Combining computer vision and multimodal analysis: a case study of layout symmetry in bilingual in-flight magazines

Tuomo Hiippala

Centre for Applied Language Studies
University of Jyväskylä

1 Introduction

Regardless of the approach, the currently blooming research on multimodality is always in need of empirical scrutiny, because establishing which of our theories are found wanting, in which contexts, how and why, requires forging a strong link between theory and data. The principles of empiricism, however, are often neglected due to the time-consuming nature of multimodal analysis. Anyone working within the field knows the time and resources needed to pull apart the meanings in a film scene, a manuscript page or a single advertisement, and the effort of putting them back together during analysis. Yet these painstaking analyses hold the promise of bringing novel perspectives to the study of human communication and meaning-making in general. If this is indeed the case, then multimodal research should strive to be empirically responsible and to show that the field can also address research questions that fall outside its traditional concerns.

For this reason, this chapter aims to make a twofold contribution. Firstly, I address certain concerns about layout that have been raised in studies of written code-switching and multilingualism. More specifically, I ask whether the spatial placement of content in bilingual documents can carry additional meanings, that is, do the designers use layout to signal the reader that the bilingual content is semantically equivalent? In this chapter, this design will be termed *layout symmetry* and theorised in the following discussion.

Secondly, to explore how the proposed symmetry is used in bilingual written discourse, I develop a method that takes advantage of computer vision techniques to analyse page layout and to determine its degree of symmetry. By using the Python programming language to access the OpenCV computer vision library, I show how off-the-shelf software can be applied to study layout features in a data set comprising 1373 double-pages in in-flight magazines. In this way, I demonstrate how computational approaches can increase the volume of data in multimodal research by handing some analytical tasks over to the computer.

The chapter begins with a discussion of recent research on written multilingualism, contrasting methodological requirements identified within this field of study with those that are currently given consideration in multimodal research. I then explore some of these points of contact in more detail, before proceeding to discuss how computer vision may be leveraged in the analysis of layout symmetry in multilingual texts. Then, after describing the data and explicating the method, I present an analysis of a data set containing 1373 layouts from *Blue Wings*, the bilingual English/Finnish in-flight magazine published by Finnair. I then conclude the chapter with a discussion and point out several avenues of further research.

2 Multilingualism and multimodality

Mark Sebba, who has done pioneering work in the area of multilingualism in writing, notes that such texts "have been documented and analysed going back as far as ancient times" (Sebba, 2013, 100). Despite the prevalence of multilingual texts, written language-mixing remains severely under-researched and -theorised, due to the emphasis placed on code-switching in spoken language. Reflecting on how to address these shortcomings, Sebba (2012, 2) proposes that written language-mixing must be studied within a broader semiotic framework capable of looking beyond the mixed-language written text. Specifically, he argues that in order to fully appreciate mixed-language written texts, the analysis must account for their production and consumption.

Sebba's proposal has found support in recent multimodal research, which has emphasised the impact of production and consumption on the structure of multimodal texts (cf. Hiippala, 2015). Sebba (2012), in turn, points out that the analysis of multilingual texts must be conducted with multimodality in mind, that is, it must attend to other modes of communication besides written language. This involves, for instance, examining the visual characteristics and spatial placement of the mixed-language text on the page. In short, the points of contact between emerging approaches to the study of multilingual written discourse and multimodality seem to be numerous: I will explore these points in greater detail below.

Sebba (2013) sketches a framework to support the step from linguistic to multimodal analysis in the study of multilingual texts. According to Sebba (2013, 106-108), an analytical framework geared towards the study of multilingual texts must account for four issues. Firstly, at least three types of analytical units are required: (1) grammatical, (2) genre-specific and (3) visual/spatial. Grammatical units range from morphemes to sentences, whereas genre-specific units consist of headers, paragraphs and other typographic units. Visual and spatial units, in turn, include unified bodies of text, which comprise, for instance, a header and multiple paragraphs of text. Together with the possible graphical contributions, these visual units constitute the basic building blocks of a multimodal page, which need to be taken apart and put back together in various ways during the analysis, in order to understand the principles behind their organisation (cf. Bateman, 2008, 116-117).

Secondly, the framework must account for language–spatial relationships on the page. These relationships involve observing the placement of visual units in the layout, and determining whether these units occupy equal amounts of space. The placement and size of the visual units may be either symmetric or asymmetric. In a symmetric configuration, the visual units of two languages 'mirror' each other: in any other case, their configuration is asymmetric. The symmetry or asymmetry in layout is also related to the third issue, that is, language–content relationships. Although mirror-symmetric layout designs may replicate typographic features in both languages, these features do not guarantee that their contents are semantically equivalent. Some multilingual texts feature what Sebba (2013, 107) terms a "disjoint", in which different languages contain different content. Another possibility is that the content presented in two languages overlaps partially.

Fourthly and finally, the framework needs to consider the kind of language mixing at play. Languages can be mixed on the page in several ways: individual visual units may be monolingual, or contrastingly, they may be multilingual and mix languages within the same visual unit. Moreover, certain units may be "language-neutral" and thus difficult to assign to any particular language. These include, for instance, brand names, headers and words that take advantage of similarities across vocabulary. Sebba (2013, 108) gives an example from the packaging of a household hygiene product in English, French and German:

> "Dermatolog. tested, pH-skin-neutral, alcohol-farbstoff-colorant free"

As Sebba (2013, 108) points out, the abbreviated word 'Dermatolog.' can stand for 'dermatologically/dermatologisch/dermatologique'. Although such abbreviations highlight linguistic creativity and show how language-mixing occurs on several levels of linguistic structure, Sebba's proposed move towards multimodal analysis requires that an equally tight analytical grip may be achieved on multimodal phenomena on a more abstract level, that is, for the visual units. These visual units bring together both linguistic and graphical contributions, forming an additional level of organisation. Coincidentally, it is the household hygiene products that help to bridge the gap between linguistic and multimodal analyses: Thomas (2014) develops a framework for studying the cross-cultural and multimodal aspects of 'pack messages' in shampoo and toothpaste packaging.

Emphasising the need for increased empiricism in multimodal research, Thomas (2014) calls for attention to two methodological issues: firstly, to the need to avoid under- and over-interpretation, and secondly, to the need to delay the interpretation process to minimise the analyst's subjective experience of the data. Relying too much on initial impressions runs the risk of conflating different multimodal structures, which weakens the capability to explain how these impressions and meanings arise in the first place (cf. Thomas, 2014, 169). These kinds of conflations resulting from over-interpretation include, for instance, Kress and van Leeuwen (2006) proposal for information value zones, which assigns predefined meanings to specific layout areas, regardless of the content present in these areas. As an analytical tool, the information value zones are far too coarse for defining the analytical units needed for a description of multimodality in multilingual texts. The oppositions put forward by Kress and van Leeuwen (2006, 175-203) – given–new, ideal–real and centre–margin – reveal little about specific language–spatial relationships, which Sebba (2013, 107) posits as essential for understanding multilingual texts. To put it simply, finer descriptions are required to segment the multimodal texts into what Sebba (2013) terms visual and spatial units.

Thomas (2014) pursues this kind of description by applying the *Genre and Multimodality* (GeM) model, which provides multiple analytical layers for a systematic description of page-based multimodal texts (cf. Bateman, 2008). The layers of the GeM model account for various aspects of a multimodal text: the content, its appearance, hierarchical organisation and placement in the layout, its rhetorical organisation and navigation structure. The model also offers an annotation schema for creating multimodal corpora, which cross-references the analytical layers, making it possible to observe how the different layers of description relate to each other. Thomas (2014) argues that the GeM model is particularly apt for tackling the methodological challenges outlined above, as the model enables the analyst to work with annotated data instead of raw, hand-picked examples. To facilitate the creation of multimodal corpora, Thomas (2014, 175) applies optical character recognition – a computer vision technique for detecting letterforms and text – to identify layout blocks in the shampoo and toothpaste packages, which correspond closely to Sebba's (2013) definition of visual and spatial units.

The GeM model has been previously used for cross-cultural investigations. For instance, Bateman and Delin (2003) seek to identify appropriate units of comparison for performing contrastive analyses of multimodality across cultures. Arguing for the need to account for multimodality in cross-cultural communication and translation, Bateman and Delin put forward several proposals that resonate with Sebba's (2013) framework for studying multilingual texts. They point out that multimodal texts are never 'ideal' organisations of meaning, but subject to a number of constraints, and propose that the page constitutes a suitable basic unit for contrastive analysis, which needs to be taken apart and put back together during the analytical process (Bateman and Delin, 2003, 252-253).

To support the methodological toolkit with a theoretical framework, Bateman and Delin (2003) deploy the notion of genre to explain the patterns detected using the GeM model in multimodal texts across different cultures. This proposal has been put to work, for instance, by Kong (2013), who exemplifies how the GeM model can be used to deconstruct the multimodal structure of global news items in English and Chinese tabloid newspapers, showing how local genre conventions shape the multimodal presentation of global news.

The work presented in Bateman and Delin (2003), Kong (2013) and Thomas (2014) suggests that the GeM model presents an effective analytical framework for studying multimodality in cross-cultural contexts. Additionally, given the many points of contact with the framework proposed in Sebba (2013), the model appears to hold considerable potential for studying multilingual texts. However, due to the rich description provided by the model, using the GeM framework to create a multimodal corpus consumes both time and resources Hiippala (cf. 2015, 213). For this reason, the GeM model can only be used to evaluate the hypotheses put forward in Sebba (2013), if an appropriate level of description may be identified. One testable hypothesis concerns layout symmetry, which is arguably a relatively abstract phenomenon, as the symmetry emerges among the visual units and their placement on the layout. At the same time, studying these layout patterns obviously requires more than hand-picked examples. Kong (2013, 179) observes that:

> "... attempting to understand how information is arranged as a pattern, making any claims on empirical grounds requires a large number of texts."

With this observation in mind, the following section explicates how computer vision techniques can be deployed to study layout symmetry, while also enabling a considerable increase in corpus size.

3 Leveraging computer vision

Computer vision is a vast field of study, which develops techniques for retrieving, processing, analysing and interpreting images. What are broadly understood as images in computer vision may include, for instance, photographs of objects, scenes, faces, documents and many more. As exemplified by Thomas (2014, 173-175) above, one important area of computer vision is optical character recognition, which is a technique for detecting and retrieving text from document images. This technique has had a considerable impact on the field of linguistics, freeing time and resources from the laborious process of manual input by converting printed texts into a machine-readable form, thus allowing an increase in the size of corpora.

Other computer vision techniques are currently being introduced to multimodal research as well. O'Halloran et al. (2014), for instance, use automatic face detection to analyse photographs shared on a social networking site. Their data set comprises 301865 geographically located photographs taken in Singapore, which the authors combine with other data sets of Twitter messages and FourSquare check-ins to describe patterns of human behaviour in social media across the city. Given the volume and complexity of the data, this kind of work naturally presupposes the involvement of computer scientists. The same applies to more complex tasks, such as using computer vision and machine learning to automatically annotate and create multimodal corpora, which represents a long-term goal far beyond the scope of the current chapter.

In the current study, however, relatively simple computer vision techniques may be used to test Sebba's (2013) hypothesis about layout symmetry in multilingual documents. Because layout symmetry is reflected on the 'surface' of the multimodal text and therefore does not necessarily require semantic analysis of content, this feature of written multilingual texts is susceptible to computer vision. In plain words, layout symmetry leaves visible traces on the document surface, which the computer can perceive as pixels. A pixel is the smallest element in a digital image: computer vision typically attends to the properties of each pixel and in many cases, to their groupings and their properties as well, in order to establish how different pixels relate to each other. This represents the level of detail which is rarely given consideration in multimodal analysis, and for this reason, bridging the gap between multimodal research and computer vision requires establishing how individual pixels and their groupings reflect semiotic phenomena.

4 Data and methods

This section explicates which computer vision techniques were used and how to examine layout symmetry in a particular kind of multilingual text: the in-flight magazine. Sebba (2013, 100) notes that airlines often provide the passengers with multilingual magazines for in-flight entertainment. Thurlow and Jaworski (2003) have drawn similar conclusions about in-flight magazines as a genre, but emphasise that the in-flight magazines are also characterised by multimodality. These two features – multilingualism and multimodality – make the in-flight magazines a suitable target of analysis for studying layout symmetry. Moreover, the in-flight magazines include various types of content, ranging from travel and destination information to feature articles on lifestyle and culture, which are likely to reflect different multimodal and multilingual characteristics (cf. Thurlow and Jaworski, 2003, 584).

However, in order to apply computer vision techniques to study layout symmetry in the in-flight magazines, the data must be in a digital form. Fortunately, the Finnish flag carrier Finnair also makes its in-flight magazine *Blue Wings* available online.[1] The airline has published the in-flight magazine since 1980. Until April 2012, the magazine was bilingual, including content in both English and Finnish. Since May 2012, the magazine is published in English only. For this study, a total of 24 issues of *Blue Wings* published between January 2010 and April 2012 were collected and compiled into a data set containing 1373 double-pages of content in Finnish, English or in both languages. Magazine covers were excluded from the data set, together with three double-pages from the June 2011 issue, which contained an encoding error that prevented their retrieval from the website.

The contents of each magazine were stored as screenshots, which were captured automatically from the *Blue Wings* website by a Python script. Because the screenshots captured the entire screen, another script was then used to crop each double-page from the screenshot. All scripts were written in the IPython environment for rapid prototyping and testing (cf. Pérez and Granger, 2007). Moreover, the decision to use Python was motivated by easy access to OpenCV (Open Source Computer Vision), a free open source computer vision library also used for multimodal analysis by O'Halloran et al. (2014).

OpenCV provides a range of programming functions for common computer vision tasks, such as for manipulating images, which are directly applicable to the study of layout symmetry in bilingual documents. Applying OpenCV, however, requires understanding how computers perceive images, and what this implies to multimodal analysis. What humans perceive as unified bodies of text or images – the units which Sebba (2013, 106-108) terms 'visual' and 'spatial' – the computer perceives as groups of pixels. Each pixel has a value, which determines its luminous intensity: for a greyscale image, the value for each pixel ranges from 0 (black) to 255 (white) with 254 shades of grey in between.

At pixel level, differences begin to emerge already when contrasting single letters. Compare, for instance, the capital letters 'A' and 'L' on your screen: the pixels that make up these letterforms overlap only partially, mainly at baseline and ascender height, that is, bottom and top. For this reason, if the computer were instructed to calculate a Structural Similarity Index, which measures the degree of similarity between two images (hereafter SSIM, see Wang et al., 2004) for these two letters, the algorithm would return a relatively low value. Moreover, in addition

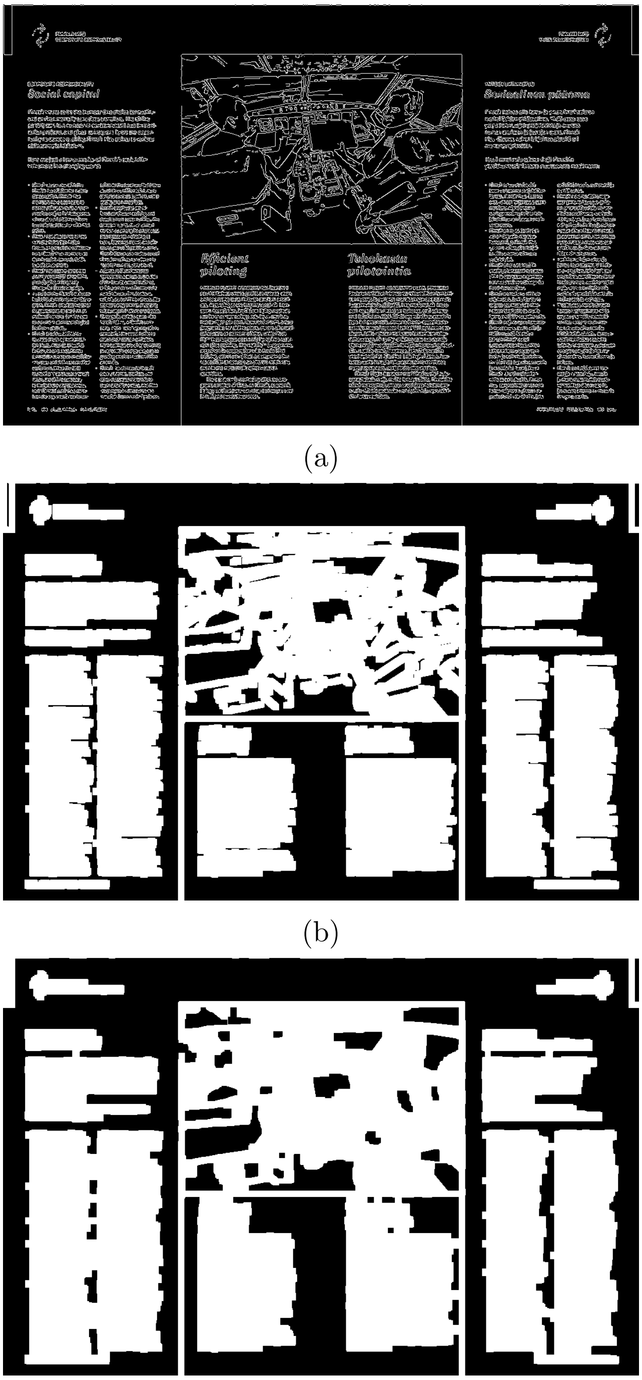
[1]The *Blue Wings* magazines are available at: http://www.digipaper.fi/bluewings [last accessed: 22 March 2015].

to evaluating their placement, calculating an SSIM also involves comparing the intensity of pixels placed in corresponding positions. What this means is that the algorithm would not consider two instances of the same letter, such as 'L', to be perfectly similar in terms of structure, if one were black and the other were grey.

These two factors – a comparison of positioning and intensity at pixel level – set two requirements for applying SSIM to examine layout symmetry: single letters need to be turned into solid blocks to represent text, and the number of colours have to be reduced to a bare minimum, that is, black and white. Meeting these two requirements constitute a prerequisite for calculating an SSIM for each image using the scikit-image package for Python (van der Walt et al., 2014).

To achieve this using OpenCV, I applied the following procedure to the data:

1. Load the original image for each double-page and classify them into three categories according to the languages present:
   - Monolingual double-pages with content in Finnish or English only.
   - Mixed-language double-pages with content in Finnish and English, positioned across the entire double-page spread: this category is consistent with Sebba's definition of *complementarity* (2012, 14).
   - Bilingual double-pages with content in Finnish and English, positioned on opposite sides of the double-page spread: this category contains the pages that correspond to Sebba's definition of *parallelism* (2012, 15).
2. Identify the outlines for each letter and image using the Canny edge detector in OpenCV. The detector produces a binary image, that is, an image that contains only black and white pixels, as shown in Figure 1a.
3. Dilate the image using the OpenCV dilate function. Perform five passes with a rectangle-shaped 3x3 pixel kernel. This turns the individual letters and words into text blocks and reduces detail in photographs and other graphic elements, as shown in Figure 1b.
4. Close the black areas contained within white areas using the OpenCV close function, performing two passes with a rectangle-shaped 5x5 pixel kernel, as shown in Figure 1c.
5. Split the image into two halves, left and right, and use the OpenCV flip function to mirror the right-hand side. This step is necessary, because SSIM values are calculated at pixel level: the computer is unable to perceive symmetry in the same way as humans do.
6. Calculate a value for Structural Similarity Index for the left half and the flipped right half using scikit-image.



(a)

(b)

(c)

Figure 1: (a) Canny edge detector applied to a double-page; (b) dilating the resulting image; (c) closing the gaps in the dilated image

7. Calculate the percentage of white pixels on each double-page to approximate how much layout space is taken up by content.

8. Store the filename and values for SSIM and the percentage of white pixels into a file as comma-separated values.

9. Divide the comma-separated values into three files according to their category: monolingual, mixed-language or bilingual.

This procedure resulted in three files, which serve as the foundation for the analysis presented in the following section.

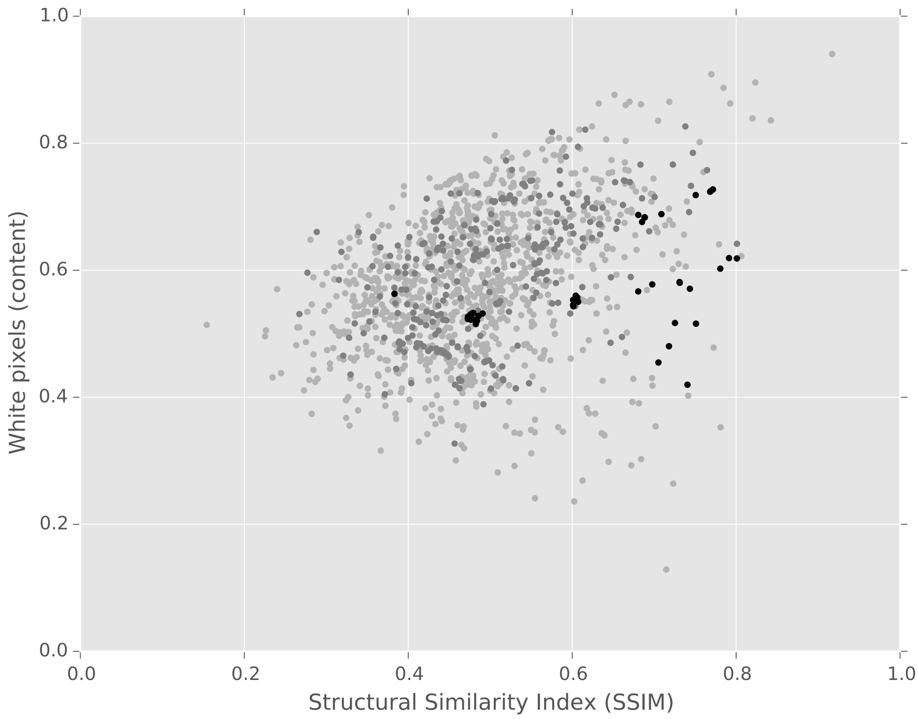


Figure 2: SSIM values and white pixels on 1373 double-pages from *Blue Wings*

5 Analysis

To begin with, Figure 2 shows the entire data set containing 1373 layouts in a scatter plot. The SSIM values are represented on the horizontal axis, where 1.0 indicates perfect symmetry. In the case of perfect symmetry, the position of each pixel on the left- and right-hand side pages would correspond precisely. On the vertical axis, which measures the percentage of white pixels present in the layout, the maximum value of 1.0 would mean that the double-page spread is filled with content. The black dots mark the bilingual double-pages (n = 44), while the dark grey dots stand for mixed-language double-pages (n = 297). The light grey dots indicate monolingual double-pages (n = 1032).

I will begin by focusing on the bilingual double-pages, before moving on to consider the mixed-language and monolingual double-pages. A closer examination of the data revealed that the bilingual double-pages deal with a rather limited set of topics, which may be roughly divided into three categories:

1. in-flight entertainment, service and on-board purchases,
2. check-in, transfer and services at the Helsinki-Vantaa airport,
3. information on Finnair's actions towards corporate social responsibility.

While the first two topics constitute common content categories in the in-flight magazines, which convey practical information to the passengers (cf. Thurlow and Jaworski, 2003, 584), the third one – corporate social responsibility – is a relatively new feature which is likely to arise from increased demands for corporate accountability and transparency (cf. Nielsen and Thomsen, 2007).

A glance at Figure 2 reveals three clusters of bilingual double-pages. Firstly, practical information about the flight and descriptions of corporate social responsibility have the highest degrees of layout symmetry, with the SSIM values ranging between 0.7 and 0.8. Their scattering on the horizontal axis may be explained by variation in the amount of content. The double-pages related to corporate social responsibility tend to include more negative (empty) space (average 0.54, SD = 0.06) than those related to practical information (average 0.69, SD = 0.02). In terms of Sebba (2013, 107), the content presented in English and Finnish is semantically equivalent. On the level of the double-page, their positioning is parallel (cf. Sebba, 2012, 14) and for this reason, the layout of both pages mirror each other, as reflected by their high SSIM values.

Secondly, a smaller cluster in the 0.6 SSIM range contains several double-pages with practical information about the Helsinki-Vantaa airport. Despite mirrored positioning of text paragraphs, the top of the double-page includes a world map that extends across the layout, which is not mirrored. Moreover, although the general design reflects the principles of parallelism and equivalence, two photographs placed in corresponding positions on both left- and right-hand pages contain different content. Thus, it appears that layout symmetry does not mirror graphic elements: I will discuss how certain graphic elements can be 're-used' in the following section.

Thirdly and finally, yet another cluster may be identified in the 0.5 range. As the SSIM value suggests, although these double-pages are bilingual, they do not exhibit layout symmetry. These pages are a part of a sequence offering practical information in both English and Finnish, which spans over three double-pages and dedicates three pages of content for each language. With a total of three pages for each language in this sequence, the first and the third double-pages are monolingual, while the second includes content in both English and Finnish. For this reason, there is a clear disjoint in terms of content (cf. Sebba, 2013, 107).

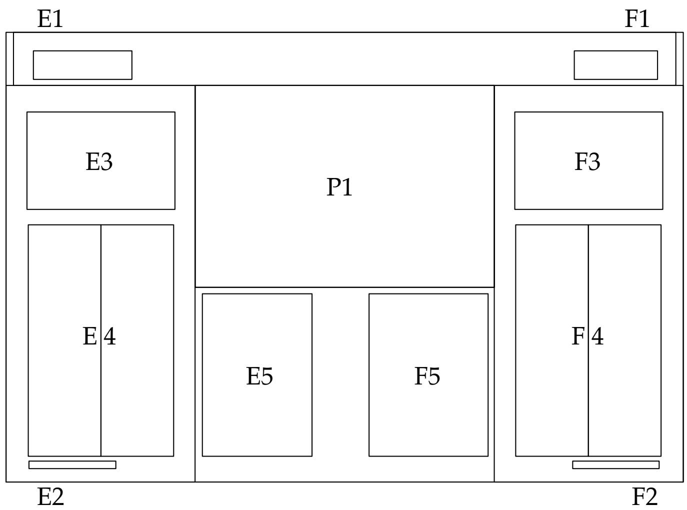
Given the high number of mixed-language double-pages, it is necessary to consider why layout symmetry is reserved for communicating specific kinds of content. The mixed-language double-pages, which make up 22% of the data, also include a variety of bilingual content, such as indices, captions that summarise the photographs on otherwise monolingual double-pages, parallel bilingual paragraphs providing travel and event information and brief summaries of feature articles in either English or Finnish. These double-pages reflect the kind of complementarity described by Sebba (2013, 109), in which the language–spatial relationships are characterised by asymmetry and language–content relationships by disjoint.

What comes to the monolingual double-pages, their SSIM values and the amount of content present in the layout, it is not surprising to observe the SSIM values range from less than 0.2 to over 0.8. A feature article that consists mainly of written text laid out in paragraphs will have a high SSIM value due to the presence of monolingual text blocks on opposing pages. This is, of course, completely normal: rendering the content using a semiotic mode such as text-flow, which organises written language into paragraphs and entire texts Bateman (cf. 2011), will occupy large parts of the layout space and appear to the SSIM algorithm as structurally similar. However, unlike the monolingual double-pages, their bilingual counterparts use layout symmetry to signal language–content relations. According to Sebba (2013, 109), layout symmetry is a "visual metaphor for equality". To better understand how this symmetry emerges, I will now examine this particular design from a multimodal perspective using the GeM model.

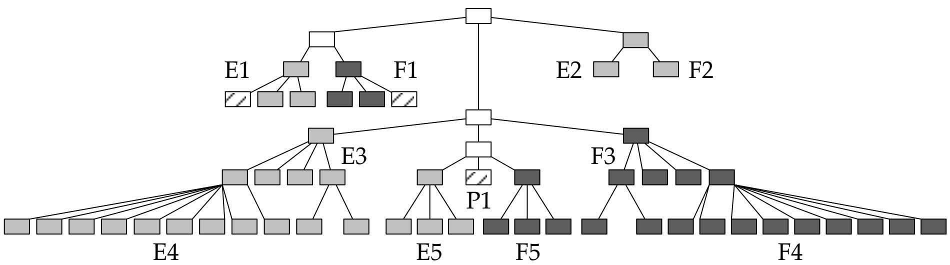
Having identified the bilingual double-pages that take advantage of layout symmetry, the next step is to consider their multimodal structure. From the perspective of page layout and design, layout symmetry should be understood as a deliberate rhetorical strategy, which involves manipulating multiple semiotic modes with a certain communicative goal in mind Hiippala (2014, 114). In this case, the goal is to signal the reader that the bilingual content is semantically equivalent. However, adopting this rhetorical strategy involves more than mirroring the spatial placement of content in the layout: additional "cue structures" Holsanova and Nord (cf. 2010, 83) are required to invoke the interpretation associated with layout symmetry. In the following discussion, I propose that some of these cue structures may be captured using the analytical tools provided by the GeM model.

To begin with, Figure 3a shows an area model, which is used within the GeM model to represent the physical layout (cf. Bateman, 2008, 124-126). For comparison, the double-page in Figure 3a is also shown in Figure 1. Focusing on the area model, the bilingual double-page in Figure 3a clearly shows a symmetrical layout. The content in English, marked and numbered using capital letter E, is spread across the left-hand side of the double-page. The corresponding Finnish content, indicated by F, is positioned on the right-hand side. Additionally, a photograph, P1, extends across the centre of the layout. Again, it is worth noting here that this photograph is the only content element not mirrored in the layout, an issue to which I will return shortly.

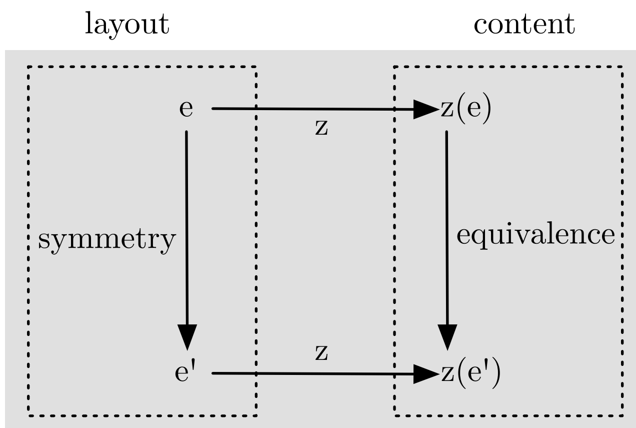
For a more precise look at the content on the double-page, Figure 3b shows its layout structure (cf. Bateman, 2008, 122-123). In the GeM model, the layout structure captures the hierarchical organisation of the content, in order to examine, for example, how paragraphs combine to form larger bodies of text. Whereas the topmost node in Figure 3b represents the entire double-page, the terminal nodes within each branch of the tree diagram stand for headers, text paragraphs and list items, to name but few examples. They are positioned across the layout, as indicated by the matching identifiers shown above in Figure 3a. In Figure 3b, the light grey nodes mark the English content. Contrastingly, the dark grey nodes stand for Finnish. Nodes with diagonal stripes, such as P1, indicate the graphic elements. Finally, the white composite nodes serve to indicate entry points into multilingual branches of the layout structure.



(a)



(b)



(c)

Figure 3: (a) area model for a double-page in *Blue Wings* 3/2012, pp. 108-109; (b) layout structure for a double-page in *Blue Wings* 3/2012, pp. 108-109; (c) discourse semantics for layout symmetry

I argue that by examining the area model and layout structure, we can already make headway into understanding layout symmetry as a phenomenon that is strongly rooted in the underlying structure of a multimodal artefact (cf. Hiippala 2015). In this case, the layout reflects a high degree of parallelism, that is, “matched units, symmetrically arranged and containing identical content in each language, without any language mixing” (cf. Sebba 2013: 109). The area model and layout structure allow making increasingly precise observations about how this kind of parallelism emerges.

The units E1–F1, which consist of the double-page headers in English and Finnish occupy parallel positions in both area model and layout structure. They both reproduce the same symbol with two arrows, which is often understood as a symbol for recycling. Contrastingly, the units E2–F2, which contain the page numbers, magazine title and month of publication, are both written in English! Considering the criteria used to group the pages during the analysis, this kind of configuration would actually suggest that the double-page in question is a mixed text. However, this kind of mixing does not involve the CSR content on the page, but page numbers, which are a feature of the magazine medium (cf. Hiippala, 2015, 113-114). They serve the purpose of helping the reader to navigate the magazine, instead of working towards convincing the reader of Finnair’s CSR activities. For this reason, I argue that it is reasonable to treat this double-page as an example of a parallel, symmetrical design, despite the language mixing within the page numbers.

The actual CSR content, in turn, is nearly perfectly mirrored. The units E3-F3 contain the respective section headers in English and Finnish, followed by introductory sentences to a list of socially responsible actions undertaken by Finnair (E4-F4). The most interesting example, however, emerges within the branch that contains the photograph P1, positioned above two parallel bilingual paragraphs E5-F5. The photograph shows two pilots in the cockpit, whereas the two paragraphs below highlight the economical flying techniques taught by Finnair to its pilots. Although no caption links the photograph and the two paragraphs explicitly, the cohesive ties between the verbal and visual content join them together.

From the perspective of layout symmetry, the re-use of the photograph is particularly interesting, as certain graphic elements, such as the recycling symbols within the units E1-F1, are mirrored as a part of the design, but the photograph is not. It appears that the photograph P1, which contributes towards the discourse about corporate social responsibility, is considered culturally appropriate for both English and Finnish audiences. Professional photography is an established mode of corporate communication, whose deployment must take the context of culture into account, but in this case, the same image works for both languages. It may even be suggested that corporate communication favours these kinds of culturally 'neutral' images (cf. Machin, 2014).

Finally, one more aspect of layout symmetry needs to be considered: the cue structure that guides the reader to consider the content in two languages as semantically equivalent. As noted above, many monolingual layouts exhibit visual features similar to those found in bilingual layouts with layout symmetry. Therefore, it is necessary – again – to turn towards the combination of the area model and layout structure.

Bateman and Schmidt (2012, 53) argue that taking the layout structure apart and reconstructing it is an essential step for making sense of multimodal artefacts. In simple terms, this basic interpretation process involves establishing which parts of the content belong together. In order to support the interpretation that the content in Finnish and English correspond to each other, the symmetrical double-pages take advantage of entire layout space and the underlying layout structure to highlight parallelism: the bilingual nodes are contrasted *spatially* in the area model and *hierarchically* in the layout structure.

To do so, the pages draw on discourse semantics that guide the interpretation of the entire double-page, which are set out in the back-and-forth diagram shown in Figure 3c. Figure 3c shows two symmetrical entities – e and e' – in the source domain **layout**, which the mapping relation z connects to the entities $z(e)$ and $z(e')$ in the target domain *content* (cf. Bateman, 2011, 28). This establishes a formal mapping, which yields the discourse semantic interpretation that symmetrically rendered bilingual content is equivalent in terms of meaning.

This general formulation of the principles that guide the interpretation of bilingual symmetrical layouts concludes the analytical part of this chapter. In the following section, I present a brief summary of the chapter and its implications.

6 Discussion and concluding remarks

This chapter has examined the structure and layout of bilingual double-pages using computer vision and multimodal analysis. Taking the work of Sebba (2012, 2013) as the point of departure, the chapter explored a specific phenomenon – layout symmetry – and its use in bilingual in-flight magazines. A total of 1373 layouts in *Blue Wings* were analysed using computer vision techniques to identify how often symmetrical, 'mirrored' designs were used for organising bilingual content across double-page spreads.

A closer examination using the GeM model revealed that adopting layout symmetry as a rhetorical strategy for communicating bilingual content does not only involve its spatial positioning and mirroring, as represented by the area model, but also extends to its hierarchical organisation in the layout structure. Moreover, layout symmetry extends to both content arising from the underlying medium (magazine), such as page numbers and running heads, and the genre realised using this medium (an in-flight magazine) and its various stages (cf. Thurlow and Jaworski, 2003).

Methodologically, it must be noted that the computer vision techniques used in this chapter have been relatively crude, focusing exclusively on layout symmetry across double-page spreads. In order to extend the analysis to single pages, the techniques must be complemented by machine learning for more precision in identifying text blocks, different graphic elements, their combinations and positioning. The detected elements and their positioning could be then evaluated locally, in order to automatically identify parallel and complementary structures (Sebba, 2013, 109). Additionally, the text blocks could be passed to an OCR engine for language detection. This would facilitate locating the corresponding units in bilingual documents and to evaluate their translations (Bateman and Delin, 2003).

Generally, computer vision holds immense potential for multimodal analysis. As Thomas (2014) notes, handing certain tasks, such as layout segmentation, over to the computer helps to delay the analyst's interpretation and makes the process less subjective. Moreover, this kind of automation is critical, if multimodal research wishes to take full advantage of the powerful methods used within traditional corpus linguistics. A long-term goal should be to automate as much of the analysis as possible in order to increase corpus size. Such endeavours will likely require collaborating with computer scientists, but given the widely available tools, libraries and tutorials, the researchers working on multimodality should not be afraid to explore the possibilities already by themselves.

References

Bateman, John A. (2008): *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents.* London: Palgrave Macmillan.

Bateman, John A. (2011): "The decomposability of semiotic modes". In: O'Halloran, Kay L. and Smith, Bradley A. (eds.), *Multimodal Studies: Multiple Approaches and Domains*, London: Routledge. pp. 17–38.

Bateman, John A. and Delin, Judy L. (2003): "Genre and multimodality: expanding the context for comparison across languages". In: Willems, Dominique/ Defrancq, Bart/ Colleman, Timothy and Nöel, Dirk (eds.), *Contrastive Analysis in Language: Identifying Linguistic Units of Comparison*, Houndsmills: Palgrave Macmillan. pp. 230–266.

Bateman, John A. and Schmidt, Karl-Heinrich (2012): *Multimodal Film Analysis: How Films Mean.* London: Routledge.
Hiippala, Tuomo (2014): “Multimodal genre analysis”. In: Norris, Sigrid and Maier, Carmen D. (eds.), *Interactions, Images and Texts: A Reader in Multimodality*, Berlin and New York: De Gruyter Mouton. pp. 111–123.
Hiippala, Tuomo (2015): *The Structure of Multimodal Documents: An Empirical Approach.* New York and London: Routledge.
Holsanova, Jana and Nord, Andreas (2010): “Multimodal design: Media structures, media principles and users meaning-making in printed and digital media”. In: Bucher, Hans-Jürgen/ Gloning, Thomas and Lehnen, Katrin (eds.), *Neue Medie - neue Formate: Ausdifferenzierung und Konvergenz in der Medienkommunikation*, Frankfurt/New York: Campus. pp. 81–103.
Kong, Kenneth C. C. (2013): “A corpus-based study in comparing the multimodality of Chinese- and English-language newspapers”. *Visual Communication* 12(2), pp. 173–196.
Kress, Gunther and van Leeuwen, Theo (2006): *Reading Images: The Grammar of Visual Design.* London: Routledge, 2nd edition.
Machin, David (2014): “Multimodality and theories of the visual”. In: Jewitt, Carey (ed.), *The Routledge Handbook of Multimodal Analysis*, London and New York: Routledge. 2nd edition, pp. 217–226.
Nielsen, Anne Ellerup and Thomsen, Christa (2007): “Reporting CSR — what and how to say it?” *Corporate Communications* 12(1), pp. 25–40.
O’Halloran, Kay L./ Chua, Alvin and Podlasov, Alexey (2014): “The role of images in social media analytics: A multimodal digital humanities approach”. In: Machin, David (ed.), *Visual Communication*, Berlin: De Gruyter Mouton. pp. 565–588.
Pérez, Fernando and Granger, Brian E. (2007): “IPython: A system for interactive scientific computing”. *Computing in Science and Engineering* 9(3). doi:10.1109/MCSE.2007.53, pp. 21–29.
Sebba, Mark (2012): “Researching and theorising multilingual texts”. In: Sebba, Mark/ Mahootian, Shahrzad and Jonsson, Carla (eds.), *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, New York and London: Routledge. pp. 1–26.
Sebba, Mark (2013): “Multilingualism in written discourse: An approach to the analysis of multilingual texts”. *International Journal of Bilingualism* 17(1). doi:10.1177/1367006912438301, pp. 97–118.
Thomas, Martin (2014): “Evidence and circularity in multimodal discourse analysis”. *Visual Communication* 13(2), pp. 163–189.
Thurlow, Crispin and Jaworski, Adam (2003): “Communicating a global reach: Inflight magazines as a globalizing genre in tourism”. *Journal of Sociolinguistics* 7(4), pp. 579–606.
van der Walt, Stefan/ Schönberger, Johannes L./ Nunez-Iglesias, Juan/ Boulogne, Francois/ Warner, Joshua D./ Yager, Neil/ Gouillart, Emmanuelle/ Yu, Tony and the scikit-image contributors (2014): “scikit-image:

image processing in Python". *PeerJ* 2(e453). doi:10.7717/peerj.453.
Wang, Zhou/ Bovik, Alan Conrad/ Sheikh, Hamid Rahim and Simoncelli, Eero P. (2004): "Image quality assessment: From error visibility to structural similarity". *IEEE Transactions on Image Processing* 13(4), pp. 600–612.