumably towards the reporter on or off camera.

In the simplest cases, this involves a conventional head and shoulders shot (medium close-up) of the newsworker speaking to camera. In other cases, still or moving images or words may appear behind the news presenter in studio, e.g. on the maxiscreen during the introduction to a news report, or may be visible behind the reporter speaking to camera on location. In these cases the background may trigger processes of inference on the part of the viewers. An example may be found in BBC coverage of the G20 summit (31 March 2009, Example 2) where, in transition to a reporter stand-up on issues to be covered at the international meeting, the camera pans a council estate condominium from top to bottom to then frame the reporter, walking and speaking to camera.

Example 2. BBC, 31 March 2009.

Verbal text	Visual text
<R VO> Of course you won't have to go very far from <R to camera> the meeting to find areas where jobs have never been that easy to find, especially during a recession and some will be asking whether this summit really can deliver measures to boost employment and livelihoods.	Camera pans from oblique angle down a high rise condo to R entering centre of screen from right with medium shot, to camera.

It is the combination of words and the panning on the building that allows viewers, familiar with the architecture typical of low income areas in London, to draw the implicature that the building in the background is used to add factuality to the speaker's location. In other words, although no explicit verbal reference to the area is made, the image functions as verification of the reporter's deictic centre. More than that, however, the more significant implicit implicature viewers may draw is that the council estate in the background is relevant to the reporter's comment as an indirect reference to the location within his speech. In other words, the viewer is led to believe that the building possibly functions as a synecdoche for the poor part of London, 'where jobs have never been that easy to find' near the venue of the G20 summit, described by the reporter.

Text addresses the image with various strategies

This category comprises two sections representing different ways in which, more or less explicitly, in a more or less stylistically marked form, the verbal text addresses the images in voice-over segments over actuality film.

Through explicit exophoric deixis. As is well known, deixis is 'a category whose very purpose is to link uses of language to the context in which they occur' (Chapman, 2011: 40). Person, time, place, social and discourse deixis is a pervasive language phenomenon 'that sits right at the borderline between semantics and pragmatics'. In other words, contextual, hence pragmatic, information is needed to interpret the correct meaning of deictic expressions even in the case of a simple personal pronoun, although we have the linguistic knowledge of it as speakers of a language. The specification of 'exophoric' in our classification is meant to underline the relationship between the language and its visual representation in news; it is in this way that the deictic expressions function as marked imaginary lines or 'vectors' (Kress and Van Leeuwen, 1996) either towards the image on screen as a whole, or to specific elements therein. A moment in a BBC report illustrates the former case. In Example 3, soldiers run towards a target, shooting, the images accompanied by the reporter's words, referring to the place of the action and the action in which the soldiers are involved (the raiding).

Example 3. BBC, 18 February 2009.

Verbal text	Visual text
<R VO> British soldiers, seen here raiding a Taliban compound in Helmand Province	Soldiers in action, running, shooting

An illustration of the latter type is a BBC report on the case of Fritzl (Example 4), the Austrian man charged with incarceration and multiple abuse of his daughter. Here the reporter draws the viewers' attention to a specific moment in the trial when the man, surrounded by guards, moves his eyes left and right and up and down probably due to uneasiness, in any case in a way worthy of notice.

Example 4. BBC, 18 March 2009.

Verbal text	Visual text
<R VO> And this is the moment. <i>Look</i> at his eyes.	Fritzl moves eyes to left and right, up and down.

The reporter's words guide the viewers' perception of the images through time and person deixis respectively. Also, as Martinec and Salway (2005: 348) point out, the use of the present tense on the part of the reporter points to the action of Fritzl rotating his eyes as a sign of uneasiness and hence fulfils a deictic function of its own beyond the reference demonstrative 'this'. Moreover, in addition to the traditional deictic indicators (pronominal reference, demonstratives, verb tenses, etc.), we understand verbs or verbal phrases as direct summons to the viewers vis-à-vis the image, constructing a 'vector' (Kress and Van Leeuwen, 1996) between text and moving image on screen and highlighting to viewers the relation existing between the two. This is evident in a BBC report on teenage pregnancies in the UK, where the reporter's voice-over 'Meet baby Serena' (Example 5) is accompanied by a visual of a crawling baby in close-up. As in the early example of Fritzl, in which the imperative 'Look' appeals to the viewers asking for their direct involvement with the image, so the imperative 'Meet' is a similar directive that functions as a personal reference and deictic introduction to the actor.

Example 5. BBC, 26 February 2009.

Verbal text	Visual text
<R VO> Meet baby Serena, the latest in three generations born to teenage parents in one family.	A baby crawling on the floor with a dummy, appears in extreme close-up, gazes into the camera.

Through the representation of actors, actions/events and circumstances. In this sub-section, we include cases in which the verbal texts are accompanied by corresponding immediately identifiable images in the absence of any explicit verbal indicator depicting one or more of the following: the actors (subjects), the action/event being reported, the place or modality of its occurrence (circumstances). Some examples are:

Representation of actors. In the BBC report on teenage pregnancies in Example 5, the reporter's voice-over 'Her mother, Stephanie, was just 15 when Serena was born' coincides neatly and immediately – although without explicit deixis—with an image of the mother. A similar correspondence is evident in the following segment of the report, as may be seen in Example 6.

Example 6. BBC, 26 February 2009.

Verbal text	Visual text
<R VO> Her mother, Stephanie, was just 15 when Serena was born. Her two aunts and her grandmother were also teenage mums. Pregnancies, unplanned at some cost to the state as well as their own futures.	Close-up of woman in profile, smiling and looking downward at baby (not on screen) Three women arrive from left one after the other, and take their seat on the sofa, holding babies in their arms, joined finally by Stephanie holding Serena, who sits on arm of sofa.

Although the verbal text includes material that is not visually illustrated (e.g. pregnancies, cost to the state), the verbal-visual fit is accomplished by virtue of the continuity of the one to one correspondence between the words referring to the image and the image itself representing the principal actors.

Representation of action/event. This sub-category is exemplified in the following segment from an ITV report on the proposed increase of American troops to Afghanistan (Example 7).

Example 7. ITV, 18 February 2009.

Verbal text	Visual text
<R VO> Like everything else he has to deal with, President Obama has discovered that the situation in Afghanistan is worse than he thought.	Soldiers firing guns from behind a protective barrier, a wooded mountain partially obscured by smoke, a helicopter in air seen from below, a soldier diving into a trench directly in front of the camera, soldiers firing machine guns, soldier firing a bazooka

The verbal text is accompanied by a rapid succession of very brief shots (covering a total of 10 seconds), showing images of soldiers firing a variety of weapons in different circumstances and at different distances ranging from

close-up to long shots. All are visual representations of the ‘event’ mentioned in the voice-over: the situation in Afghanistan (warfare). The insistent sound of shots being fired which is heard over the compelling visuals underlines the gravity of the situation (‘worse than he thought’).

Representation of contextual circumstances. The BBC report of Dutch parliamentarian Geert Wilders’ unsuccessful attempt to enter Britain to show his anti-Muslim film is a good illustration of this sub-category, the image representing the location referred to in the reporter’s voice-over (Example 8).

Example 8. BBC, 12 February 2009.

Verbal text	Visual text
<R VO> It was, of course, all extraordinary publicity for a film show being held this afternoon in the House of Lords.	A ‘postcard’ image of the historical building appears in long shot.

RELATIONSHIP OF IMAGES TO TEXT: FUNCTIONS OF THE IMAGES VIS-À-VIS THE VERBAL TRACK

In this second part of the model, we take a different perspective starting from the images and focusing on the meaning constructed by them in conjunction with the verbal text. Furthermore, as we do this we also consider factors affecting the relationship between visual and verbal tracks and the viewer’s comprehension process. In other words, here we pursue a more cognitive pragmatic approach essentially in line with Montgomery (2007).

Illustration

While the previous examples provide neat synchronization between the verbal and visual tracks, in many other cases the correspondence between the two planes is not as straightforward. While still addressing the topic being reported, in many instances the visuals do not transparently represent the information conveyed by the words. As Montgomery (2007: 97) points out, ‘identifying a visual correlate of a verbal referring expression is the outcome of sometimes complex and cumulative inferential work’, as in Example 9 from the BBC item on teenage pregnancies.

Example 9. BBC, 26 February 2009.

Verbal text	Visual text
<R VO> Today’s figures show that teenage pregnancies are now rising for the first time in five years, fuelling the debate about [what’s gone wrong].	Man shown from back of head looking down at and playing with baby, who is on floor, another close-up of baby to camera with dummy, reaching out (to camera).

One possible visual realisation of the information conveyed by the reporter's text would have been a graph indicating the increase in the number of British teenagers having babies in 2009. The images instead portray an adult playing with a baby that, in line with information contained in the report to that point, viewers are inferentially led to interpret as a baby born of a teenage parent. In consideration of the verbal information, the visuals provide an extension of the meaning as they reflect on the possible interaction between a presumably teenage parent and a child, rather than focusing on the factual increase in pregnancies registered in the country. At the basis of the meaning-making operation is the expectation of relevance, that is the promise, in this case, that the baby in the report is interesting and worthy of our attention because it is related to the topic in hand. Put another way, we treat this image as combined with the previous visuals of teenage parents and with the words being uttered by the reporter. Therefore we treat the whole ensemble as a message that, according to Sperber and Wilson (1995[1986]: 158), 'communicates a presumption of its own optimal relevance'. Such 'ostensive-inferential communication' conveys a 'stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions' (p. 63).

As discussed earlier, viewers process the ensemble of sounds, words and images contained in television news in exactly the same way in which they make sense of any text. While not engaging with the issue of explicit communication and intentionality in news editing, Montgomery (2007: 97) explains the 'supposition of co-reference between the verbal and visual tracks' with the Principle of Intelligibility, whereby, as indicated above, the viewer: (a) expects to find a correspondence between any verbal referring expression and its relevant illustration on the visual plane; and (b) vice versa treats any image as potentially referring to the accompanying words. While we cannot argue beyond doubt in favour of the deliberate nature of a specific news item, we assume that some degree of explicitness and intentionality is present in news bulletins and, above all and most importantly, that, like in any communication, this is at the basis of the pragmatic inference process that is activated in the viewer's mind.

We see the nature of the visual-verbal co-reference as best understood in terms of a gradient of more or less transparent overlap. This follows Montgomery for what concerns the specific area of television news and more broadly Blommaert's (2007: 15) concept of scalarity that applies to all socio-linguistic phenomena, which 'cannot be captured in horizontal metaphors of spread and distribution; they need to be imagined as stratified, layered and unequal phenomena that reveal systemic features of (unequal) social structure'. Example 10 illustrates a much looser connection between the two tracks which imposes a greater onus on the viewers in terms of information processing. This ITV report in relation to the opening of the G20 summit in London

focuses on the possible disagreement and division between world leaders regarding the response to the global financial crisis. Covering Brown's speech at St Paul's Cathedral, the report includes excerpts of his talk interspersed with voice-over comments by the reporter.

Example 10. ITV, 31 March 2009.

Verbal text	Visual text
1.38 <Brown> And our task today is to bring financial markets into closer alignment with the values held by families and business people across the country.	Medium shot of Brown speaking and smiling/looking at audience on 'families'.
1.49 <R VO> He seemed to acknowledge, though, there was a real danger.	Long shot of Brown speaking on right of majestic hall with massive lit-up chandeliers. Audience becomes visible.
1.51 <R VO> Many might view this summit as crushingly irrelevant.	Medium shot on Brown's wife in the audience very attentively listening to him, looking concerned.
1.56 <Brown> And I recognise that for too many families, anxious about jobs...	Medium shot of Brown speaking

Our focus is on the third row in Example 10 where the co-reference between the reporter's comment and the medium shot of Brown's wife, the function of which is not immediately transparent, triggers a number of possible pragmatic inferences on the part of the viewers. What sense can an idealised or implied viewer make of this particular combination? Is the wife concerned about the irrelevance of the summit and her husband's speech? Is attention being called to her place in the audience among the vast public and, if so, is it because she doesn't have a privileged place? Does her presence as Brown's wife tie in with his previous reference to families? And since her eyes then turn (surreptitiously?) to the left, possibly to the person sitting next to her, is she checking on the general level of attention or the audience's reaction to her husband's words? As in the previous example, the visuals extend the verbal meaning and construct an independent thread that originates from Brown's wife. However, if compared with the previous examples, the case described here requires a much more demanding cognitive process and yields uncertain results. More importantly, if positioned along a gradient, this case is definitely collocated at the far end of the continuum between a close and tenuous fit according to the notion of scalarity of sociolinguistic phenomena mentioned earlier (Blommaert, 2007).

Further still along the continuum are cases of what appears to be a total lack of fit between words and images, cases which in our own data of 22 items are so rare as to be considered editing anomalies (see Examples 11 and 12):

Example 11. ITV, 31 March 2009.

Verbal text	Visual text
<R VO> Let's hope it's all worth it. Tom Bradby, ITV News, at <i>Westminster</i> .	Door at 10 Downing Street with two people peeping out of it.

Example 12. ITV, 18 February 2009.

Verbal text	Visual text
<R VO> Faced with such aggression, the Americans think there's no alternative to sending in extra forces.	American flags, people cheering American soldiers in white uniforms at an unspecified celebration in a stadium. Dignitaries on stage.

From a pragmatic Relevance Theory perspective, the issue of such missing co-references or, as Meinhof (1994) terms them, dichotomies, is worthy of attention. Even if these cases can be explained in terms of the haste to which news editing is frequently subjected, the problem still remains of what viewers make of such seemingly incoherent verbal–visual relations. It is realistic to assume, with Montgomery, that viewers take anything that appears on screen as informative and therefore establish some degree of reference between the two tracks even though the relevance between them is highly implicit or apparently missing. Therefore it is plausible to assume that, in Example 11, Downing Street is taken as a synecdoche for the whole government that we know resides at Westminster. In Example 12, the equally vague reference can be taken to indicate a generalised illustration of official military function that goes hand in hand with the notion of US troops to Afghanistan mentioned by John Irvine.

At the same time, clearly, in these seemingly inconsistent cases, the images can be interpreted as representing at least (most transparently) actors, concrete events or components of events or circumstances regarding them in accordance with the relation of trust between news producers and audience, and the assumption that news is in any case and in any way informative and newsworthy. Moreover, viewers treat news as 'ostensive communication' and draw pragmatic inferences to determine explicit as well as implicit meanings, and, finally, in light of Montgomery's Principles of Intelligibility, viewers assume a degree of relevance in the combination of verbal and visual tracks. As a further assumption on which news processing is based, however, viewers do not and cannot expect the visuals to display things that cannot be figuratively portrayed, e.g. mental processes, future projections, hypothetical constructions, or asides and comments by news workers (as in the examples above and Example 13).

Example 13. BBC, 16 February 2009.

Verbal text	Visual text
<R VO> Those who were let go today hope that their jobs will come back with better days.	Sign indicating Oxford Mini factory, images of cars driving left to right.

While this may sound like a limitation to the decoding process, it is instead an important rule that viewers bear in mind. News is after all an account of events that have occurred in another time and another place, presented and described as if they were still real, by someone who has authority and credibility. However, news is talk accompanied by images even though the talk may be constructed upon and led by the images. Viewers are aware of that and know that news talk involves comments, reflections, projections, evaluations and the like, all things that cannot easily be visually represented especially not at the fast pace in which news programmes are constructed and delivered. Therefore, if we can talk about the Principle of Intelligibility (Montgomery, 2007) or the more general Principle of Relevance (Sperber and Wilson, 1995), we must also assume a Principle of Non-Representability according to which abstract and mental processes cannot be represented visually. The principle of Non-Representability also applies to another important dimension, that of the viewers' expected *partial* representation of information, according to which only part of the news item receives visual attention. An illustration is Example 14 which we can classify as a representation of event, in that the act of flying is figuratively present but not the projected criticism of world leaders at the G20 summit.

Example 14. ITV, 31 March 2009.

Verbal text	Visual text
<R VO> but as he left the US, President Obama was bound for a summit where it seems few other leaders will back his approach.	Plane taking off.

It would appear therefore that viewers are predisposed to accept a degree of lack of coherence in the news. Higgins (1991) has identified the ability of news recipients to withstand and interpret as normal and expected the fact that 'while each single news item should have an internal coherence and unity, the broadcast as a whole has no apparent coherence' (p. 150) due to the recurring topic shifts between the reported items. We suggest that this ability applies as well to the single news item, and that in addition to an expectation of general relevance and intelligibility and to the process of implicature generation, news viewers have an expectation of lack of representability and coherence.

Adding meaning

So far we have discussed cases of more or less close co-reference between the words and the images; however, undeniably, the visual track preserves a degree of autonomy inherent in its medium. Apart from the discussed cases of presumably low accuracy in the editing in which images send a message independent of the words, images speak for themselves and viewers have leeway in interpreting visuals, so it is not implausible to conceive that footage and still pictures may suggest unexpected or unwanted meaning. We make clear once again that we are interested in the comprehension issue while the modalities of news production are beyond the scope of this article. In some cases in which the images speak more than the words, although the latter are still instigated by them, at the conceptual/ideational level in which the information is encoded, we are guided in our sense-making by the usual combination of words and images. However, at another level, the images may convey other meanings that plausibly the words had not completely catered for. A revealing example is offered by the ITV report on the increase of American troops sent to Afghanistan (18 February 2009) in Example 15.

Example 15. ITV, 18 February 2009.

Verbal text	Visual text
<R VO> The Americans, the British and their allies were sent there to take the fight to the Taliban, but now it's the other way round.	Soldiers ascending rocky slope to dwellings, moving away from camera, backs to camera, then walking

In this report, the visuals that accompany the verbal text represent soldiers moving away from the viewer's deictic centre, showing their backs to the viewers, and hence suggesting the act of distancing themselves or getting away, and that movement may trigger in the viewers' mind the idea that they are retreating from the enemy they were supposed to attack. This of course is an elaboration on the visual text that is encouraged by the lack of frontality of the soldiers, who seem to be exiting the shot through the upper right corner of the screen.

It is plausible that a pragmatic inference may be drawn by viewers leading to an extended or additional meaning. Such an interpretation that the visuals permit is very probably not explicitly or intentionally communicated but made possible by elements on the interpersonal level, e.g. the view of the soldiers' backs and their distancing from the audience; also, the support of one small linguistic trigger, 'the other way round', allows the viewer's elaboration of the message as negative.

Style: artful conjunction of verbal and visual tracks

The above discussion has taken into account the degrees of explicit and implicit relevance that words establish in their co-reference to visuals in news reports. Another case of meaning-construction worth noting is when the verbal text engages in a stylistic relation with the visuals constructing metaphors and word play. This can occur both within a single scene and between scenes, and contributes to the dynamic development of images in a report. An illustration of the former case appears in a BBC item on G20 in which the reporter's verbal track covers images of police on motorbikes manoeuvring and departing from a parking area completed by a quick glimpse of a Mini driving by (Example 16).

Example 16. BBC, 31 March 2009.

Verbal text	Visual text
<R VO> Security preparations are moving into overdrive and, with the police operation gathering momentum so too the last-minute work to try to ensure ...	Police on motorcycles, motorised police in <i>motion</i> , shot of a Mini driving by.

Within the single scene, the verbal text maps a correspondence between constituent elements of a cognitive metaphor in which the target domain is the semantic domain structured in such a way as to understand the source domain (Lakoff and Johnson, 1980). Here we are dealing with a 'specific rhetorical strategy' whereby 'the blending of visual images and verbal text in a metaphoric relationship is intended to influence the viewers' comprehension (Wiggin and Miller, 2003: 269) and helps them to 'think abstractly' (Messaris, 1994, in Wiggin and Miller, 2003: 270) by suggesting a figurative relationship. Forceville (1996) revealed how verbal and visual texts can be combined to create metaphorical meaning. In other words, 'preparations moving into overdrive' and 'police operation gathering momentum' is what viewers have to understand and for which the scene creates another conceptual domain figuratively as an illustration. Thus 'overdrive' and 'momentum', synonymous with dynamism, are made accessible in the mind of the viewer by a series of vehicles, motorbikes and a car that suggest the speed usually associated with movement.

In place of a one-to-one reference between transparent words and images, here we have an elaborate word play that is similar in conceptual complexity to the case of irony, in which the wit contained in an ironic utterance produces a failure of interpretation that forces the reader to reanalyse the text and finally identify the writer's intention (Furlong, 2011). The reader/listener is rewarded for the processing labour with the pleasure that witty irony offers. Like a case of irony, the 'overdrive' metaphor is rewarding for the stylistic

pleasure it brings to the viewers. From a conceptual metaphor theory (Lakoff and Johnson, 1980) therefore, the target domain being constructed visually in the viewers' minds to understand the verbally expressed source domain lightens the cognitive load that stylistically marked language requires.

In another ITV item on Brown's first visit to the newly elected Obama (Example 17), the word play is constructed on the figuratively new appearance that Brown is trying to achieve, hence political 'makeover', and the actual cosmetic preparation he undergoes before disembarking from the plane.

Example 17. ITV, 3 March 2009.

Verbal text	Visual text
<R VO> The new face of Gordon Brown, a Prime Minister in search of a makeover. He has to look his best here and it was with not a hair out of place [that he made his way down the steps onto American soil].	Brown on plane, profile visible through the airplane window as if in a cameo, a hand applying make-up to his face and combing hair.

As in the earlier case, therefore, the visuals function to explicate the pragmatic inference that determines this explicitly and intentionally formulated ostensive communication. In a way, images guide the viewers' interpretation and their comprehension process, making the appreciation of the metaphor on the one hand more immediate because more accessible, while on the other adding an additional level to the figurative meaning.

In the case of adjacent scenes, this kind of artful conjunction operates to ensure thematic continuity throughout the report. An example is the ITV report on Obama's review of military strategy in Afghanistan (Example 18).

Example 18. ITV, 18 February 2009.

Verbal text	Visual text
1.27 <R VO> President Obama had wanted to conclude an extensive Afghan strategy review before making any decisions on troops levels	Obama shaking hands with people after descending steps of airplane.
1.37 <R VO> but his hand has been forced by the increasing violence	Helicopter mid close-up moving on ground, another helicopter behind it takes off.
1.42 <R VO> and so 17,000 extra combat soldiers will be deployed.	Inside helicopter in flight, soldiers before open door.

Throughout the first shot, Obama, who has just disembarked from an airplane, is seen greeting well-wishers, shaking hands repeatedly. At the end of the shot we see a mid-shot of a helicopter moving on the ground and another

taking off, as the reporter says: But his 'hand has been forced' by the increasing violence. In this particular case, the word play establishes a connection between the physical act of shaking hands, not mentioned in the verbal text, and the abstract concept 'forcing his hand' pronounced in the following scene which, although in a different context (Obama being pushed into a decision), refers back to the previous one.

CONCLUSION

The attractiveness of a study of news reports that takes into account the multiple aspects of the messages conveyed to the readers is undeniable. The numerous studies on the topic attest to the willingness of scholars to track down the complex comprehension process that the ensemble of moving images, sounds and words trigger in the viewers' mind. Following on from Montgomery's (2007) attention to the inferential processes that watching televised news involves, in this article we have shown the many combinations and points of contact between the visual and the verbal, and how their interaction suggests a range of possible meanings to the viewers. In so doing, we have investigated a number of excerpts from news programmes and shown how the 'majority of the inferences drawn from a text are the result of combining textual elements with themselves and/or with contextual elements' (Shiro, 2005: 167). We have also pointed to what in news messages is explicitly communicated and what is the implicit and plausibly inferable meaning. We also have suggested that the type of inference drawn on the basis of a specific piece of textual information derives from the function and role of that information in the text. In the case of the building behind the reporter in Example 2, for instance, it is the complex role that the image plays in relation to the journalist's mention of the high unemployment area of London that makes it possible for the viewers to interpret the dwellings in the background as an indication that the reporter is speaking from deprived southeast London. Therefore, the interaction between the verbal and visual in this case triggers a cognitive process that moves from the explicit to the implicit.

Comprehension is based on inferring and providing the inexplicit links between textual elements. As analysts we are aware that the possible meanings we read in the texts we analysed were based on our own inferences and generalised to an average viewer with a degree of familiarity with TV discourse. Obviously, as we have made clear throughout this article, other possible meanings can arise from a viewer's interpretation of any verbal-visual combination in a multimodal text.

Drawing on the possible textual meanings of news reports and the seminal work by Montgomery (2007), and based on a consideration of the elements present in the texts (while deliberately abstaining from considerations regarding journalistic practice and intention), this article has proposed a model for the classification of verbal-visual interplay that aims to capture the

reciprocal relation between the two tracks and contribute to an understanding of verbal–visual interaction in TV news. The model addresses such interaction from two perspectives. The first consists of the description of the relationship in two types of cases: when text and image are conjoined, and when the verbal text is in an explicit and direct relation with the images. The second perspective is concerned with the function of the verbal–visual relationship and considers cases in which the images elaborate on and enrich the verbal meaning, by illustrating, by adding meaning, or stylistically through the artful conjunction of verbal and visual texts. Our model is dynamic in that it considers the images in motion and takes into account the links between separate scenes. It also considers the fact that meanings are primarily social and based on the knowledge of the world viewers have; it proposes a range of inferential interpretations that possibly become available to TV news addressees without forgetting one of the difficulties analysts face, namely that ‘inferences are elusive because once they have been drawn they do not appear to be inferences any more’ (Shiro, 2005: 167).

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

NOTES

1. For a detailed review of the literature on the topic, see Martinec and Salway (2005).
2. Martinec (2013) further develops this system and discusses its application to new media.
3. In this regard, Martinec and Salway (2005: 342), commenting on Barthes’ remarks on relay-text in film note that ‘Barthes’ text advancing the action by setting out new meanings sounds very much like the logico-semantic relation of extension and perhaps also enhancement (see Halliday, 1994: 120–129)’.
4. Especially for the concepts of ‘elaboration’ and ‘extension’, which he sees as similar to Barthes’ (1977) ‘anchorage’ and ‘relay’.
5. Skilled journalists assemble their report on the basis of the available images, according to Maria Pia Farinella, Rai TV journalist (personal communication, and BBC manual). Note also Altheide (1987) in Griffin (1992) who suggests that the news depends on what images are accessible and can be shown.
6. Viewers are actively involved in the process of meaning-making. With reference to a literary context, see McIntyre and Jeffries (2010: 127) and Semino (1997: 125), who argues that comprehension involves projection and construction, in the sense that a text projects a potential meaning while readers construct it. The agency that these two authors

recognise in literary contexts, we believe, also applies to the substantially different case of viewers of television news reports.

7. The programmes were recorded in connection with the IntUne project (Integrated and United: A quest for citizenship in an ever closer Europe) financed by the Sixth Framework Programme of the European Union, Priority 7, Citizens and Governance in a Knowledge Based Society (CIT3-CT-2005-513421), in which both authors participated.
8. Although this (typo)graphical convention may suggest an implicit assumption that the visual data are secondary to the verbal text, our analysis is based on the assumption that the two tracks are simultaneously decoded by the viewer. Our thanks to an anonymous reviewer for this observation.

REFERENCES

- Andén-Papadopoulos, K. (2008) The Abu Ghraib torture photographs: News frames visual culture and the power of images. *Journalism* 9(1): 5–30.
- Andrews, D. (2007) Softcore as serialized (and feminized) featurette. *Television and New Media* 8(4): 312–340.
- Altheide, D. (1987) The format of TV network news. In: Vidal-Beneyto, J. and Dahlgren, P. (eds) *The Focused Screen*. Strasbourg: AMEDLA/Council of Europe.
- Barthes, R. (1977) *Image, Music, Text*. London: Fontana.
- Bednarek, M. and Caple, H. (2012) *News Discourse*. London: Continuum.
- Blommaert, J. (2007) Sociolinguistic scales. *Intercultural Pragmatics* 4(1): 1–19.
- Caple, H. (2008) Intermodal relations in image-nuclear news stories. In: Unsworth, L. (ed.) *Multimodal Semiotics: Functional Analysis in Contexts of Education*. London: Continuum, 125–138.
- Caple, H. (2010) What you see and what you get: The evolving role of news photographs in an Australian broadsheet. In: Rupar, V. (ed.) *Journalism and Meaning-Making: Reading the Newspaper*. Cresskill, NJ: Hampton Press, 111–133.
- Chapman, S. (2011) *Pragmatics*. Basingstoke: Palgrave Macmillan.
- Dahlgren, P. (1988) What's the meaning of this? Viewers' plural sense-making of TV news. *Media, Culture & Society* 10: 285–301.
- Fiske, J. (1987) *Television Culture*. London: Routledge.
- Forceville, C. (1996) *Pictorial Metaphors in Advertising*. New York: Routledge.
- Furlong, A. (2011) The soul of wit: A relevance theoretic discussion. *Language and Literature* 20(2): 136–150.
- Graber, D.A. (1990) Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication* 40(3): 134–155.
- Grice, P. (1975) Logic and conversation. In: Cole, P. and Morgan, J. (eds) *Syntax and Semantics Volume 3: Speech Acts*. New York: Academic Press, 41–58.

- Griffin, M. (1992) Looking at TV news: Strategies for research. *Communication* 13: 121–141.
- Griffin, M. (2004) Picturing America's 'War on Terrorism' in Afghanistan and Iraq: Photographic motifs as news frames. *Journalism* 5: 381–402.
- Halliday, M.A.K. (1985) *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, M.A.K. (1994) *An Introduction to Functional Grammar*, 2nd edn. London: Arnold.
- Halliday, M.A.K. and Hasan, R. (1976) *Cohesion in English*. London: Longman.
- Hartley, J. (1992) *The Politics of Pictures: The Creation of the Public in the Age of Popular Media*. London: Routledge.
- Higgins, C. (1991) Broadcast news: A linguistic mode of analysis. *Continuum* 5(1): 149–165.
- Jayyusi, L. (1991) The equivocal text and the objective world: An ethnomethodological analysis of a new report. *Continuum* 5(1): 166–190.
- Juluri, V. (2002) Music television and the invention of youth culture in India. *Television and New Media* 3(4): 367–386.
- Kress, G. and Van Leeuwen, T. (1996) *Reading Images*. London: Routledge.
- Lakoff, G. and Johnson, M. (1980) *Metaphors We Live by*. Chicago: University of Chicago Press.
- Lemke, J. (1998) Metamedia literacy: Transforming meanings and media. In: Reinking, D., et al. (eds) *Literacy for the 21st Century: Technological Transformation in a Post-Typographic World*. Hillsdale, NJ: Erlbaum, 283–301.
- Martinec, R. (2013) Nascent and mature uses of a semiotic system: The case of image–text relations. *Visual Communication* 12(2): 147–172.
- Martinec, R. and Salway, A. (2005) A system for image–text relations in new (and old) media. *Visual Communication* 4(3): 337–371.
- McIntyre, D. and Jeffries, L. (2010) *Stylistics*. Cambridge: Cambridge University Press.
- Meinhof, U. (1994) Double encoding in news broadcasts. In: Graddol, D. and Boyd-Barrett, O. (eds) *Media Text: Authors and Readers*. Clevedon: Multilingual Matters, 212–223.
- Meinhof, U. (1998) *Language Learning in the Age of Satellite Television*. Oxford: Oxford University Press.
- Messaris, P. (1994) *Visual Literacy: Image Mind and Reality*. Boulder, CO: Westview Press.
- Montgomery, M. (2007) *The Discourse of Broadcast News*. London: Routledge.
- Nichols, W. (1976) Documentary theory and practice. *Screen* 17(4): 34–48.
- Oguss, G. (2005) 'Whose barrio is it?' Chico and the man and the integrated ghetto shows of the 1970s. *Television and New Media* 6(1): 3–21.
- O'Halloran, K. (1999) Towards a systemic functional analysis of multisemiotic mathematics texts. *Semiotica* 124(1/2): 1–29.

- O'Toole, M. (1994) *The Language of Displayed Art*. London: Leicester University Press.
- Royce, T. (1998) Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text *JASFL Occasional Papers* 1(1): 25–50.
- Semino, E. (1997) *Language and World Creation in Poems and Other Texts*. London: Longman.
- Shiro, M. (2005) Inferences in discourse comprehension. In: Coulthard, M. (ed.) *Advances in Written Text Analysis*. London: Routledge, 167–178.
- Sperber, D. and Wilson, D. [1986] (1995) *Relevance*. Oxford: Blackwell.
- Thompson, J.B. (1984) *Studies in the Theory of Ideology*. Cambridge: Polity Press.
- Thornborrow, J. and Haarman, L. (2012) Backstage activities as frontstage news. *European Journal of Communication* 27(4): 376–394.
- Thornborrow, J., Haarman, L. and Duguid, A. (2012) Discourses of European identity in British, Italian, and French TV news. In: Bayley, P. and Williams, G. (eds) *European Identity: What the Media Say*. Oxford: Oxford University Press, 84–117.
- Tolson, A. (1996) *Mediations: Text and Discourse in Media Studies*. London: Arnold.
- Van Leeuwen, T. (1991) Conjunctive structure in documentary film and television. *Continuum* 5(1): 76–114.
- Wiggin, A. and Miller, C. (2003) 'Uncle Sam Wants You!': Exploring verbal-visual juxtapositions in television advertising. In: Scott, L. and Batra, R. (eds) *Persuasive Imagery*. London: Routledge, 267–295.
- Wojcieszak, M. (2009) Three dimensionality: Taxonomy of iconic linguistic and audio messages in television news. *Television and New Media* 1(6): 459–481.
- Zelizer, B. (2002) Journalism photography and trauma. In: Zelizer, B. and Allan, S. (eds) *Journalism after September 11*. New York: Routledge, 48–68.
- Zelizer, B. (2004) When war is reduced to a photograph. In: Allan, S. and Zelizer, B. (eds) *Reporting War: Journalism in Wartime*. Milton Park: Routledge, 115–135.

BIOGRAPHICAL NOTES

ROBERTA PIAZZA is Senior Lecturer in English Language and Linguistics at the University of Sussex. Her research interests range from anthropological linguistic studies of identity, as she explores in *Marked Identities* (Palgrave Macmillan, 2014, co-edited with Alessandra Fasulo), to pragmatics and media discourse, as she investigates in *The Discourse of Italian Cinema and Beyond* (Continuum, 2011), *Telecinematic Discourse* (John Benjamins, 2011, co-edited with Monika Bednarek and Fabio Rossi) and *Values and Choices in Television Discourse* (Palgrave Macmillan, 2015, co-edited with Louann Haarman and Anne Caborn). Another article with Louann Haarman 'Toward a definition

and classification of human interest narratives in television war reporting' appeared in the *Journal of Pragmatics* (2011).

Address: University of Sussex, Falmer, Brighton BN1 9QN, UK. [email: r.piazza@sussex.ac.uk]

LOUANN HAARMAN is Alma Mater Professor of English Language and Linguistics at the University of Bologna, Italy (retired). Her principal research interest is in the field of discourse analysis with particular reference to media discourse and television news. Publications include *Evaluation and Stance in War News* (Continuum, 2009, with Linda Lombardo), a cross-cultural comparative study of British, American and Italian coverage of the Iraq war, and 'Backstage activities as frontstage news' (*European Journal of Communication*, 2012, with Joanna Thornborrow) dealing with features of media self-referentiality and artfulness in visual and verbal synchrony as resources in television news reporting. Most recently she has co-edited (with Roberta Piazza and Anne Caborn) *Values and Choices in Television Discourse* (Palgrave Macmillan, 2015).

Email: louann.haarman@gmail.com