Introduction: Multimodality as challenge and resource for translation
Carol O’Sullivan, University of Portsmouth

Translation is usually thought of as being about the printed word, but in today’s multimodal environment translators must take account of other signifying elements too. Words may interact with still and moving images, diagrams, music, typography or page layout. Multimodal meaning-making is deployed for promotional, political, expressive and informative purposes which must be understood and accounted for by technical translators, literary translators, copywriters, subtitlers, localisers, publishers and other professionals working with language and text.

The original title of the conference from which some of the contributions to this special issue emerged borrowed the title *Image, Music, Text* from Stephen Heath’s 1977 compilation and translation of essays by the semiotician and literary theorist Roland Barthes. It seemed to us as organisers that the various relations between image, music and text would constitute sites where interesting translational challenges and solutions might arise. The contributions which follow demonstrate indeed that this is the case, but also that multimodal challenges and resources for translation constitute a much vaster and richer field than this initial tripartite ‘tag’ would suggest.

1. Multimodality as a new/old phenomenon

Many scholars have remarked on an increase in textual multimodality with the rise of information technology and the web. Gunther Kress and Theo van Leeuwen, for instance, argued in 2001 that prestige forms of communication in Western culture were in the past conspicuously monomodal. They suggested that

> the most highly valued genres of writing (literary novels, academic treatises, official documents and reports, etc.) came entirely without illustration, and had graphically uniform, dense pages of print. Paintings nearly all used the same support (canvas) and the same medium (oils), whatever their style or subject. In concert performances all musicians dressed identically and only conductor and soloists were allowed a modicum of bodily expression. The specialised theoretical and critical disciplines which developed to speak of these arts became equally monomodal [...]. More recently, this dominance of monomodality has begun to reverse [...] (Kress and van Leeuwen 2001: 1).

This is at least in part an issue of perception: we think of different modes as discrete and clearly bounded as a function of the critical and analytical vocabulary we have to describe them. Thus translation, whose theory remained until recently almost exclusively word and script-based, is generally conceived as the rendering of written text into written text; the
particular resources used to write the text, and the other semiotic modes used to construct meaning around the text, have been all but ignored.

Lacking a critical vocabulary with which to engage with multimodal texts, we may be less likely to be aware of them, but as Eija Ventola and Martin Kältenbacher observe, although multimodality has long been ignored by scholars with an interest in reinforcing the boundaries of disciplines and research fields, it has “been omnipresent in most of the communicative contexts in which humans engage” (2004: 1). Medieval manuscripts, for instance, are often highly multimodal, as Jones (2013) acknowledges, including calligraphic and illustrative elements. Sara Oberg Stradal (2012) shows how the Practica Chirurgiae by the fourteenth-century physician John of Arderne included marginal illustrations with a variety of functions. Illustrations included visual/verbal plays on words in the body of the text. Stradal theorises, for instance, that the owl [Latin: bubo] drawn in the margin beside the term for a cancerous growth or boil [in Latin, also bubo] would have helped medieval readers to navigate the specialised terminology in the text. Should the text then be translated from Latin into the vernacular, this verbal/visual punning would constitute a challenge for translators.

Even after Gutenberg’s introduction of the printing press, the rectangular block of text that characterised the printed page had its own dimension of material signification, as Martin Janssen (2010) has shown. The space available for the text also placed its own constraints on the translator. In her seminal essay “Towards a media history of translation” Karin Littau discusses the notable case of the Nuremberg Chronicle, a history of the world from the Book of Genesis to the present, published both in Latin (as the Liber chronicarum) and in German (as Die Schedelsche Weltchronik) in 1493. The German translation by Georg Alt was widely criticised for its infidelity to the text, as it strayed from the principle of close translation, cutting out much of the detail and the scholarly apparatus from the text (Littau 2011: 270–271). Littau shows, however, that much of the omission in the German translation can be ascribed to the need to preserve the visual layout of the text, which is heavily illustrated. Because “German requires more words to express the Latin content” the German text had to be more heavily edited to fit the space. She concludes that “Alt translates text and image in tandem” (2011: 271). The writing-based theory used to evaluate and criticise the translation failed to take layout and images into account².

2. Typography

It is not just about the conspicuous ‘accompanying’ or juxtaposed semiotic elements such as illustrations. Theo van Leeuwen’s recent work on typography (2005, 2006) argues that typographical elements constitute a further modality for what have conventionally been considered monomodal print texts. He quotes Bellantoni and Woolman’s distinction
between the ‘word image’ and the ‘typographic image’ (van Leeuwen 2006: 142). The argument could be summarised as “font also signifies.” This meaning-making potential of typefaces was recognised by Beatrice Warde as long ago as 1955, in an essay which explicitly compares printing to translation. Warde argues that “[t]he most important thing about printing is that it conveys thought, ideas, images, from one mind to other minds. [...] Type well used is invisible as type, just as the perfect talking voice is the unnoticed vehicle for the translation of words, ideas” (1955: 13). Translation has often been theorised via the conduit metaphor, and in her essay Warde makes clear that she also sees typography as a conduit, by which “the mental eye focuses through type and not upon it” (1955: 16). Type must not call attention to itself, but (like translation) it must provide perfect access to the content of the text while itself remaining invisible:

The book typographer has the job of erecting a window between the reader inside the room and that landscape which is the author’s words. He may put up a stained-glass window of marvellous beauty, but a failure as a window; that is, he may use some rich superb type like text gothic that is something to be looked at, not through. Or he may work in what I call transparent or invisible typography (Warde 1955: 15–16).

Warde’s words echo Lawrence Venuti’s discussion of the strategy of fluency (1995) as a way to de-emphasise the text’s translated status. To date, typeface choice and other printing decisions have not been taken much into account by scholars of translation. This is perhaps a pity, because typography has, at different times and in different media, been quite active as a translation issue. The issue of the ideal typeface for subtitles, for example, has long been discussed in the industry; sans serif fonts are generally preferred as being easier to read (Díaz Cintas and Remael 2007: 84). Media consumers may have strong opinions about the adequacy of particular typefaces, colours, and so on; some fans draw a distinction, for instance, between the ‘ugliness’ of player-generated subtitles for films on DVD and the elegance of laser-engraved theatrical subtitles. Liu (2011: 209) notes that fonts for bilingual English and Chinese texts in in-flight magazines have to be carefully considered because for linguistic reasons the same block of text frequently has to accommodate both Chinese characters and English letters, with consequences for kerning, font size and so on.

In sixteenth-century Germany, the question of the best font with which to translate was also at issue, as part of the debate about the relative merits and usage of gothic and roman typefaces. Georg Rörer, who supervised the printing of the Wittenberg editions of Martin Luther’s translation of the Bible, developed typographical ‘aids’ for the reader in the form of roman typeface in certain words. The idea (and we must remember the Reformation context in which this was taking place) was that roman typeface was used for negatively connoted words, while positively connoted words were presented exclusively in gothic (Flood 1993: 133–
135). George Flood hypothesises (*ibid.*) that this typographical device was part of Protestant anti-papal propaganda which sought to link certain Biblical elements with the Church of Rome; so for instance in Revelations 17, the ‘Whore of Babylon’ is presented in the text using roman type. Here typography is an important meaning-making resource for ideological shifts in translation.

3. New media, new tools

The perennial existence of multimodality notwithstanding, it is true that with the rise of new media, in particular the world wide web and multimedia forms of communication and entertainment, the multimodality of texts has become increasingly conspicuous. N. Katherine Hayles observes of experimental electronic literature that:

> screen design, graphics, multiple layers, color, animation, etc. are signifying components essential to the work’s effects. Focusing only on “the actual order of words and punctuation” would be as inadequate as insisting that painting consists only of shapes and ruling out of bounds color, texture, composition, perspective, etc. (2003: 267).

These remarks are also valid for pragmatic texts, whether they be bilingual in-flight magazines (Liu 2011), webpages (Rike, in this issue), women’s monthly magazines (Chueasuai, in this issue), board games (Evans, in this issue) or term banks (López Rodríguez *et al.* in this issue).

Indeed one of the difficulties of working on multimodality may be the problem of categorisation. Translation Studies has struggled at times with the concept of multimodality. One example of this is Katharina Reiss’s initial classification of the ‘audio-medial’ function of language as supplementary to the informative, operative and expressive functions — a position she later modified (see Reiss 2000: 164–165), arguing that multi-medial texts must be considered a ‘hyper-type’ which could, in turn, be informative, operative and/or expressive in function. This revisiting of her approach speaks to the difficulty of incorporating multimodality into theories of translation which had until then been exclusively text-based.

Mary Snell-Hornby has suggested that we can define four different genres of multimodal text (2009: 44, some emphasis added):

1. **multimedial** texts (in English usually called audiovisual, but not to be confused with “multimedia” in its loose everyday usage) are conveyed by technical and/or electronic *media* involving both sight and sound (e.g. material for film or television, sub-/surtitling);
2. **multimodal** texts involve different *modes* of verbal and nonverbal expression, comprising both sight and sound, as in drama and opera;
3. **multisemiotic** texts use different *graphic sign systems*, verbal and nonverbal (e.g. comics or advertising brochures);
4. **audiomedial** texts are those written to be spoken (e.g. political speeches).
Multimodal texts are, according to this definition, those written to be performed live on stage (and, of course, for an audience). The distinction between media, modes and sign systems is of course important, as it is important to acknowledge the possibility of different evaluative frameworks for the same text, e.g. considering its medium, its mode or its sign system. At the same time, texts defined by their medium, such as films or television programmes, may also incorporate both visual and acoustic modes of signification, as well as different graphic sign systems. Given the permeability of these categories, for the purposes of this issue, film, television, drama, opera or comics are equally considered under the broad heading of multimodality.

As Littau (2011) has persuasively argued, with changing media technologies (manuscript, print, changes in paper quality and bookbinding; web-based texts and hyperlinking) come changing theories of translation. It makes sense then that the saturated multimodality of many texts today would require both a new, or at least a rethought, critical and analytical toolbox, and potentially also new approaches to translation. Rick Iedema has argued that multimodality “provides the means to describe a practice or representation in all its semiotic complexity and richness” (2003: 39). The analysis of multimodal texts will therefore require:

- a methodology produced for forms of description in which all modes are described and describable together. From an occasional interest in other semiotic modes this project moves to a norm where all texts are seen as multimodal and are described in that way (Kress and Ogborn, 1998, quoted in Iedema 2003: 39).

Within linguistics, discourse analytic approaches and approaches based on Hallidayan systemic functional grammar have rapidly gained ground (see e.g. Chueasuai in this issue). Within Translation Studies, some strides have been made in developing methodologies for research on multimodal texts in translation. The creation of tools for this purpose faces a number of technical and logistical challenges, and hence also of financial challenges. One striking case is research into the multiple-language versions of films produced in the early sound period between the late 1920s and early 1930s. Access to archival copies, and to the means for comparing them, is crucial for the pursuance of such research. But prints are difficult to get hold of, and sometimes even to locate; these versions were in many cases never released on VHS and DVD; the technical infrastructure for projection or flatbed viewing of film prints is expensive. And yet it has been undertaken, with great success. Nataša Šurovičová describes the difficulties of:

- locating two versions of the same film (itself a challenge, given standard archiving as well as cataloging practices [...] the effort of bringing together the two titles (which by the fiat of distribution were meant to be mutually exclusive – to see the German version of Anna Christie was usually meant to pre-empt seeing the
American version), not to mention arranging for two flatbeds next to each other [...] (2004: 7).

She goes on to describe the assembling of the first International Film Studies Summer School at Gradisca in Italy in 2003, where a group of researchers assembled and ran, “in the ad-hoc space of a 17th century palazzo, a 35 mm projection booth, a multimedia lab with dozens of titles, a document and book library and, above all, a revelatory series of films screened in two to three versions” (2004: 8). The workshop and its successors can be rightly said to have transformed this field of research.

This is, admittedly, an extreme example of the infrastructure which may be deployed to explore multimodal forms of translation (and the astute reader will have noted that it is much more oriented towards ‘old media’ (35mm film) than new media). A more easily replicable example is the multimodal transcription technique developed by Christopher Taylor and colleagues at the University of Trieste, which aims to provide a multimodal research tool for the analysis of the interrelationship of different semiotic modes in audiovisual texts (see e.g. Taylor 2004). We also observe research rooted in the methods of social science (e.g. Cambra in this issue) which works with audiences and empirical, quantitative methods to generate reliable experimental data.

On the technical side, such research has real potential to create impact which will improve the experience of consumers of multimodal products. For instance, the mobile application for accessibility of live performing arts discussed by Oncins et al. in this issue demonstrates how essential the collaboration between translation researchers, technicians and programmers can be in order to explore new horizons in multimodal translation. On a slightly different note, Brian Mossop’s article shows how translation researchers can conduct ‘action research’ into the applicability of translation theory to solving real-life multimodal translation problems.

One of the issues raised by multimodality is that of agency. Who are the people who are involved in making decision about the different multimodal elements of a text in translation? Liu usefully shows in a recent article on bilingual in-flight magazines how many agents are involved in the production of the image-heavy, layout-intensive bilingual texts of the corpus under analysis: publishers, editors, sub-editors, copywriters, translators, graphic designers and so on (2011: 203–206). Soler Pardo (in this issue) discusses the problem of swearing in film, showing how a number of different interests (of censors, of translators, or distribution companies) are served by the translation decisions which are ultimately taken. Pellatt (in this issue) shows how complex such relationships can be. Her article on the simultaneously published and produced Chinese and English versions of a memoir by former Chinese premier Zhao Ziyang demonstrates the challenge for the researcher in working with texts which have been produced by such a large team of transcribers, editors,
publishers and translators (not to speak of photographic researchers, designers and printers) and further paratextually framed by a number of different interested parties with particular political standpoints. Rike (in this issue) acknowledges how in the production of a commercial website translation is only one part of a complex transcreative operation in which ‘domestic’ and ‘international’ sections of the website address different target constituencies. Romero Fresco (in this issue) shows how for accessible filmmaking to function, it requires collaboration between audiovisual translators and also the many agents involved in filmmaking, at all stages of the process. This also requires new methods in research, crucial among which is contact with professionals in the industry (Liu 2011: 201; Chueasuai, Romero Fresco, Mossop in this issue). The workflows of multimodal translation are therefore an important future area for translation research.

4. Word and image

The question of word and image relations is difficult to separate from the wider issue of paratext. Keith Harvey (2003) supplements Genette’s notion of paratext with his own notion of ‘bindings’: specifically, the outward presentation of texts in the form of book covers and blurbs. Genette sees illustration as strictly paratextual (1997: 406) but also as a very large field which is beyond the scope of his study of paratext.

In some ways, the very notion of multimodality puts in question Genette’s notion of paratext, with its distinction between what is ‘text’ and what is on the fringe of that text. With many texts we have an intuitive sense that certain semiotic modes stand in an ancillary relation to the text, with a framing function; thus the relation between images and written text in comics seems more ‘integrated,’ more essential, than the relationships between images and written text on, say, a book cover or in an illustrated story for children. We can imagine the book cover or the children’s book with their images excised, or replaced; we cannot imagine the comic without its essential combination of text and image.

But this is perhaps an intuition about which we should be cautious. Mary Elizabeth Leighton and Lisa Surridge point out that “the material form of a text always signifies,” quoting Jerome McGann’s observation that “apparitions of text, its paratexts, bibliographical codes and all visual features […] are as important in the text’s signifying programs as the linguistic elements” (Leighton and Surridge 2008: 65). Leighton and Surridge’s study of illustrated serial fiction in the Victorian period persuasively demonstrates how the location of illustrations within the text interacted with the serial presentation of the text in crucial ways. Many of these texts are now read in unillustrated editions (cf. Kress and van Leeuwen’s observation about monomodality, quoted at the beginning of this introduction). Leighton and Surridge argue that the illustrations were not merely ancillary, but in fact constitutive of plot, and that reading in
editions which lack these illustrations leads to readers “failing to generate the visual knowledge bank that would have informed and guided the interpretive strategies of Victorian readers” (2008: 97).

The relations between image and text are highly complex, and may or may not be reproduced in translation. In an article on translations of Hans Christian Andersen’s short story ‘The Steadfast Tin Soldier,’ Cecilia Alvstad (2008) shows how the ambiguities of the text are sometimes reproduced, sometimes emphasised and often partially or fully resolved in translation through the different approaches to the text taken by illustrators.

We must also beware of thinking of images as elements of the text which, by contrast with the written text, do not get ‘translated.’ On the contrary, if we take the example of comics, not only does the interaction between text and image change in the translation (see e.g. Kaindl 2004), but the images themselves are subject to alteration, editing and even removal (cf. Zanettin 2011). The anti-hero of the long-running Dylan Dog horror comic book series in Italy has a sidekick named Groucho. In the Italian comic, the sidekick is clearly visually presented as Groucho Marx. In English translations of the comics, Groucho’s trademark moustache was edited out, in order to avoid problems with intellectual property relating to the character of Groucho (D’Arcangelo and Zanettin 2004: 194).

In the field of scientific publishing, Liangyu Fu (2013) has shown how in the late nineteenth and early twentieth century, diagrams and illustrations in scientific books translated into Chinese were frequently edited, recontextualised and indigenised in various ways, both for technical reasons and in order to align the illustrations with the Chinese visual tradition.

Multimodality can be heightened in translation, for the purposes of reframing a text for a new audience. Raymond Queneau’s 1959 novel Zazie dans le métro was translated in the same year by a Paris-based publisher, Maurice Girodias. His publishing house, the Olympia Press, produced books aimed at the market for literary (and not-so-literary) pornography. The Traveller’s Companion series, with their plain green covers, were infamous among English-language readers. Queneau’s satirical novel hardly fit the bill for a ‘naughty book,’ however, and so Girodias commissioned an illustrator, Jacqueline Duhème, to enrich the book’s pages with marginal illustrations emphasising the more suggestive elements, and the Parisian location, of the narrative (a few examples of these can be seen at http://blog.ink-stainedamazon.com/?p=2201). The illustrations were part of an attempt to ‘rebrand’ the novel, with the help of some racy textual choices in the translation by Akbar del Piombo and Eric Kahane, as a piece of suggestive literature (O’Sullivan 2002). The multimodal affordances were necessary in order to supplement what the text could not supply.
In this issue, Valerie Pellatt’s article shows how the different selection and placement of photographic imagery in Zhao Ziyang’s memoir, and differences in cover design, correspond to a particular framing of the narrative for different audiences. Pasakara Chueasuai shows how small changes in colour and placement can affect the consumption and the word-image relations in the sex advice columns of the Thai translation of *Cosmopolitan*. His article also discusses how ‘non-translation’ of photographic imagery (i.e. the inclusion of unaltered photographs of Caucasian models from the American edition of the magazine) in fact enacts a profoundly translating movement within the Thai context, by distancing Thai readers from the sexual activities pictured within the magazine by framing these acts as things which are done by Caucasian, rather than Thai, women and men.

5. Accessibility

One of the major fields of multimodal research in translation studies has been that of accessibility, particularly the accessibility of multi-medial experiences (e.g. museums — Soler Gallego and Jiménez Hurtado in this issue) and entertainment products, e.g. films (Romero Fresco, Maszerowska in this issue), television (Cambra in this issue) and live performing arts (Oncins et al. in this issue).

By definition, the multimodality of these texts places specific demands on the translator, but also creates a need for certain forms of access translation, in the form of audiodescription for spectators who are blind or partially sighted, as well as subtitles for spectators who have difficulty in hearing. Because European and national legislation has supported the wide availability of such translations, a booming industry has sprung up, which constitutes one of the ways in which translation is becoming visible to the general public. The importance of this area to the media and translation industries overall is highlighted by the fact that no less than five of the articles in this special issue deal with this topic.

The article by Oncins et al. introduces a practical application of research into accessible translation in the form of an app which has the potential to make the translation of live performance accessible in a much more flexible way to audiences with a variety of sensory impairments. The article by Soler Gallego and Jiménez Hurtado discusses how close analysis of the audioguides of museums is necessary in order to prepare the ground for their translation. Anna Maszerowska takes a component of audiovisual texts which has received to date very little critical interest, namely lighting design, and shows how it constitutes a recurrent challenge for audiodescribers, and at the same time how the connotative ambiguities of lighting design may be interpreted in a variety of ways.

Accessible forms of translation need constantly to take into account the needs of their audience. Romero Fresco in this issue shows how lack of
attention to the requirements of the audience for a subtitled film can lead to subtitles which fail to achieve their objective, e.g. through something as simple as the colour scheme which forms the background to the subtitles. He argues powerfully for filmmakers to take into account the likely requirements of end users of media products who need to take advantage of forms of accessible translation; they constitute a large potential market. But he also acknowledges that this is likely to be an uphill struggle, since translation is traditionally an afterthought at best in film production. The reader of this special issue also has the opportunity to see this discussion in action, as the film is also available on the journal website. (The inclusion of Romero-Fresco’s film incidentally illustrates the importance for academic work of multimodal affordances. Audiovisual translation studies and indeed the analysis of all kinds of multimodal texts would be made immeasurably easier by the development of instruments for the easy juxtaposition of still and moving images. The legal infrastructure to allow this, which is currently lagging far behind the technologies available to accomplish it, also needs revision in order that the research in, for instance, translation studies can engage the widest possible audience through the presentation of vivid examples of translation in all of its multimodal glory.)

The needs of target audiences are also the subject of the article by Cambra on the visual attention and understanding of hearing impaired children watching subtitled cartoons. The dynamic relationship between image and subtitles must constantly be taken into account in order to optimise children’s engagement with and understanding of subtitled animation.

6. Multimodality as challenge and resource

The discussion so far has tended to focus on the translation of multimodal texts and the challenges posed by this multimodality for the translator and indeed for the reader/consumer of these texts. But it is also worth remembering that multimodality is also a resource for translation. This is demonstrated, in a very literal sense, in the article by López Rodríguez et al. in this issue where they discuss the crucial role played by images in today’s term-banks. The combination of visual and verbal elements is essential in order for terminology to be properly understood and translated.

In subtitling, the multimodality of the audiovisual text is both a challenge and a resource for subtitlers. The image may impose severe challenges on the translator, e.g. through instances of verbal/visual puns, but through verbal/visual redundancy the other modes of the audiovisual text can also provide sufficient context to make certain verbal elements redundant, and thus make it easier to condense the text (see e.g. Cambra in this issue).
Brian Mossop contributes an article on developing tools to help singers to sing in foreign languages. The aim is to make the text both comprehensible and pronounceable by engaging with the material and acoustic features of the language, more than the typographical image. Mossop draws on Catford’s translation theory to propose an intuitive transliterative model, adapting the medium of the printed sheet of music as a channel for the purposes of this intermediary form of translation.

Jonathan Evans’ article on the translation of contemporary board games shows how multimodality, while having the potential to constitute a challenge for translation, can also serve as a solution to the problem of translation. The case study of Caylus Magna Carta shows how a board game can be conceived using a combination of plastic, visual and verbal elements to minimise the need for translation, particularly of expensive game pieces, and maximise the likelihood of a game’s popularity across languages and cultures.

It is hoped that this special issue of *JoSTrans* will constitute a small step forward in scholarly discussions of the relationships between translation and multimodality. My co-editor Caterina Jeffcote and I would like to thank the several colleagues who have supported the issue over the course of its compilation, notably the JoSTrans editorial team, and Ian Kemble, Begoña Rodríguez and Margaret Clarke of the University of Portsmouth.

**Bibliography**


**Biography**

Carol O’Sullivan is a Senior Lecturer in Translation at the University of Portsmouth. Her research interests include audiovisual translation, translation history and literary translation. She is the author of articles on the translation of classical literature, censorship in translation, Victorian translation publishing and subtitling. She is a co-editor with Luc van Doorslaer, Denise Merkle and Michaela Wolf of *The Power of the Pen: Translation & Censorship in Nineteenth-Century Europe* (LIT Verlag, 2010). Her monograph *Translating Popular Film* was published by Palgrave Macmillan in 2011. She is Associate Editor of the journal *Translation Studies*. E-mail: carol.osullivan@port.ac.uk.

**Notes**

1 The conference was held at the University of Portsmouth on Saturday 6 November 2010.

2 Liu (2011: 209) identifies a very similar issue for in-flight magazines; because Chinese texts take “approximately 55 percent less space” than equivalent English texts, the texts must be edited down for layout.