Ryoko Sasamoto
Dublin City University, Ireland
Minako O'Hagan
The University of Auckland, New Zealand

Relevance, Style and Multimodality:

Typographical Features as Stylistic Devices

Abstract: Brightly coloured textual inserts, which often occupy a sizable part of the TV screen, have become a key feature in Japanese TV. This paper unpacks the contribution of such multimodal stimuli to inference, and the consequences this has for the interpretation process. Using data derived from a mixed-methods approach (i.e. eye-tracking and a multimodal content analysis), we evaluate the relationship between cognitive processing and communicative stimuli. We demonstrate how typographical features (colours and fonts) are used as highlighting stylistic devices by TV producers to manipulate the viewer comprehension process by guiding the audience to an intended interpretation. The results suggest how editorial choices regarding typographical features to trigger certain effects might be subsumed under the current view of style in relevance theory.

1. Introduction

With the increasing use of digital devices, communication is now taking place more and more in a digital space. This shift in communication modes comes with an interesting use of written communication via text on screen.

Digital devices, compared with traditional writing devices such as pen and paper, are the driving force for an enhanced form of written communication and people are engaging in more visually appealing written communication, exploiting different fonts[1], colours, textual effects such as animations, and, in some cases, emoticons. Indeed, as van Leeuwen (2006) points out, typographical features are playing a more prominent role in modern day communication than it has previously been the case. This paper seeks to address the role of typographical features in the interpretation of multimodal contents, namely, how features such as font and colour affect the viewer interpretation process. We use Japanese TV programmes as our source of stimuli, as such programmes often consist of a complex layer of multimodal representations. In particular, we focus on *telop* (pop-up captions) which contain the outlandish use of typographical features. Our main goal is to demonstrate how a relevance-theoretic framework can be usefully applied in shedding light on the role played by such features used as stylistic devices. This paper also briefly refers to our preliminary study using eye-tracker to gauge the impact of these features on the recipients based on their gaze pattern. The paper is structured with the current Section 1 providing a background on telop used on Japanese TV, followed by a discussion in Section 2 on stylistic features from a perspective of relevance and multimodal analysis. Section 3 covers the treatment of typographical

[1] Strictly speaking, font and typeface should be distinguished. However, for the purpose of this paper and for simplicity, we use the term *font* to refer to both the design of letters and a particular presentation (i.e. size and weight) of the letters.

features as pragmatic effects within the framework of relevance theory. Next, a case study using a Japanese TV programme is presented to demonstrate the application of relevance theory and multimodal analysis of telop in Section 4. We will see how typographical features could give rise to enhanced pragmatic effects such as communication of weak implicatures and propositional attitudes. Section 5 briefly introduces a preliminary eye-tracking study in an attempt to present some empirical supposed based on viewer data. Section 6 provides a summary of the findings in relation to our research questions.

1.1 Background: Telop on Japanese TV

In the last decade or so, the excessive use of written props in the form of captions has become a key feature of certain TV programmes in Japan as well as in other parts of Asia. Compared to other types of captions, such textual inserts (or *Telop* as known in Japan) are immediately conspicuous and dominate the TV screen (see Figure 1[2]).

[2] Throughout this paper, parts of figures have been pixelated for copyright and portrait rights.

Figure 1: Multi-coloured captions used in a typical Japanese programme (Honmadekka!? 2013)

Figure 1 shows a typical scene from a Japanese entertainment programme, belonging to a genre called *a variety show*. The scene contains three clusters of captions: the caption at the top left shows the title of this particular episode, while the caption at the top right shows the topic of the discussion taking place in the scene. The caption at the bottom, which is an example of *telop*, is a verbatim rendering of an utterance produced by one of the studio panelists[3]. Telop are usually added during the post-production editing process and presented in multiple colours. In this scene, captions at the top of the screen are written in red, white, and black, while telop appearing at the bottom of the screen is written in red, green, yellow, and blue with a white 'glow' effect.

[3] This does not mean telop are always verbatim representations of utterances produced in a programme. They can be a verbatim production of an utterance, a summary, or a TV producer's interpretation of a certain utterance/thought.

Although telop are often (but not always) verbatim representations of utterances produced by the speakers in the programme, they are essentially not intended as an aid for the hearing-impaired, unlike 'closed captions' designed specifically for accessibility[4]. Colours and fonts used in telop are therefore not in accordance with norms established for such subtitles prioritising legibility and often contain more than one colour within one caption. These captions may not even be faithful representations of the spoken elements of the programme. The added stylistic features are expected to give rise to enhanced pragmatic effects. This raises the questions of what such effects are, and how they arise.

A few attempts have been made to shed light on the use of telop, and scholars such as Shiota (2003) and Kimura *et al.* (2000) present taxonomies of types and uses of telop. While there is no agreed taxonomy for telop, it is often claimed that telop are either verbatim representations of utterances, or an interpretation or summary of discourse, and that they are used to add extra effects as well as facilitating viewer comprehension. Another common acknowledgement is the presence of mediators who come between the speaker and the viewers (e.g. Sasamoto 2014, O'Hagan 2010; Shiota 2003). O'Hagan (2010) and Sasamoto (2014) highlight that telop are inserted in order to guide the viewers' cognitive interpretation processes to suit the

[4] These are intralingual subtitles known as subtitles for the deaf and the hard-of-hearing (SDH) which the viewers can turn on and off, hence "closed" as opposed to open captions which the viewers cannot turn off.

programme directors' needs for enhanced viewing experiences, rather than merely being a comprehension aid for viewers.

Departing from descriptive typologies of telop, O'Hagan (2010) presents a case study of a popular game show in Japan and concludes that telop have a humour-framing function which a director can exploit in order to emphasise a humorous effect. This can be done, for example, by "dramatizing the trivial" (O'Hagan 2010: 85). Drawing upon this study, Sasamoto (2014) has investigated the relationship between telop' framing function (or 'highlighting function') and the viewer interpretation process. Sasamoto argues that telop are a highlighting device that a programme director can use to draw the viewers' attention to chosen elements for a specified purpose such as to induce laughter. In so doing, programme directors can ensure that there is a common platform that enables TV directors to guide viewers in processing programmes in ways that suit their – not the viewers' – needs. The relevance-theoretic notions of the Cognitive and Communicative principles of relevance (Sperber and Wilson 1986/1995) also justify this analysis, as these principles predict that the use of telop contributes to relevance by raising cognitive and affective mutuality by way of shared assumptions and affective responses. That is, ostensively-used telop demand the viewers' attention. This is empirically shown by an eyetracking study which demonstrated that the viewers cannot help but pay

attention to telop (Sasamoto and Doherty 2013)[5]. As a result, TV programme directors can assume that viewers are extremely likely to process what the directors want them to. This ensures viewers access assumptions related to the highlighted (i.e. captioned) elements. In other words, the purpose of using telop is to manipulate the interpretation process by urging viewers to search for relevance in a way that suits the TV producer.

Let us now see in further detail how telop are used - see Figure 2.

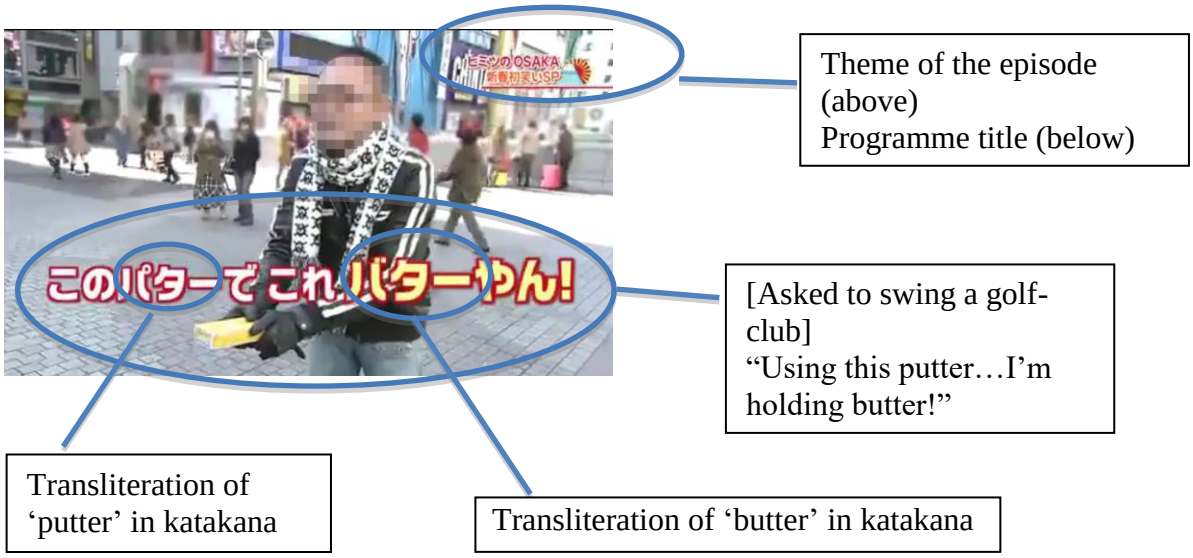


Figure 2: A typical telop use in an entertainment programme (Kenmin Show, Shinshun Special, 2013)

[5] Sasamoto and Doherty (2013) conducted an eye-tracking experiment with 16 native speakers of Japanese and demonstrate that viewers do pay attention to captions regardless of information communicated by such captions' content

Figure 2 is a scene taken from a special episode of a popular TV series, the *Himitsu no Kenmin Show*. The theme of this TV show is regional culture, and most often concerns food. In this particular episode, they were reporting on the comedy-loving, humorous nature of people from Osaka. The TV crew asked members of the public to mimic the action of playing golf on the street but gave them a selection of objects which were not golf clubs. Figure 2 shows a passer-by who was asked to participate with the TV crew saying to him: "Can you swing this putter?". The only problem is that none of the objects offered to him is a putter. Just before the moment in the scene, which features the punchline, the man is given a block of butter, which the crew had suggested he should swing. He pretends to swing the butter while saying "using this putter…" and stops halfway through his swing, declaring "I am holding butter!". The verbatim representation of this utterance is given in the telop at the bottom of the screen. This telop highlights this particular utterance, by focusing on the word play contrasting between the word "putter" in white lettering and the word "butter" in butter-like golden-yellow[6]. By highlighting this particular element with the telop, the TV director ensures that the viewers' attention is drawn to the same points as s/he wants to stress, triggering an affective response by the viewer to the obviously humorous situation. This increases the likelihood that the at-home viewers will react in the same way as the studio guests, who are also

[6] Note that Japanese uses transliterations of the English words *butter* (バター, *batā*)and *putter* (パター, *patā*, a type of a golf club) as loan words, which leads to a pun in Japanese in the same way it does in English.

watching this clip and finding it funny. In other words, in some cases, telop are associated with affective communication (c.f. Blakemore 2008). We will return to this point in Section 3.

2. Typographical Features, Stylistics and Multimodal Analysis

2.1 Suitability of typographical features

The nature of TV medium means that telops are used simultaneously with other stimuli delivered to the audience via different communicative channels. Every communicative stimulus, such as an utterance, sound effect and even the choice of font, seems to interact with the others, forming an overall multimodal content. On the one hand, telop can be treated as a single unit (mode) while it, on the other hand, is itself multimodal, if counting typographical features such as colours and fonts that are embedded in a caption just like a matryoshka doll.

As shown in Figures 1 and 2, telops are often displayed with various typographical features, availing of the use of different fonts, colours and backgrounds. It is generally acknowledged that such typographical choices influence the way a message is interpreted by a recipient, as metaphorical associations in the recipient's mind play a significant role in visual communication (Van Leeuven 2005). Indeed, Meier *et. al*. (2007, 2008) explore links between font and affect, and found that people often associate a positive attitude with brighter colour (Meier *et. al*. 2007), and a positive attitude with a bigger font (Meier *et. al*. 2008). Lakens *et. al*. (2012) also

found that there is a link between negative words and the colour black. Such links, or *embodied affect-space links* (Bornstein and Becker-Matero 2012), is based on the assumption that 'there are inborn, pre-existing connections between affect and bodily experience' (*ibid*. 95). This would suggest that humans are capable of exploiting such metaphorical associations to create similar effects (such as the brighter colour evoking a positive attitude) by presenting a particular visual cue.

Indeed, many scholars report the metaphorical association between font and affect and suitable uses of various fonts in specific contexts (c.f. Doyle and Bottomley 2008, Henderson et.al. 2004). For example, the use of the playful font 'comic sans' in a legal letter is likely to look incongruous with the formal nature of the intended communication. At the same time the use of such a font is likely to be perceived suitable for a child's birthday party invitation. On television, a producer typically takes into account the nature of a message and matches it to a selection of fonts. Indeed, there seems to have emerged a convention in which certain font types are associated with certain types of text and communication. The typographical selection of telop could therefore reveal a link to the intended affective responses which certain captions are expected to trigger in the recipients (as indicated above).

Issues with font selection have been discussed in various disciplines, including Audiovisual Translation (AVT), Web studies, and aesthetics. In AVT, it is generally agreed that sans-serif fonts such as Arial and Helvetica are preferred to serifed fonts for subtitles because of their better legibility (Díaz Cintas and Remael 2007: 84). On the other hand, in a more general font design context, serifed typefaces (serif) are preferred for longer sentences and, hence, for the main body of a text, while sans-serif is preferred for headings (c.f. Takahashi and Katayama, 2012). Takahashi and Katayama (2012) explain that serifed typeface contributes to readability while sans-serif typeface promotes visibility. Indeed, as we will see in Section 4, serifed fonts (MS Mincho being a typical example of serifed font for Japanese) are used for a longer telop employed to add experts' explanations while san-serif (MS Gothic being a typical example of san-serif for Japanese) is used for telop that is added to enhance humour such as ones in Figure 1 and Figure 2.[7]

2.2 Social semiotic approach to typography: Multimodal analysis of fonts

In the previous section, we saw how studies in AVT and web-studies focused on the suitability of fonts in various contexts. In contrast, scholars such as van Leeuwen (2005, 2006) describe typographical features as

[7] Note that each TV station typically has an in-house font for telop and rarely uses fonts widely available on Windows such as MS Gothic or MS Mincho.

semiotic resources for meaning-making from a multimodal perspective. Scholars with a special interest in multimodality and multimodal analysis (e.g.van Leeuwen 2005, 2006; Stökl 2005; Norgaard 2009) see typography as a mode, which has its own grammar and lexis, and often take a descriptive approach to gain insight into the anatomy of typography. Van Leeuwen (2006:151) provides the system of a network of distinctive features such as weight, expansion, and curvature of letters which have the potential to be developed into "grammar". Similarly, Stökl (2005) presents his grammar of typography, again, based on the distinctive features of each letter. The main aim of these multimodality scholars is to provide a systematic theory of "meaning making", and they argue that typography, which can be presented as a mode that has a system, or grammar, has three "metafunctions" as devised by Halliday (1976). For example, using the image of the words *DEAD MAN* written in a font that resembles bones, van Leeuwen (2006) argues that the bone-like font plays an ideational metafunction, which Halliday (2007, 183) defines as "content function of language". He also demonstrates that a change from standard to italics, or standard to bold could have an interpersonal metafunction, which is understood as the function of language which expresses the relationship between the speaker and hearer. For example, "ice" written in bold and upper case, as in "**ICE"**, may be taken as communicating a demand, while it may be taken as a question if written in italics (with a question mark), as in "*ice?*". Finally, van Leeuwen (2006) argues that the use of bold and

different colours, which is often seen in magazine articles, demonstrates a textual metafunction, which is the last of three meta-functions Halliday proposes. Textual metafunction manages how ideational and interpersonal functions are linked and presented in text (cf. Halliday 2007, 184).

2.3 Issues with the multimodal approach

As we have seen, multimodal approaches provide a description of the set of tools available to the communicator. And indeed, a framework based on multimodal analysis of typography provides us with an insight into how the communicator can choose semiotic resources for meaning making (using bold for emphasis, for example). However, multimodal analysis does not *explain* how readers recover the intended meaning, or how the use of these tools on screen gives rise to certain stylistic effects for viewers. In addition, the "semiotics" of each typographical feature is not always guaranteed as it can vary from context to context. For example, italics may be used for an emphasis in one context but as a marker of a quotation in another. For this reason, it is not a matter for semiotics or semantics to provide an explanation. That is, it is not entirely clear how multimodal description or classification could enable us to explain the context-specific nature of typographical features.

Furthermore, it is not clear whether we need such artificial levels of, or even psychological motivation for such "meaning" as ideational, interpersonal and textual metafunctions in order to explain how we interpret

communicative stimuli such as utterances, let alone to account for the features of typography, which are part of written stimuli. For example, while the use of certain fonts is likely to influence a reader's or viewer's interpretation, it is not clear how each kind of influence can be distinctly explained. For example, if, as van Leeuwen (2006) argues, a bone-like font plays an "ideational" function and leads the reader to recover a sense of death, while the use of bold plays "interpersonal" function which leads the reader to recover a sense of "demand", how is it that typographical features play a particular role at different levels of communication, and how should the reader or viewer ascertain which role and level to select? Or, is it the case that the role of typographical features can be explained independently of these so-called metafunctions? As it stands, current multimodal analysis does not provide answers to these questions. In the next section, we will see how relevance theory, a cognitively grounded theory of communication, can provide a theoretical architecture in which the relationship between "meaning making" tools and pragmatic effects can be accounted for.

3. Relevance Theory and Pragmatic Effects

So far, we have seen how typography has been accounted for in various approaches. We have also seen that there is little agreement as to how typographical features affect overall meaning despite the consensus that typographical features *do* affect overall meaning. However, if we look at phenomena other than typography, problems with overall stylistic effects

are nothing new. In the traditional, descriptive literature, intuitions about the relationship between style and linguistic form were often accounted for by distinctions such as given-new and theme-rheme (c.f. Halliday and Hasan 1976; Green 1980; Reinhart 1981). In relevance theory, the so-called "stylistic effects" are explained as pragmatic effects, which are communicated as a result of the use of a particular linguistic form. One of such cases which are often discussed is emphasis[8] (cf. Sperber and Wilson 1986/1995, 2002 and Clark 2013). Consider example (1):

(1)

a) Alfie plays the drums.

b) What Alfie plays is the drums.

c) It is the drums that Alfie plays.

In (1), all three examples communicate the fact that Alfie plays the drums. However, the use of different linguistic structures puts an emphasis on a different constituent. In contrast with (1a), which seems most 'neutral' (i.e. without any particular emphasis), the emphasis is on Alfie in (1b), while in (1c), the emphasis is on the drums. In contrast with the traditional approaches to stylistic effects, which often rely on artificial levels such as theme-rheme, topic-comment, given-new, presupposition-focus, and background-foreground distinctions, in relevance theory, stylistic effects emerge from a natural linkage between linguistic form and pragmatic effects. Sperber and Wilson (1995, 2002) explain that these stylistic effects are the

[8] This is not to say that all cases of stylistic effects are linked to word order.

result of the speaker exploiting structural features of utterances to convey the intended interpretation at a minimum (justifiable) cost. They argue that linkages between linguistic form and the overall interpretation such as in (1) above can be explained in their cognitively-grounded relevance-theoretic framework, as "given that utterances have constituent structure […] and given that they are processed over time, the most cost-effective way of exploiting these structural features will give rise to a variety of pragmatic effects" (Sperber and Wilson 2002, 202-217). In other words, the speaker exploits whatever structures are available to him, in order to communicate at a minimum cost possible. Similarly, Wilson and Wharton (2006) demonstrate that so-called stylistic devices such as stress patterns can be used as natural highlighting devices to draw attention to particular utterance constituents, guiding the hearer to an optimally relevant interpretation in those contexts.

While relevance-theoretic stylistics enables us to account for how "particular linguistic formulations give rise to particular kinds of effects" (Clark 2013, 352), the focus has largely been on verbal communication. The linkage between non-verbal features of communication and stylistic effects is under-researched, except, perhaps, for Wharton's (2009) work on non-verbal communication, Yus's (2009) work on multimodal metaphors and Scott and Jackson's (this volume) work on typography. Scott and Jackson's work in this volume is particularly relevant to the current study. Taking a relevance-theoretic approach, they examine a range of typographical

features, capitalisation and italicisation in particular, and argue that the typographical choices are driven by the author's communicative intensions and considerations for relevance. Like Scott and Jackson (this volume), we also consider the role of typography as the guide for the reader to reach the intended interpretation. However, with increased interest in multimodality, it is imperative that we take into consideration the fact that communication more often than not comes in more than one mode. It is thus important to examine how ostensively used multi-modal stimuli, such as telop on Japanese TV programmes, interact with other programme elements, and influence the viewer interpretation process.

Blakemore (2008) examines a range of apposition cases from a variety of texts and explains how apposition gives rise to non-propositional effects. These effects are ineffable and will be lost under paraphrase. Blakemore (2008, 5) defines affective communication as "communication of [non-paraphrasable] impressions, or non propositional effects". According to Blakemore, while cases of appositions can be analysed in terms of "an impression of amplification or heightened vividness" that they give rise to, the effects are not always derived in the same manner in each case. She demonstrates two ways in which the amplifying effects can be derived:

(2) He felt *depressed, flattened*. (SEU w.1.16.6.239-40. Cited originally by Meyer 1992:67)

(3) He *made a complete mental retreat*; *went far away*. (Maurice Gee, *In My Father's Den*, p.171).

(examples from Blakemore 2008: 19)

According to Blakemore (2008), example (2) is a case of INTENSIFICATION, where the first and second segments of the apposition both provide the audience with access to distinct sets of assumptions. The audience is hence guided to compare the effects that each segment gives rise to. For example, in (2), the first segment *depressed* yields a representation of a state of mind in a general sense, while the second segment *flattened*, communicates a similar state of mind, but perhaps a more intense feeling. Notice that in (2), *flattened* is used figuratively, i.e. we see the use of language that has gone through the process of lexical adjustment (in this case, broadening) and "its encyclopaedic entry includes representations of a state of mind" (Blakemore 2008:24). The second segment thus encourages the audience to access assumptions related not only for a person's mental state but also those related to the encoded meaning of *flat*. It is the audience's responsibility to decide what these properties are, based on the comparison between two concepts derived from the two segments the speaker produces. In so doing, the speaker can provide the basis for deriving an ad-hoc concept, which would yield a range of weak implicatures, giving a more vivid representation (and hence a more intensified representation). In other words, according to Blakemore (2008), intensification type appositions are a case of the speaker guiding the audience to access two sets

of assumptions, where the first might yield cognitive effects yet the intended interpretation should be derived from the second set. The relevance of such cases lies in the way that these appositions enable the speaker to draw attention to the differences in the range of weak assumptions that can be derived from the first and second segments.

In contrast, example (3), according to Blakemore, is analysed as a case of HYBRID REPRESENTATION. Here, unlike the case of intensification where two concepts are processed separately and compared with each other in order to give rise to an intensified interpretation, the two segments are combined for the communication of a single intensified concept rather than two individual concepts. In such cases, the range of implicatures that are derived from the combined concept cannot be recovered from each segment processed independently. According to Blakemore (2008), the sequence in (3) is intended as a description of someone accused of a terrible crime. Here, both *a complete mental retreat* and *went far away*, due to the word *mental* in the first segment, will result in interpretations where the character is understood to seek mental refuge instead of physically going away. However, according to Blakemore (2008), by providing the second segment, the writer can guide the audience to further explore the context to recover a wider range of contextual assumptions. In the case of (3), the second segment would encourage the inclusion of assumptions about distance. Blakemore (2008, 27) explains that the first segment will lead readers to access concepts such as WITHDRAW,

RETIRE, REFUGE, GO BACK, while the second segment would lead readers to extend the context on the basis of contextual assumptions associated with *far* away, which then creates the sense of distance between the character and the narrator. As the audience is given access to these two distinct sets of assumptions, they will be able to explore these assumptions which will then result in the recovery of the wider array of weakly communicated assumptions. This wider range of assumptions cannot be retrieved on the basis of each segment taken independently, hence Blakemore's term HYBRID REPRESENTATION. The combined representation will enable the audience to derive a wider range of weak implicatures and can thus lead to more vivid representations.

Blakemore's (2008) analysis is focused on apposition cases and their affective interpretation and thus may seem irrelevant to cases of typography and their effect on viewers. However, the fact that both the apposition cases and the typography cases involve the communication of intangible feeling (and hence affective communication) makes it worthwhile to explore the similarities between the phenomena. Blakemore's (2008) analysis of apposition and affective communication demonstrates how such stylistic meanings can be explained in general terms without the need for being treated as a special case of communication. The analysis thus suggests that the choice of typography and the consequences it has for communication, does not require a treatment as a special case of communication either. It is, rather, a case of exploiting whatever structure (or "semiotic resources" in

multimodality terms) the communicator has at hand to achieve the intended effect at a minimum justifiable cost; in a similar way that a writer takes advantage of apposition. Of course, there is a fundamental difference: while Blakemore's apposition cases are single-modal and involve appositions of two linguistic representations, the typographical choice, which is the focus of this study, is multimodal and will not have a concrete linguistic representation as such. However, a typographical choice can easily be treated as a communicative stimulus, which will evoke a certain interpretation, as it gives quick-and-dirty access (or a shortcut) to a set of assumptions that are associated with the particular typographical feature[9]. For example, in some cases, typographical features may provide an extra access point to a wider set of weak implicatures in a similar way to hybrid representations, i.e. when combined with the linguistic input delivered by the base caption. That is, it may be possible to argue that the very nature of being multimodal is what enables telop to deliver a hybrid representation. This point will be illustrated later in the next section.

4. Case study: Theoretical explanation and reception study of viewer behaviour

4.1 Research Question and Research Design

[9] In some cases, it may be a set of assumptions that are extremely weak which amounts only to communication of impressions (c.f. Sperber and Wilson 2015)

As stated in Section 1, the aim of the current study is to analyse the effects that typological features have on viewers' interpretation processes, namely, how ostensively used, multi-modal stimuli such as telop interact with other programme elements, and influence the viewer interpretation process. In order to answer this question, we will investigate how typographical features of telop contribute to the interpretation of multiple stimuli and also how typographical features of telop influence interpretations.

We begin with an in-depth observation of multimodal data from a selected Japanese TV programme. Drawing upon our observations and the relevance-theoretic explanation of affective communication (Blakemore 2008), we propose a cognitively-grounded explanation of the role that typographical features play in the interpretation process. We then combine our explanation with some empirical data gained from a preliminary eye-tracking study in Section 5 where our relevance-theoretic explanation will underpin objective physiological data from the eye-tracker.

4.2 Telop and Typographical Features

As illustrated in Figure 3, the use of more than one font and one colour is common in most telops in certain types of entertainment programmes in Japan:

(a)

Caption in sans-serif font in keyword-labelling background

(b)

Telop in sans-serif, in two colours. A half of the telop has a purple 'glow' in the background

(c)

Caption on blackened screen, presented in red and white

(d)

Figure 3: Captions with various fonts and backgrounds (Honmadekka!? 2013)

In Figures (3a) and (3b), sans-serif is used, while serifed font is used in figures (3c) and (3d). In Figure (3a), a keyword labelling the background is used, while a text-box background is used in Figure (3d). Figure (3b) does not use any background, while, in Figure (3c), the whole screen is blackened and used as a background. It is not unusual to see the screen filled with nothing but writing like this; such uses of captions are often seen as a summary panel, or as a cohesive device to facilitate a narrative flow such as a change of topics.

In contrast with conventional subtitles, these telops are presented using vivid colours while varying in font size. Except for Figure (3a), where the caption is written in white, all other captions are presented in more than one colour. Moreover, these captions may use multiple colours as in Figure 1 where four colours were used in a single caption. Furthermore, telops often deploy animations (or text effects):

"I need to see him mentally suffer…"

Figure 4: Telop with various colours and pulse effect (Honmadekka!? 2013)

In Figure 4 above, the female presenter is discussing how she would need to see her cheating lover mentally suffer before she can forgive him. The telop, which is written in black with the 'glow' effect in purple, pulsates as she speaks. Similarly, in Figure 5 below, the telop at the bottom (in red) expands and then returns to its original size (to the size of the first letter, which remains unchanged during the effect) as if to emphasise the anger or irritation.

[As a response to a panellist's off-the-cuff comment]
"Why are you offering that [cold drink] to me!?"

Figure 5: Telop with enlargement effect (Honmadekk

In this programme we counted over 50 occurrences of font changes with over 100 cases of different effects being used within a 20-minute sequence

and this seems to be a typical frequency across other similar entertainment shows. As we can see from these examples, there does not seem to be any limit to the use of 'semiotic resources', and indeed, the TV directors seem to exploit typographical features to the maximum of their potential. As Yamamoto (personal communication) explains, the TV directors are determined to take advantage of any editorial prop in order to create a programme that draws viewers' attention, so that they can stay ahead of the game with different programmes broadcast at the same time. If they did not keep providing new stimuli, captions might lose their effects through familiarity. Bombarding viewers with telop is one of the methds the TV directors employ to keep the programme engaging. This raises the question of what effects the typographical features give rise to and how they are conveyed. In the next section, we will draw on the relevance-theoretic notion of pragmatic effects and non-propositional effects and outline how they can be applied in an analysis of the role of typographical features.

4.3 Typographical Features and Interpretation – Telop and Relevance Theory

4.3.1 "Affective" Interpretation and Typographical Features

This section provides an explanation on how such uses and typographical features of telop can be accounted for in terms of relevance-theoretic stylistics. Writers often use italics or bold to put an emphasis on their meaning. It is therefore not surprising to see similar devices being used to

create the effect of emphasis on telop. The difference is, unlike with two-dimensional writing, TV directors can exploit time and space for such a "form" and add animations (textual effects). Such additional features as vibrations and enlargements of telop can simply be treated as a choice of form that a programme director can exploit in order to give rise to a variety of pragmatic effects.

What is interesting is that such stylistic effects as emphasis, communicated by the 'formal' device such as added animation, are sometimes accompanied by further pragmatic (or poetic even) effects, such as weak implicatures as we see below. (cf. Sperber and Wilson 1995, Pilkington 2000, Wharton 2009). There are also cases where typographical features contribute to the communication of propositional attitudes as we will see in 4.3.2. But let us first address the case of weak implicatures. Recall Figure 4, in which a telop is used to display the presenter's desire to see her lover suffer. This telop is presented in a font that is typical of the Japanese horror manga genre. It also uses a darker colour, compared to other captions. It is in black and has a purple glow around it. In contrast with typical stylistic effects derived from the use of particular linguistic forms (c.f. Halliday and Hasan 1976; Green 1980; Reinhart 1981), the use of such fonts and colours does not communicate a specific assumption such as "emphasis". Instead, it communicates a more intangible feeling, or an

impression[10], such as those often associated with such emotions as JEALOUSY, GRUDGE or REVENGE, just like the effects the hybrid representation cases discussed in Blakemore (2008) give rise to. The presence of weak implicatures is not limited to verbal communication. Wharton (2009), for example, shows how gestures and facial expressions can be used to convey a range of weaker assumptions. In a very similar way, the use of a particular font colour enables the TV director to guide the audience to recover a range of assumptions associated with horror, or, in this case, a bunny boiler. In addition, this telop has the added animated effect of vibration/pulsing. The use of pulse echoes the feeling of "horror" conveyed by the use of this particular font, resembling a bodily reaction often expected in a tense situation evoking fear. Such exploitation of resemblance in communication has been discussed extensively in relevance theory (cf. Sperber and Wilson 1995; Noh 2000; Wilson 2000; Wilson and Sperber 2012) and indeed, what such added animated effect does is to exploit the resemblance in form to communicate some other representations. In other words, as with Blakemore's (2008) hybrid representation cases, the use of multiple visual effects creates a combined representation which facilitates the communication of a wide range of weak implicatures[11].

[10] See Sperber and Wilson (2015) for a fuller account of the communication of impressions.
[11] As mentioned in Section 3, unlike Blakemore's hybrid representation cases, interpreting telops as 'hybrid representations' involves comparison of representations between different modes due to its multimodal nature.

Similarly, in Fig. 5, we can see the use of an animated effect (enlargement) and a red font colour. The utterance which the telop represents is produced by the host of this programme, who is one of the most popular comedians of all time in Japan. When an expert panellist on this programme explains how cold drinks stop people from being selfish, one of the studio (celebrity) panellists says “it is difficult to make a selfish person drink”, while gesturing towards the host. The host, taking the cue for generating laughter, counters the panellist, saying “Why are you offering that (drink) to me!?”, as if to say that the panellist implied the host was a selfish person. Here the director exploits the structure of enlargement of the telop to emphasise the show host’s “indignation”, and as a result, this emphasis encourages the recovery of certain effects via highlighting devices. The pragmatic effect of emphasis is communicated. In addition to this, the use of a blood-red colour and a dominating typeface communicates a range of ineffable feelings associated with emotions such as anger or danger (that awaits the panellist who angered the powerful host). The font enlargement might also be taken as “shouting”, echoing the feelings of anger communicated by the use of this particular colour.

In both cases typographical features provide quick access to a set of assumptions that are often associated with certain typographical elements. Such assumptions are combined with the representation delivered by the linguistic input provided by the encoded element of the telop. The combined representation would then evolve into a set of assumptions that the audience

is expected to recover. However, this is not to say that these feelings or impressions are always communicated by the use of particular fonts and colours (or coded by these typographical features). In fact, there are cases where the chosen colour does not seem to be associated with any particular emotion, unlike certain colours that are more or less conventionalised (or grammaticalised). Instead, the use of certain typographical features in a specific context would allow the viewers to access a range of assumptions, not just one assigned 'meaning', which will help them recover the intended effects. Recall Fig. 2, where a member of the public was asked to play golf using objects other than a golf club. In Fig. 2, the second half of the telop is coloured in yellow. Yellow is often associated with happiness, creativity, and sunshine in many cultures (Colourmatters 2012). However, in this case it is obvious that the colour of the telop is chosen to match the colour of butter. The use of yellow in this example highlights the fact that it is the butter, not a putter, which the speaker is holding. Unlike the previous examples where typographical features provided quick-and-dirty access to a ready-made set of assumptions about feelings of horror (in the case of Figure 4) or anger (in the case of Figure 5), the use of yellow in this example coincides with one of the encyclopaedic entries for the concept BUTTER. In so doing, the TV director can emphasise humour arising from a feigned surprise expressed by the speaker who contrasts between "butter" and "putter".

Figure 2 was the case where the colour was used not to highlight a certain emotional state, but to highlight a "gap" between an existing assumption (that it is a putter that the speaker is holding) and a new assumption (realisation that it is actually butter that he is holding), and, of course, the wordplay that links "butter" and "putter". By highlighting this gap, the TV director could emphasise the feeling of "surprise" arising from incongruity. This suggests that we cannot assign "meaning" to a typographical feature. For example, the colour red may invoke the passion and dedication of an athlete as opposed to the anger we saw in Figure 5, while the horror font may invoke, in this case, an impression of mocking. Furthermore, cross-modal associations, such as those between colours and emotions, might be influenced by culture-specific assumptions. What we are suggesting here is that viewers will have access to a range of assumptions which are associated with such typographical features and recover a range of feelings/impressions that suit this particular context. It is important to stress that we cannot simply assign the role each typographical feature plays without a specific context. Such features, as multimodality scholars seem to claim, are tools for the communicator to exploit (or "semiotic resources"). However, it seems fair to suggest that typographical features can influence the viewer interpretation process at least in two ways: as a structure (through animation/text effect) that gives rise to a pragmatic effect, and as a pointer for the recovery of a range of intangible feelings or impressions. Our

data shows that the interpretations are context-specific and recovered in line with expectation of relevance.

4.3.2 Propositional Attitudes and Typographical Features

In contrast to cases we have seen above, where the intention is not chiefly to affect explicit content of the utterance, there are cases in which the change of the font can affect the recovery of a higher-level explicature[12] by influencing the interpretation of a propositional attitude. Figure 6 demonstrates this:

(a)

"Obviously you think you are the funniest comedian."

(b)

[12] See, for example, Wilson and Sperber (1993) for a fuller discussion of a higher-level explicature.

"I am so grateful for all the job opportunities given to me."

(c)

[Above the caption in (b)] "Liar"

Figure 6: Telop with font change (Honmadekka!? 2013)

In Figure 6 (a), celebrity guests tease the host for his self-assuredness, producing remarks such as "obviously you think you are the funniest comedian (captioned)". The host responds, in Figure 6 (b), denying the allegation of being so full of himself. He says "I am ever so grateful for all the wonderful job opportunities given to me", which is captioned. Responding to the host's utterance, other panellists say "liar" which is captioned in Figure 6 (c) just above the host's line shown in Figure 6 (b) that is retained in Figure 6 (c).

The typographical features used in this sequence are particularly noteworthy, as not only does the font change from sans-serif to serifed, but the line is now italicized. Although there are no set rules or guidelines as far as telops are concerned, serifed fonts such as Times New Roman are often recommended for official documents such as legal documents or business letters for their "clean and authoritative" nature (Butterick 2008). By swapping from a sans-serif font to a more official looking font, the TV director can be taken as indicating that this is what the host is "required" to state publicly rather than how he actually feels privately. Although this is highly speculative, the change of typography may arguably be providing an additional prompt for the viewer to doubt the host is being sincere, and to conclude that it is his "official" stance. As a result, viewers would recover a higher-level explicature such as the one in (2):

(2) *The speaker (and the mediator) do not believe* [what's represented by the telop]

This suggests that the typographical features of a written stimulus can sometimes affect the explicit content of an utterance, and in this case, the change of font affects the interpretation of a higher-level explicature, guiding the viewers to a certain propositional attitude.

5. An Empirical Approach to Telop

So far, we have argued that typographical features influence the way the viewer would interpret the telop, and that the use of typographical features

gives rise to a range of pragmatic effects. The question now concerns how we could develop this theoretical analysis further to demonstrate the influence of non-verbal stimuli in multimodal contents which the audience is exposed to. In this section we provide initial empirical support for this relevance-theoretic analysis of typographical features. In essence we sought more empirical evidence to answer the following main research question: how do ostensively used multi-modal stimuli such as Japanese telop interact with other elements in the same programme, and influence the viewer interpretation process? To answer this question, we would first need to substantiate the assumption that typographical features do affect the audience. To this end, we conducted eye-tracking experiments, albeit on a pilot basis. Eye-tracking technology enables us to evaluate audience reception of multimodal content in an objective manner, and eye-tracking studies commonly involve measuring of the cognitive effort[13] of the subject, i.e. how much visual attention the subject pays to the stimuli. However, in our study the eyetracking data was sought to provide some empirical support for our theoretical discussion. So, the data was treated as supplementary evidence and analysed in a qualitative manner rather than providing answers quantitatively.

5.1 Methodology: Eye-tracking

[13] It is true that no agreement has been reached over a universal measure of cognitive processing effort (see Just and Carpenter, 1993; Schultheis and Jameson, 2004)

We conducted pilot experiments using the Tobii Eye-Glass, which is a 30 Hz, monocular eye tracker. Sixteen participants (consisting of fourteen female and two male subjects), all native speakers of Japanese, took part in this experiment[14]. The mean age of the participants was 20.6 years with all of them having normal or corrected-to-normal vision. No motor or neurological abnormalities were reported. The experiment was approved by, and conducted in accordance with the regulations of Dublin City University's Research Ethics Committee. A short clip (duration = 22 minutes: 99 seconds, including 90 seconds of TV commercials) was prepared with a Japanese primetime programme called Honma dekka!? TV (Fuji TV, July 13th, 2013), which is a popular variety show with a panel of experts and celebrity guests. The clip was displayed on a 32 inch TV screen. The experiment was conducted in a simulated living room setup with environmental controls. The particular type of eyetracker device was used for its portability advantage over a fixed-desktop type eye-tracker to ensure ecological validity.

5. 2 Results: Eye-tracking Data

Gaze data from each participant was visualised and compared. Of the total data collected we had to discard data from 4 subjects due to data corruption and technical problems. Areas of Interest (AOIs) were set for a telop with

[14] It is true that it is a small-scale study. However, eye-tracking studies of this sort typically has high attrition rate (O'Brien 2009).

added effects and font changes. Visualisations were created for the full recording (i.e. the visualisation of the gaze data for the whole duration of the clip) from which we isolated each particular segment of interest (i.e. the visualisation of the gaze data from the duration that the chosen telop were being displayed on screen). Out of 438 telops, 17 were found to have added effects such as vibration, enlargement, or entrance with special effects, such as the ones shown in Figures 4 and 5. We counted only 3 occasions where an italicised serifed font rather than the usual sans-serif, such as the one in Figure 6(b) were used and we classify them as cases of "font change". We therefore categorised the data from the full recording into 4 groups and created separate visualisations for each of the following scenarios: (a) Full recording minus the above 3 cases of font change. This was done by removing the gaze data for the duration of time that the 3 captions which were written in italicised serifed font appeared on screen. This first category therefore contained 435 captions; (b) Recording that contains only the 3 cases of font change; (c) Full recording minus text effects. This was done by removing the gaze data for the duration of the 17 captions that were found to have added animated effects (i.e. vibration, enlargement, and special entrance effect). This category therefore contained 421 captions; and (d) the recorded data for the duration of the 17 captions that had added animated effects. These were then compared for each participant. A 'heat map' was chosen for visualisation as it would represent the intensity of the subjects' visual attention most clearly. In all recordings, there were no deviations (no

extra effect, sound or visual) observed in these particular scenes except for the textual effects and font changes that appeared with the telop at the bottom of the screen.

In most cases, it was observed that the audiences' gaze patterns changed in the segments which contained telops with both font change and animated effect cases in comparison to recordings of segments that do not contain such typographical changes. In Figure 7, we can see the comparison of gaze patterns for 3 selected subjects[15]. In Figure 7, we compared segments of recordings that do and do not contain changes of fonts in telop:

[15] Images of all subjects (except for the discarded data from 4 subjects) are given in the Appendix. Due to the colour restrictions we present the gaze visualisation only.

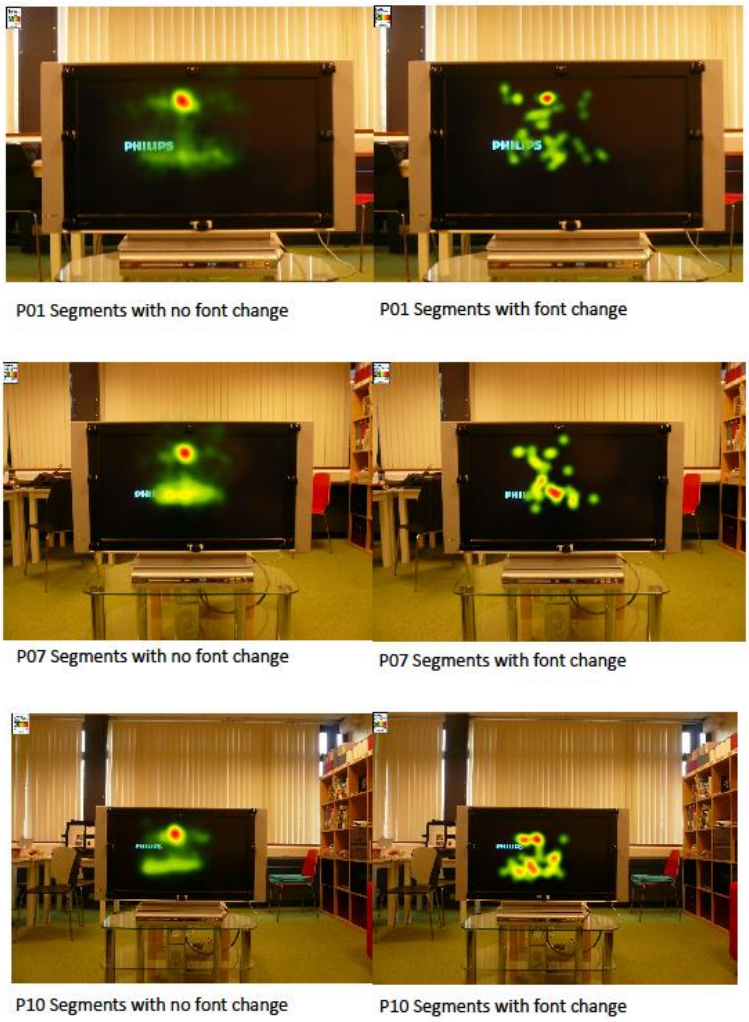


Figure 7: Font Change and Visualisation of Eye-tracking Data

A heatmap visualization such as Figure 7 illustrates where participants' fixation was concentrated. The degree of fixation is shown as hotspots - the colour red shows the most fixations followed by yellow and then green. The label starting with P is an identifier for each participant. In Figure 7, the images on the left show the heat maps for

the recording of segments that do not contain font changes (category a), while the images on the right present heat maps for the recording which contains font changes (category b). In all 3 cases, gaze patterns in the image on the left (recording with no font change) show relatively intense attentions to the upper part of the TV monitor with general coverage in the region where telops are often inserted. In contrast, on the heat maps of the duration with font changes (images on the right), all subjects show more erratic movement of their gazes, which are more widely distributed (and so is the gaze intensity).

Similar changes in gaze patterns are seen in comparisons between the segments which contained telop with animated effects such as vibration, enlargement, and entrance with special effects (category c), and recordings of segments that do not contain such text effect (category d).

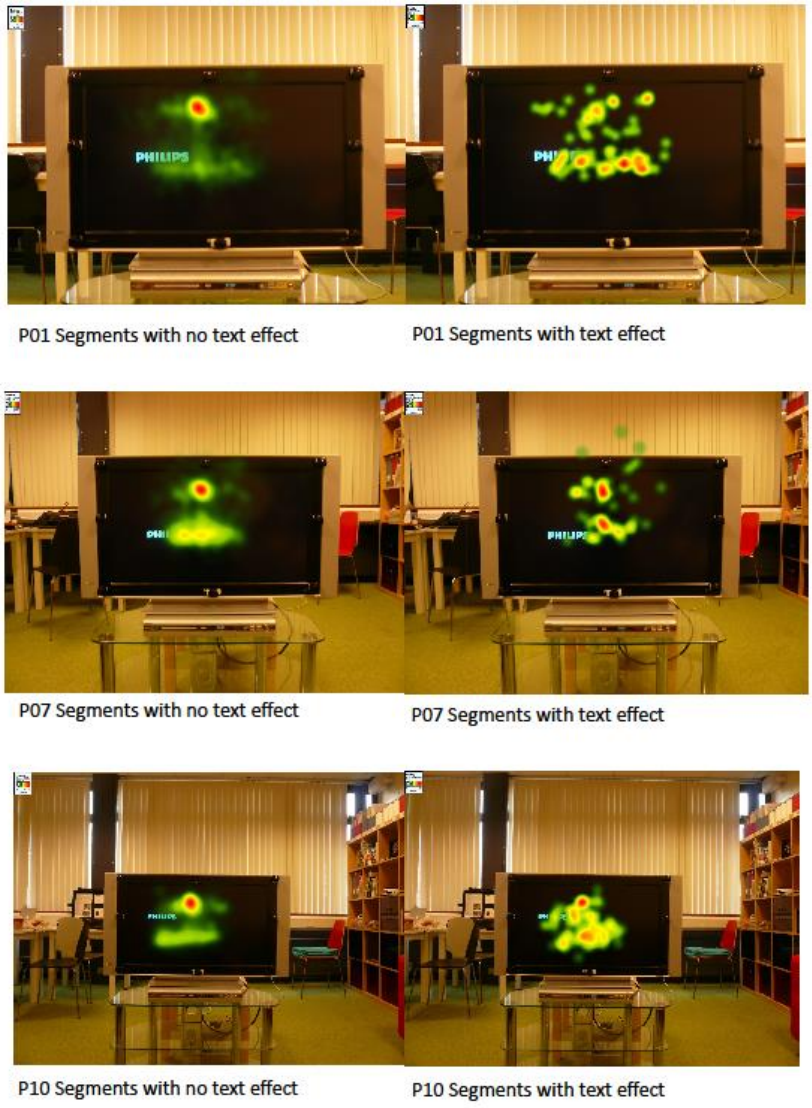


Figure 8: Text Effects and Visualisation of Eye-tracking Data

Here, the gaze pattern is even more erratic, and it seems to move downwards compared with segments with no text effect. The visualisation also shows that the intensity of gaze is more widely distributed, rather than staying in one part of the screen as it does in the images on the left (the heat maps for segments with no text effect).

While these visualisations do not provide supporting evidence in a quantitative sense, they indicate that it may be worthwhile to explore the

change in participants' viewing behaviour. In both Figure 7 and Figure 8, gaze data for the full recording seems to gather around one big cluster around the upper-centre position of the screen, often with much less-intense gaze distribution towards the lower end of the TV screen. Interestingly, the heat map for the full recording, which shows the intensity of gaze throughout the recording, seems to focus on the upper side of the screen rather than the bottom region where telop appear constantly. However, in both cases, heat maps for text effects and font changes seem to show that gaze moves in a more erratic manner than for the full recording, across the screen.

These findings suggest that typographical features could be the factor that influenced the participants' attention at least on the basis of qualitative observations. The result demonstrates how audiences' visual attention is influenced somewhat by typographical features and hence supports our analysis that typographical features do contribute to the viewer interpretation process. The findings from this study support what relevance theory predicts: these more ostensive stimuli, thanks for the layering of typographical features such as font and colour, attract more attention, so viewers pay more attention to them in pursuit of cognitive effects.

6. Conclusion

This paper is concerned with the growing importance of typographical features in the interpretation process in the context of shifts in written

communication due to the use of digital devices. While the "core" of written communication may still rest with verbal inputs, the contribution of non-verbal stimuli, such as typographical features including fonts and colours to communication is not insignificant. With the advances of technology, it has become easier to use various typographical tools such as fonts and colours to help communicate ineffable feelings, which may not be expressible in a concrete verbal representation. We demonstrated how TV directors use typographical features in order to communicate intangible feelings or impression using telops on Japanese TV programmes as a case study. We also demonstrated how the viewers gaze pattern seems to be somewhat affected by such stimuli. It was our aim to answer the question: how do typographical features such as font and colour affect TV viewers' interpretation processes? Japanese TV programmes were ideal as our data source, as these programmes involve complex layers of multimodal representations including the use of outlandish telops exploiting various typographical features.

Drawing upon Blakemore's (2008) analysis of appositions and affective communication, we have argued that typographical choices contribute to affective communication in particular, helping to communicate intangible impressions or feelings rather than one strong (or concrete) thought representation. Typographical features contribute to relevance by functioning as a highlighting device to guide viewer's interpretation process. In some cases, they guide the viewers to recover a range of pragmatic

effects at an affective level (i.e. communication of layers of weak assumptions) by contributing to the recovery of hybrid representations. In other (probably rare) cases, the choice of a particular typographical feature leads the viewer to recover a specific propositional attitude. We thus showed the important role of ostensively-used non-verbal stimuli in communication and closed a gap in the relevance theoretic literature by providing an account of an under-researched stylistic device. The observation from our eye-tracking study, albeit on a pilot basis, seems to indicate that viewers are influenced by the use of certain typographical features, although further research is necessary to ascertain the degree of influence typographical features may have on the viewers' interpretation process.

Acknowledgements

We would like to thank DCU doctoral researcher in Computing, Feiyun Hu, for his continued assistance with the Tobii Glasses, The Research Support office and the Faculty of Humanities and Social Science at Dublin City University for their funding of this project under the Enhancing Performance scheme in 2013, and Dr. Stephen Doherty for collaborating from Sydney for this project.

References

Blakemore, Diane. 2008. “Apposition and affective communication.” *Language and Literature*, **17** (1) : 37-57.

Bornstein and Becker-Matero, 2013. Metaphor and Psychodynamic Research. In Metaphor and Fields : Common Ground, Common Language, and the Future of Psychoanalysis, ed. by Katz S. Montana. 92-103. London: Routledge

Butterick, Matthew. 2008. Font Recommendations. Available from *http://typographyforlawyers.com/about.html*

Clark, Billy. 2013. *Relevance theory.* Cambridge: Cambridge University Press.

Color Matters. 2012. “The meanings of Yellow.” http://www.colormatters.com/yellow

Díaz-Cintas Jorge and Remael Aline. 2007. *Audiovisual translation: subtitling*. Manchester: St. Jerome.

Doyle, John R. and Bottomley, Paul A. 2009. “The massage in the medium: Transfer of connotative meaning from typeface to names and products.” Applied Cognitive Psychology 23(3): 396–409.

Green, Georgia. 1980. “Some wherefores of English inversions.” *Language* 56: 582-601.

Halliday, M.A.K , and Hasan, Ruqaiya. 1976. *Cohesion in English*. London: Longman.

Halliday, M.A.K . 2007. *Language and Education.* London: Continuum.

Henderson, Pamela W., Giese, Joan L. and Cote, Joseph A. 2004. "Impression Management Using Typeface Design." *Journal of Marketing* 68(4): 60-72.

Just, Marcel A. and Carpenter, Patricia A. 1993. "The intensity dimension of thought: Pupillometric indices of sentence processing." *Canadian Journal of Experimental Psychology*. 47 (2): 310-339.

Kimura, T, Hosoi, A., Honda, N., Kato, Y., Kawamura, F., Koizumi, A., Oosawa, Y., Suzuki, E. and Watabe, K. 2000. "テレビ画面に踊る文字たちの生態学" [Physiology of Letters Dancing on TV screen], *Galac 366*.

Lakens, Daniël, Semin Gün R and Foroni, Francesco. 2012. "But for the bad, there would not be good: Grounding valence in brightness through shared relational structures." Journal of Experimental Psychology: General, 141(3), 584-594.

Meier, Brian P. , Robinson, Michael D . and Caven, Andrew J. 2008. "Why a Big Mac is a Good Mac: associations between affect and size." Basic and Applied Social Psychology 30: 46–5

Meier, Brian P. , Robinson, Michael D., Crawford, L . Elizabeth and Ahlvers, Whitney J. 2007. "When "light" and "dark" thoughts become light and dark responses: affect biases brightness judgments." Emotion 7: 366–376.

Noh, Eun- Ju. 2000. *Metarepresentation: A Relevance-Theoretic Approach.* Amsterdam: John Benjamins.

Nørgaard, Nina. 2009. "The Semiotics of Typography in Literary Texts. A Multimodal Approach." *Orbis Litterarum* 64(2): 141-160.

O'Brien, Sharon. 2009. "Eye tracking in translation process research: methodological challenges and solutions." In *Copenhagen Studies in Language: Methodology, Technology and Innovation in Translation Process Research*, ed. by Inger Mees, Fabio Alves, and SusanneGopferich, 251-266. Copenhagen: Samfundsitteratur.

O'Hagan, Minako. 2010. "Japanese TV entertainment: Framing humour with open caption telop." In: *Translation, humour and the media: Translation and Humour*, Volume 2, ed. by Delia Chiaro, 70-88. London: Continuum.

Reinhard, Tanya. 1981. "Pragmatics and linguistics: An analysis of sentence topics." *Philosophica* 27: 53-94

Sasamoto, Ryoko. 2014. "Impact caption as a highlighting device: Attempts at viewer manipulation on TV." *Discourse, Context and Media* 6: 1-10.

Sasamoto, Ryoko and Doherty, Stephen. 2013. "Manipulating the mind of others - the use of subtitles as viewer manipulation." *Cognitive Futures of the Humanities*, Bangor, UK , 04-APR-13 - 06-APR-13.

Schultheis, Hoger, and Jameson, Anthony. 2004. Assessing cognitive load in adaptive hypermedia systems: Physiological and behavioural

methods. In *Adaptive Hypermedia and Adaptive Web-based Systems*, ed. by Wolfgang Nejdl, and Paul de Bra, 18-24. Eindhoven: Springer Verlag.

Shiota, Eiko. 2003. “ 関連性理論とテロップの理解” [Relevance theory and understanding of telop]. *Journal of Ryukoku University*: 63-91.

Sperber, Dan and Deirdre Wilson. 1986/1995. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Stökl, Hartmut. 2005. “Typography: body and dress of a text - a signing mode between language and image.” *Visual Communication* 4 (2): 204-214.

Takahashi T. and Katayama N. 2012. ‘伝わるデザイン：研究発表のユニバーサルデザイン[Design to communicate: universal design for research presentation]’ [Online]. Available at http://tsutawarudesign.web.fc2.com/seisaku.html (accessed 26 June).

van Leeuwen, Theo. 2005. “Typographic Meaning.” *Visual Communication*, 4(2) : 137-142.

van Leeuwen, Theo. 2006. “Towards a semiotics of typography.” *Information Design Journal* 14 (2):139-155.

Wharton, Tim. 2009. *Pragmatics and Non-Verbal Communication.* Cambridge: Cambridge University Press.

Wilson, Deirdre. 2000. “Metarepresentation in Linguistic Communication”. In *Metarepresentations: A multidisplinary perspective*, ed. by Dan Sperber, 411-48. Oxford: Oxford University Press.

Wilson, Deirdre and Sperber, Dan. 1993. “Linguistic form and relevance”. *Lingua* 90(1), 1-25.

Wilson, Deirdre and Sperber, Dan. 2012. *Meaning and Relevance*. Cambridge: Cambridge University Press.

Wilson, Deirdre and Whraton, Tim. 2006. “Relevance and prosody”. *Journal of Pragmatics* (38): 1559-1579.

Yus, Francisco. 2009. “Visual metaphor versus verbal metaphor: A unified account.” In: *Multimodal Metaphor*, ed. by Charles Forceville and Eduardo Uriós-Aparisi, 145-172. Berlin: Mouton de Gruyter.

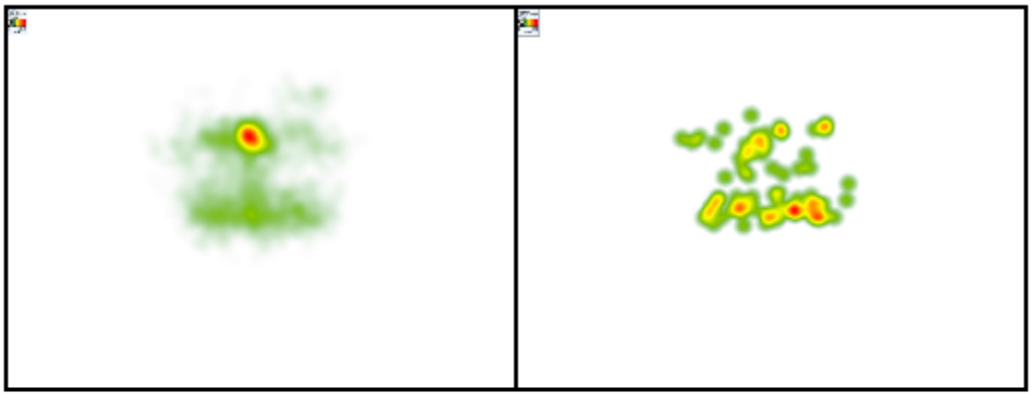
Source of visual material:

Himitsu no Kenmin Show. 2007-2015. Nihon Television Network. Produced by Shinichi Nose, Minoru Sato, and Kie Shimizu. Tokyo.

Honmadekka!? 2009 – 2015. Fuji Television Network. Produced by Michiko Kametaka and Masayuki Harashima. Tokyo.

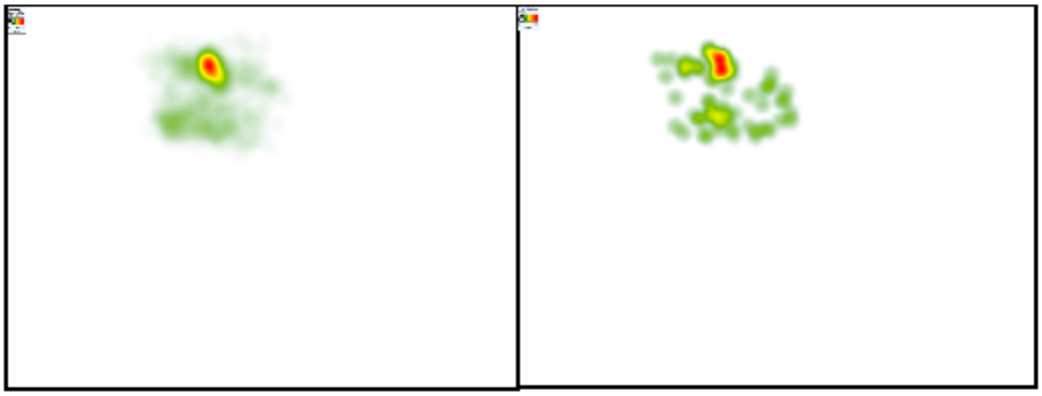
Appendix: Visualisation of all participants

(1) Comparison between segments with no text effects and segments with text effects



P01 Segments with no text effect

P01 Segments with text effect

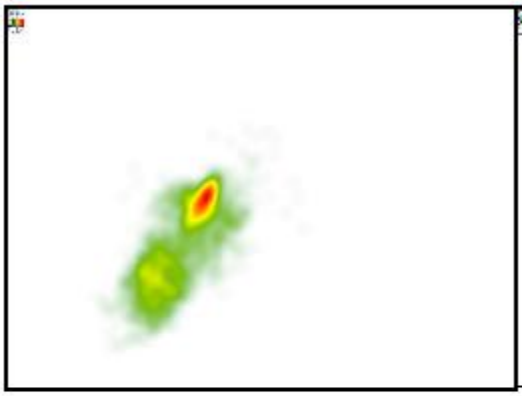


P02 Segments with no text effect

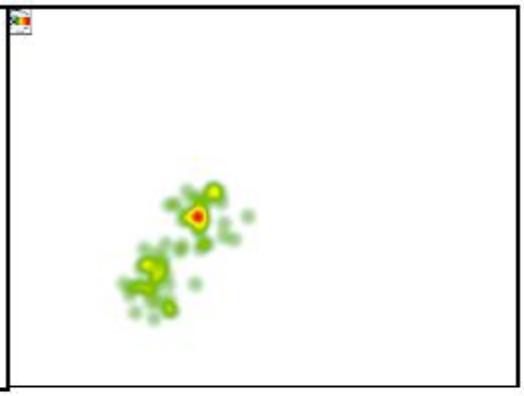
P02 Segments with text effect

P04 Segments with no text effect

P04 Segments with text effect



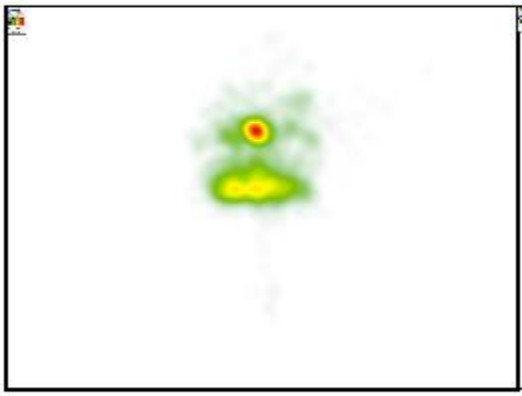
P05 Segments with no text effect



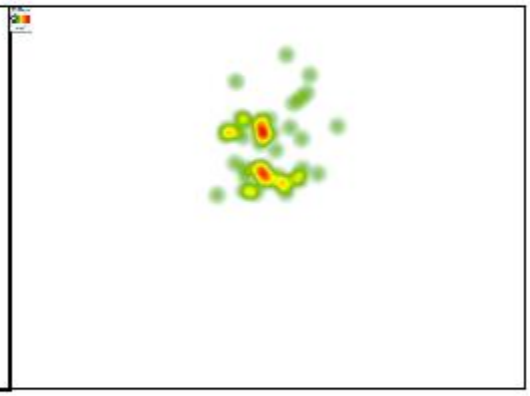
P05 Segments with text effect

P06 Segments with no text effect

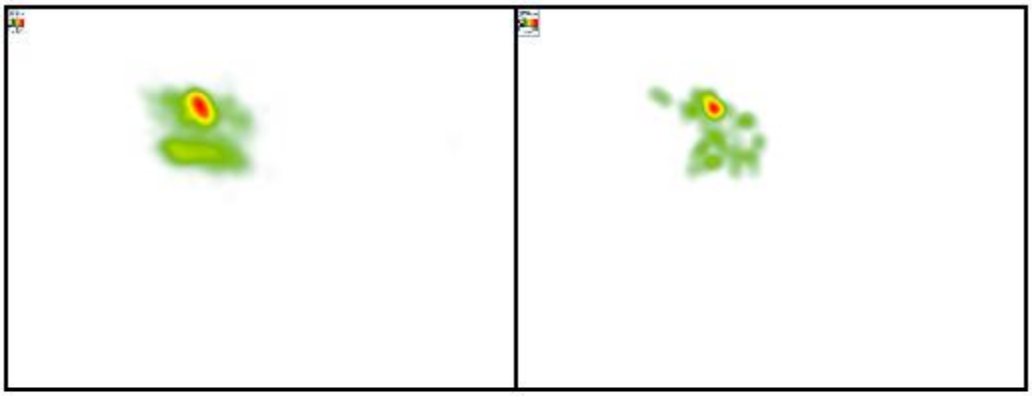
P06 Segments with text effect



P07 Segments with no text effect

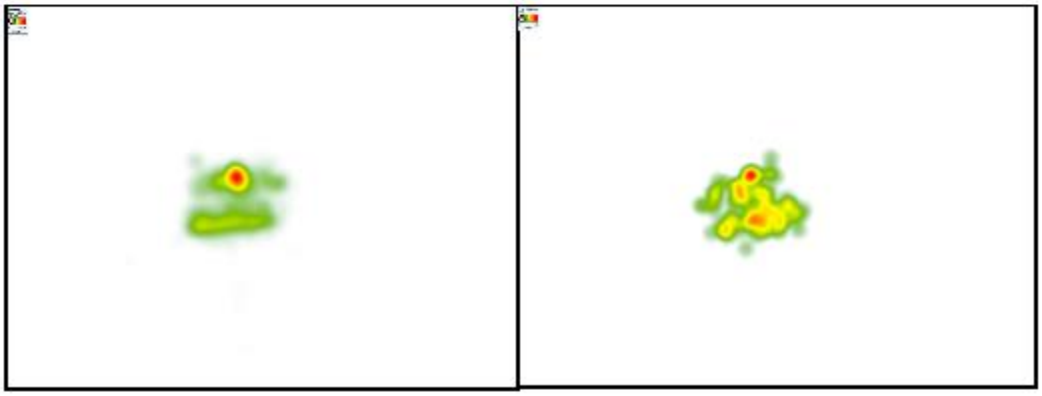


P07 Segments with text effect



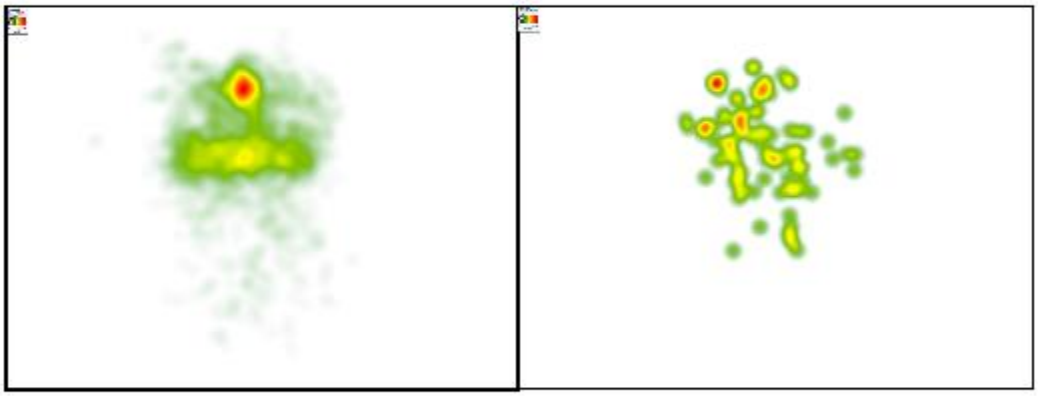
P08 Segments with no text effect

P08 Segments with text effect



P10 Segments with no text effect

P10 Segments with text effect

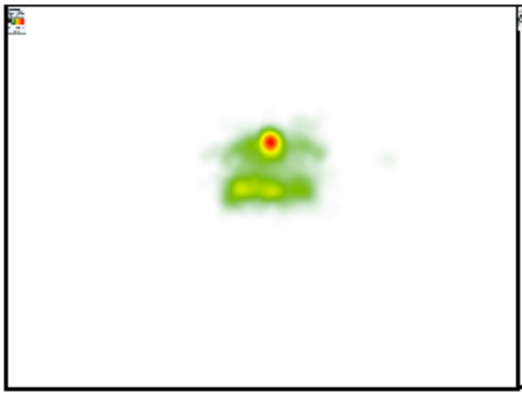


P12 Segments with no text effect

P12 Segments with text effect

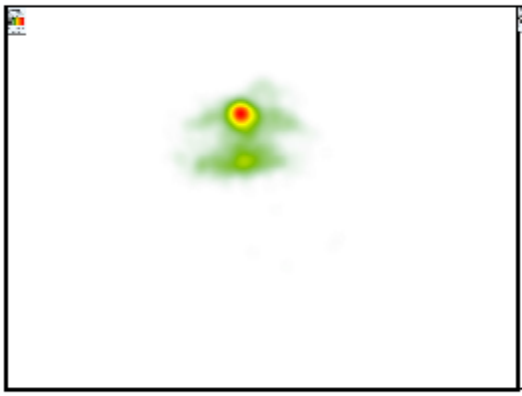
P14 Segments with no text effect

P14 Segments with text effect

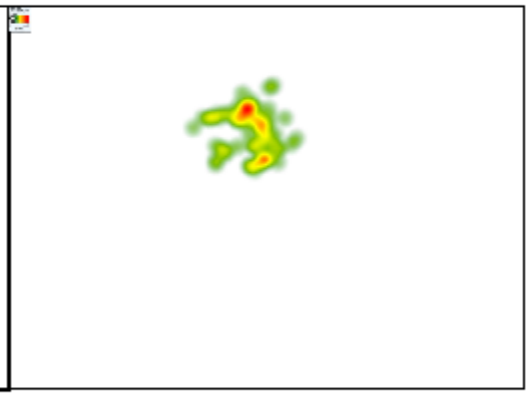


P15 Segments with no text effect

P15 Segments with text effect

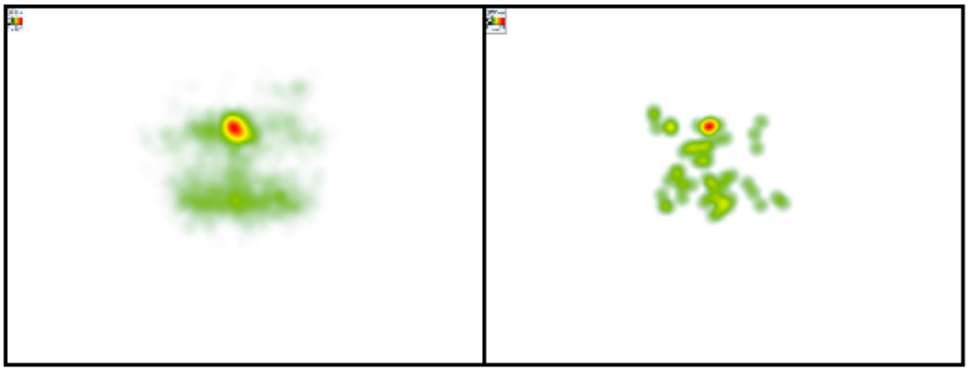


P016 Segments with no text effect



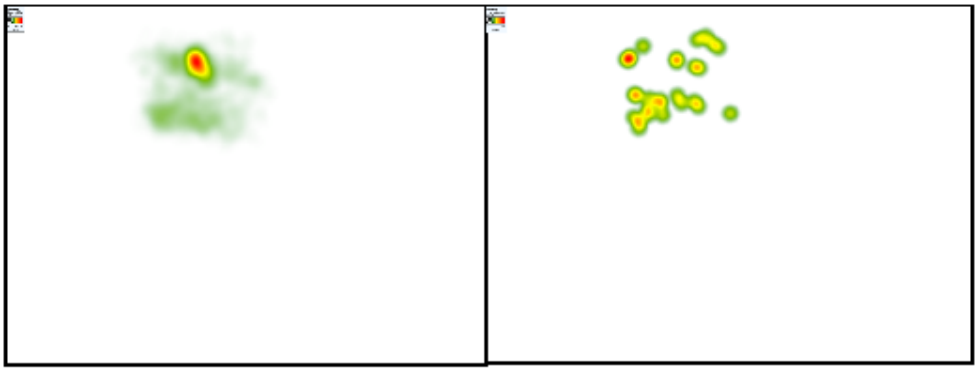
P16 Segments with text effect

(2) Comparison between segments with font change and segments with font change



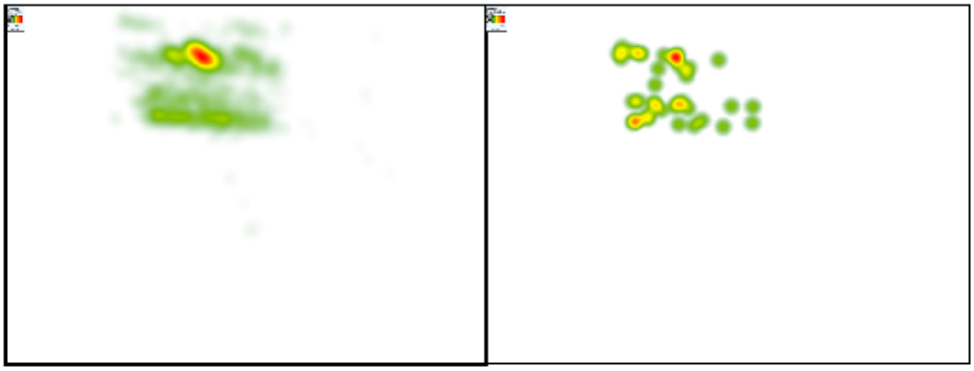
P01 Segments with no font change

P01 Segments with font change



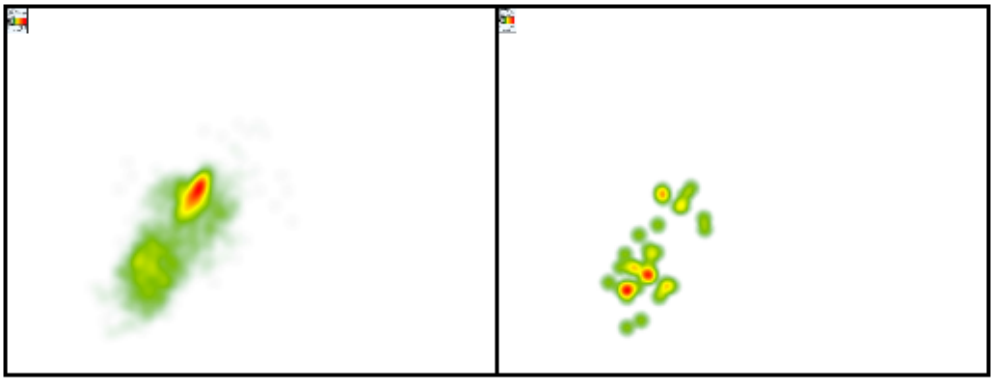
P02 Segments with no font change

P02 Segments with font change



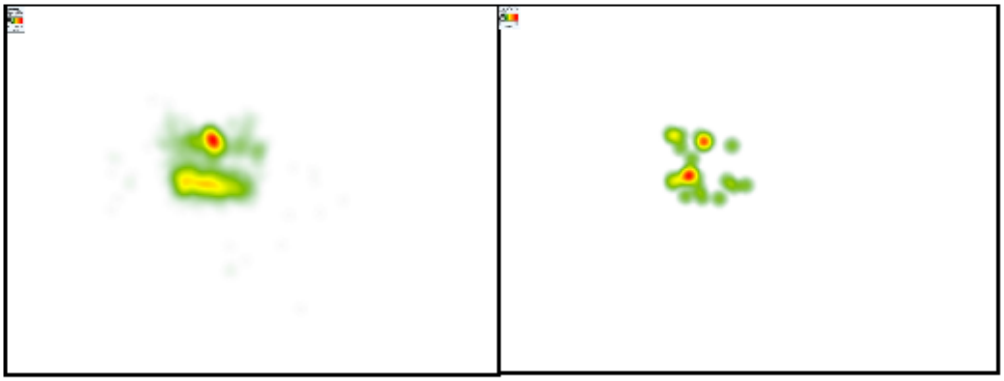
P04 Segments with no font change

P04 Segments with font change



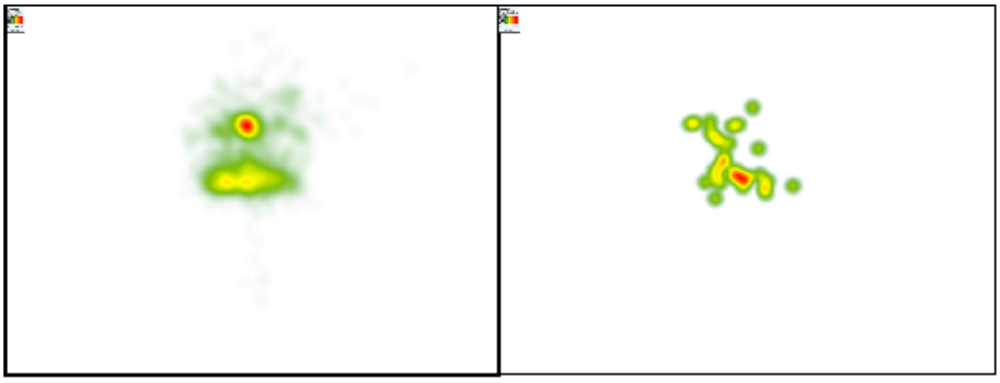
P05 Segments with no font change

P05 Segments with font change



P06 Segments with no font change

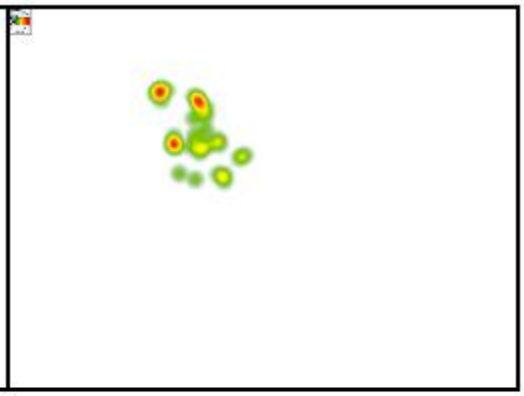
P06 Segments with font change



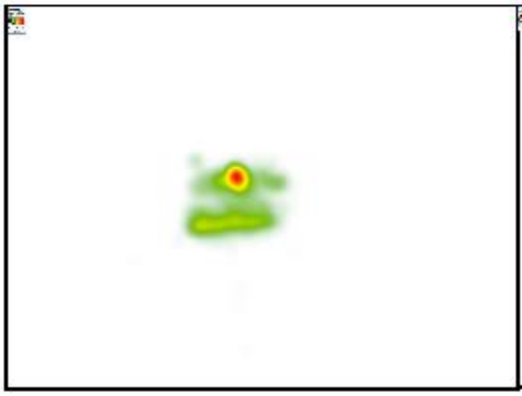
P07 Segments with no font change

P07 Segments with font change

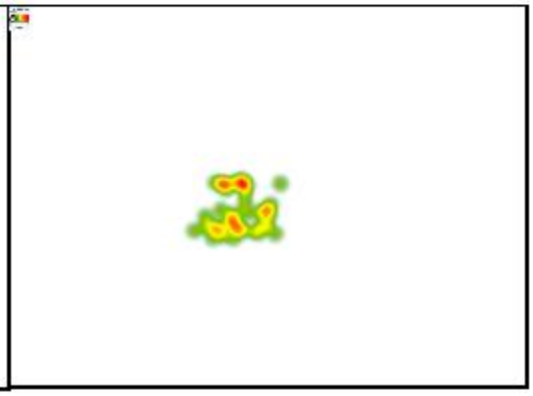
P08 Segments with no font change



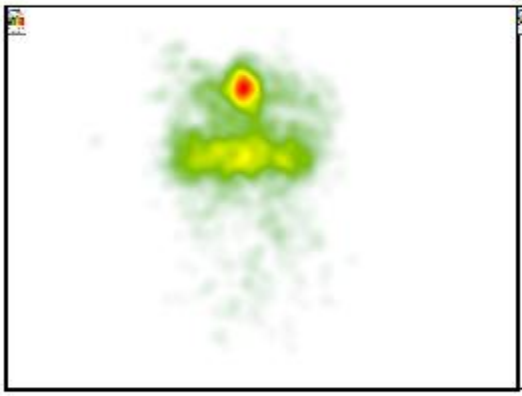
P08 Segments with font change



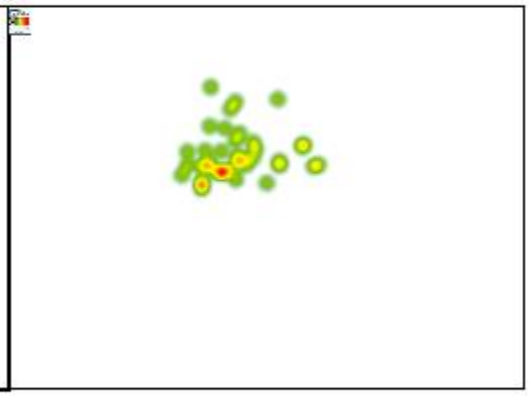
P10 Segments with no font change



P10 Segments with font change

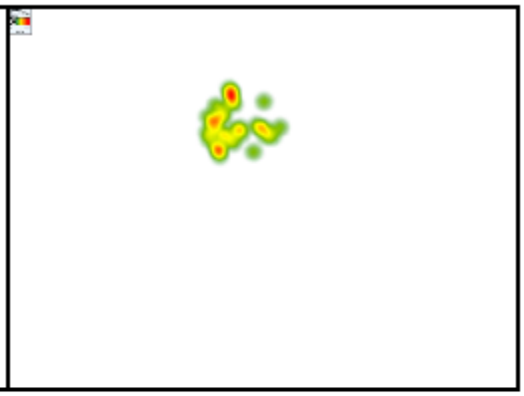


P12 Segments with no font change

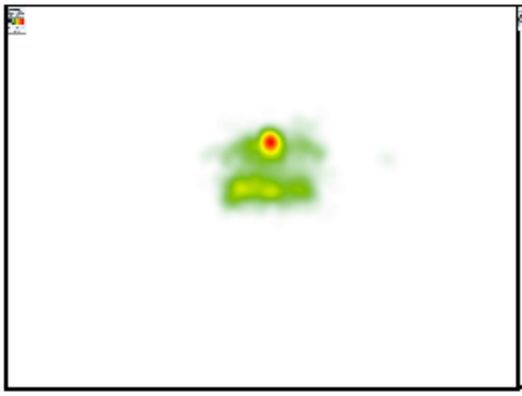


P12 Segments with font change

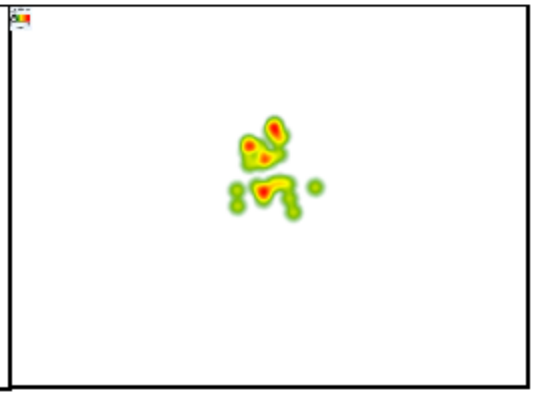
P14 Segments with no font change



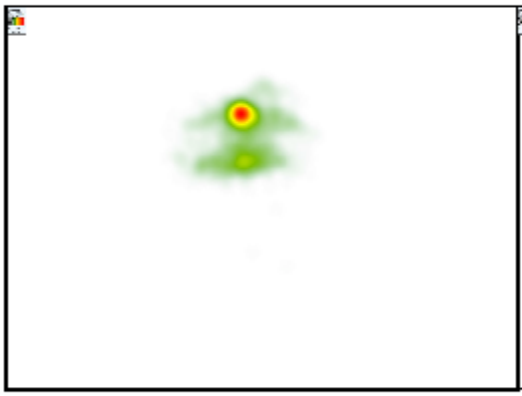
P14 Segments with font change



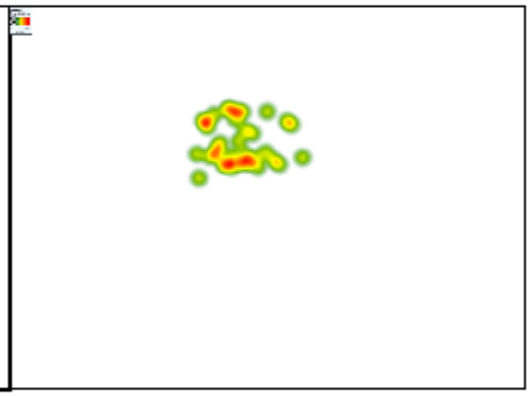
P15 Segments with no font change



P15 Segments with font change



P016 Segments with no font change



P16 Segments with font change