

The relevance of redundancy in multimodal documents

Olli Philippe Lautenbacher

University of Helsinki, Finland
olli-philippe.lautenbacher@helsinki.fi

The aim of this article is to refine the role of redundancy in deferred multimodal communication, from the standpoint of both communicators and their audiences and, by extension, translators. What is advocated here is the idea of a recursive reading process consisting of three phases (perception, construction and integration) and that this process is based on the detection of a salient series of trigger stimuli that the communicator offers as incentives. The shared ground of significance of these trigger compounds actually reveals core meanings in the document, especially when there is exophoric reference. In the translation process, any change within this redundancy system, such as a modification in the balance between endophora and exophora, might alter the overall reception experience.

1. Introduction

The use of redundancy in communication is indisputable. Redundancy occurs in monologue, dialogue and speech, in poetry, lyrics and music, in literature, graphic novels, and movies, etc., in which the voluntary reiteration of salient elements is a rule rather than a mere accident. I would like to argue here that a redundancy system, based on overlapping semantic features between co-occurring items within a document, is not only relevant in multimodal communication, but is in fact essential for it to work properly. Besides redundancy serving as a highly efficient tool in cohesion building or as a guarantee of successful communication, it is also an actual means of creating meaning. This is why the question of redundancy needs to be viewed as particularly important within the realm of translation too. In textual linguistics, the study of syntactic relations as building blocks of discourse continuity offers us a perfect example of the richness of the redundancy phenomenon as a tool for cohesion. As Halliday and Hasan (1976, p. 8) observed, cohesion is created through a semantic relation between two elements of the same discourse, with the first element being crucial to the interpretation of the second. Furthermore, the authors listed different cohesion types (Halliday & Hasan, 1976, p. 13), revealing the complexity of referential relations in substitution, ellipsis, conjunction and lexical cohesion (such as parallelism, paraphrase and total or partial reiteration). For each and every one of these relations, a common ground of reference is needed between the trigger and its antecedent (Lautenbacher, 2014, p. 57). In other words, at least some amount of semantic redundancy, be it only partial, is always at stake in these endophoric relations.

Studies that pertain more exclusively to anaphora distinguish clearly between co-referential anaphors (such as pronouns that refer to the entire antecedent) and indirect—or associative—anaphors (for the difference, see Kleiber, 2001) such as those based on part–whole relations, but where one still fundamentally looks for shared features between the trigger expression and its referential antecedent.

From a broader perspective, it is evident that redundancy is one of the fundamental aspects of embodiment and deixis in face-to-face communication (Mondada, 2016). In oratory speeches or in commercial campaigns, the simple repetition of discourse segments is consciously used as a rhetorical effect (Seliger, 2008). Even more generally speaking, the basic communication principle of Information Theory (Shannon, 1948) demonstrated that redundancy was useful to secure the transfer of meaning by countering the noise inherent in any communication situation.

In the following analysis, I attempt to show how the reading process, with multimodal documents, is based on the detection of a series of salient trigger stimuli that are given to the reader-viewers as incentives to help them to grasp core meanings of the document. This reading process builds on two main types of redundancy: one that is internal to the text, essentially useful for cohesion, and another that is external to the text, used for introducing new meanings.

From a translation viewpoint, we will see how important it is to distinguish between these two forms of redundancy and how cautious translators have to be in manipulating them, because even small changes in the balance of this redundancy system can not only have an impact on meaning to some extent but, perhaps even more crucially, alter the overall reception experience—which, needless to say, would be an unwelcome result in most translation situations.

2. Theoretical background

The theoretical basis for this paper combines two different approaches to reception, although they were conceived of almost simultaneously in the late 1980s. First, Relevance Theory (Sperber & Wilson, 1986/1995) offers a sound general account of the roles of both communicator and audience, mainly in face-to-face communication. Secondly, the Construction-Integration Model (Kintsch, 1988; further developed in 1998) develops an interesting perspective on linguistic comprehension. In this section, I briefly present the main aspects of both approaches pertaining, *mutatis mutandis*, to the role of redundancy in multimodal communication.

2.1 Relevance theory

Relevance theory conceives of communication as a bi-directional activity, but, as in ballroom dancing, one partner must take the lead and the other must adapt his or her steps to follow, using “[w]hatever code and contextual information come most easily to hand” (Sperber & Wilson, 1995, p. 43).

As the authors envisage it, communicating implies the production of stimuli with two distinct intentions: the most obvious one consists in informing the audience of something (*informative intention*), and the other—which comes first—is making the audience see that one has that informative intention (*communicative intention*) (Sperber & Wilson, 1995, p. 29).

Because one cannot exactly know what others know, one can only assume a set of facts that are manifest to them (their *cognitive environment*) as well as what might be mutually manifest both to the communicator and the audience (*mutual cognitive environment*) (Sperber & Wilson, 1995, pp. 39–42).

Of course, in communication, the given information can be old knowledge to the audience or, on the contrary, it can be completely new. Furthermore, this new information can either be somehow connected to already familiar elements within the receiver’s cognitive environment or can lack a connection to anything that the receiver knows. Each of these cases leads to different degrees of relevance. The better the connection between new and old items, the better they can serve as premises in an inference process to derive

further new information. In other words, contextualizing new information in old information is what gives rise to *contextual effects* such as reinforcing old assumptions, abandoning them or transforming them in some way (Sperber & Wilson, 1995, pp. 108–109): “When the processing of new information gives rise to such a multiplication effect, we call it *relevant*” (Sperber & Wilson, 1995, p. 48; emphasis in the original).

As “all human beings automatically aim at the most efficient information processing possible” (Sperber & Wilson, 1995, p. 49), the communicator and audience both tend in their own way to reach out to each other: the communicator adopts a “behaviour which makes manifest an intention to make something manifest”—a behaviour referred to as *ostension*—and the audience in turn trusts that “just as an assertion comes with a tacit guarantee of truth, so ostension comes with a *tacit guarantee of relevance*” (Sperber & Wilson, 1995, p. 49; my emphasis). Any assumption is therefore relevant in a given context not only to the extent that its contextual effects are pervasive, but also, at the same time, to the extent that the effort required to process it in that context is minimal (Sperber & Wilson, 1995, p. 122). By selecting the most relevant stimulus for a particular audience, the communicator attempts to diminish the processing effort demanded of the addressee: “Here, the interest of the communicator and the addressee coincide” (Sperber & Wilson, 1995, p. 157). In other words, one could say that successful communication is a bridge built on empathy.

Sperber and Wilson (1986/1995) phrase their basic *principle of relevance* as follows: “Every act of ostensive communication communicates a presumption of its own optimal relevance” (Sperber & Wilson, 1995, p. 158), *presumption of optimal relevance* being reformulated in their second edition as follows:

- (a) The ostensive stimulus is relevant enough for it to be worth the addressee’s effort to process it.
- (b) The ostensive stimulus is the most relevant one compatible with the communicator’s abilities and preferences (Sperber & Wilson, 1995, p. 270).

As we shall see later in this article, ostension and relevance are usually obtained through more or less salient forms of redundancy, which guide the audience—hence, the translator—towards the most plausible core meanings of a given document. One must also notice that in translation situations, point (b) above also concerns the translator, whose role is that of both an addressee and a communicator.

2.2 The Construction-Integration Model

According to the approach proposed by Kintsch (1988), an important factor in building discourse representations is *general knowledge*. This is because it “makes understanding processes smart: it keeps them on the right track and avoids exploring blind alleys” (Kintsch, 1988, p. 164). His objective is to posit a model of understanding that “combines a construction process in which a text base is constructed from the linguistic input as well as from the comprehender’s knowledge base, with an integration phase, in which this text base is integrated into a coherent whole” (Kintsch, 1988, p. 164).

Kintsch assumes that knowledge is conceptualized as an *associative net* with interconnected nodes that are either concepts or logical propositions (Kintsch, 1988, p. 165); but the structure of this net is not pre-established, rather it is “generated in the context of the task for which it is needed” (Kintsch, 1988, p. 164). Kintsch therefore advocates a “weak production system that generates a whole set of elements”, which can even be too large or irrelevant (the *construction* stage) but which are then processed in a context-sensitive manner (the *integration* stage) in order to select the most adequate elements from all the outputs previously generated (Kintsch, 1988, p. 164). According to this approach, meaning is always created in context, and is therefore defined as a

“momentary, subject- and situation-specific activated semantic and experiential context of a concept” (Kintsch, 1988, note p. 165). Kintsch elaborates further:

Concepts are not defined in a knowledge net, but their meaning can be constructed from their position in the net. The immediate associates and semantic neighbors of a node constitute its core meaning. ... It is not possible to deal with the whole huge knowledge net at once. Instead, at any moment only a tiny fraction of the net can be activated, and only those propositions of the net that are actually activated can affect the meaning of a given concept. Thus the meaning of a concept is always situation specific and context dependent. It is necessarily incomplete and unstable (Kintsch, 1988, p. 165).

The construction process of the Kintschian model involves (a) forming concepts and propositions which correspond to the input; (b) elaborating these elements by selecting some of their most closely associated neighbours in the general knowledge net; (c) inferring additional propositions; and (d) assigning connection strengths to the created pairs of elements (Kintsch, 1988, p. 166).

As a result of this process, one obtains a rich yet incoherent text base and so, as a consequence, it has to be further processed through the integration phase, where the actual discourse context will constrain the range of possible meanings. An important aspect of the whole comprehension process is that it occurs in cycles (which for Kintsch roughly correspond to short phrases). Each cycle gives rise to the elaboration of a new net, which includes the short-term buffer contents from the previous cycle. Then, once this net is constructed, the integration process can begin (Kintsch, 1988, p. 168).

The actual integration following each processing cycle consists of “freezing” the most highly activated nodes of the net. These highlighted nodes are what will constitute the new (temporary) discourse representation, and they can be of various natures (not only lexical or propositional, but also inferential) (Kintsch, 1988, p. 168):

The previously sketched model implies that word meanings have to be created anew in each context, that this is initially strictly a bottom-up process with context having its effects in the integration phase, and that this construction-plus-integration process takes time, with different factors influencing successive phases of the process (Kintsch, 1988, p. 169).

In the process of translating, the relative importance of the discourse context, as described by Kintsch, means that the translator will not only need to transfer the “interconnected nodes, associates and semantic neighbors” appearing in the original document, but also will have to anticipate the “semantic and experiential context” that most probably will be activated by the audience of their translation.

3. A three-stage reading process

One of the general aspects of deferred multimodal communication that is noteworthy is the intuitively obvious observation that any given document has been intentionally created in order for it to be seen, heard or read. Hence, such a document is perceived, both by the communicator and their audience, as being a *message*, and in this sense, the mere existence of that document is already proof of an underlying *communicative intention*, even in the apparent absence of the communicator herself in the reception situation. All the bottom-up features that the document reveals consequently appear to the viewers as inherent parts of an ostensive communication. Thus, as the document is an *ostensive communication artefact*, it carries—just as any act of ostension—not only “a request for attention” (Sperber & Wilson, 1995, p. 155), but also a *guarantee of relevance* in itself,

as was suggested in the principle of relevance mentioned above. As a result, any obvious salience within the set of “stimuli” (Sperber & Wilson, 1995) or “inputs” (Kintsch, 1988) of such a document will be interpreted by the audience as a clue worth processing.

The present approach to reception and comprehension broadly suggests that one should look at the process through its constitutive phases, which partly unfold in time. These are the previously mentioned construction and integration phases. In the context of visual communication, however, where image viewing relates differently to time than either text reading or listening, there might be a preliminary stage to be considered. Even before Kintsch’s first step in the construction process, there is a purely perceptual phase that involves the initial glimpse of the document.

3.1 The perception phase: setting up a background

The viewer’s first contact with a visual document amounts to a *global perception* of the given scene. In contrast to slowly unfolding linguistic communication, this first glimpse offers an almost instant background setting for any further reading of the message. This is typically a purely perceptual situation, where reception follows a bottom-up path, so essentially it is the given document that leads the audience’s attention.

When the viewers see a picture, a scene or a situation, they almost instantly gain a first overall impression of what is going on in it, using their prior knowledge. They grasp an intuitive “meaning” based on quick, bottom-up features that emerge from the scene. This is what visual recognition researchers, such as Lester Loschky’s team at Kansas State University, refer to as the “gist of the scene”, which is accurately recognisable by viewers in some 36 milliseconds (Visual Cognition Laboratory, Kansas State University, n.d.).

Almost simultaneously, the viewers may already have connected these first, essentially denotative features to very basic connotation paths, such as, in the case of a film beginning with [“dark street view” > “invisible areas” > potential menace” > “fear”]. Such connotations in turn can serve as a departure point for *genre instantiations*: “darkness” and “menace”, for instance, might lead the audience’s thoughts towards drama or film *noir*; “fear” towards thrillers or horror movies.

This perceptual step in visual document reading appears to involve some type of preliminary integration of a background (the cultural context, the mood, the genre, the general gist of the communication). The perception phase establishes a basis upon which a more detailed scrutiny might follow, by setting a semantic frame in which the more precise search for meaning will occur. In Kintschian terms, this initial phase might already activate basic associative nets that the audience will then try to confirm. The key elements in this confirmation process (or perhaps more rarely so, rejection process) will be the saliences carried by the viewed document and the relations that these salient elements maintain between each other. This is where a more active reading process begins.

3.2 The construction phase: linking saliences

Communicators basically create manifestness in a set of assumptions in two prototypical ways: by pointing it out explicitly (a case of “strong communication”) or by adopting a more subtle framing strategy (a case of “weak communication”). As Sperber and Wilson (1995) emphasize, “with weaker forms of communication, the communicator can merely expect to steer the thoughts of the audience in a certain direction” (p. 60). Nevertheless, creating manifestness always occurs by means of salience, which is why, during the construction phase, the viewer will strive to establish links between any saliences that can be detected among the co-occurring elements of the multimodal document. The viewer will suppose that each of those salient constituents ought to be treated as a partial clue to

a more central meaning, which probably was already roughly sensed during the perception stage.

Moreover, the receiver is fully or intuitively aware that one of the most basic methods that the communicator can use to emphasise an important idea—although perhaps not the most subtle one—is to repeat it. Hence, among the given saliences, those that reiterate a content in any manner will simply gain in relevance. As Forceville (2014) aptly points out, picture-based communication is predominately mass communication, which is “more complex than face-to-face communication in the sense that there is no opportunity for adapting the message in the case of misunderstanding or incomprehension. Experienced mass communicators of course anticipate this” (p. 62). Indeed, one of the most accessible tools for this anticipation is the use of redundancy. As Kintsch observed in his description of the construction process (see Kintsch, 1988, p. 166, point d, mentioned in Section 2.2), the audience is assigning connections here between paired elements and some of them acquire more strength than others because of some feature they have in common.

One important point needs to be emphasized, however: during “strong communication”, the communicator offers multiple and more or less synonymous triggers that lead to an explicit, *endophoric* signified, such as the use of a name, a nickname, a substitutive lexeme or a pronoun to refer to a particular character in a novel. On the contrary, during “weak communication”, the relation to the signified is *exophoric*. This means that the given multiple and explicit triggers lead the audience’s thoughts to an implicit common ground of significance which, because of its shared nature, appears nevertheless to be relevant. In this latter case, all the apparently various stimuli finally “say the same thing”, even if the “thing” in question is not explicitly mentioned in the viewed document.

I believe this was established by Barthes in his semiotic analysis of the *Panzani* advertisements (Barthes, 1964, pp. 40–51). For Barthes, any advertising poster, owing to the very genre it represents, always carries the same fundamental connotation, namely, that “the described product is the best”. This background is something the viewer acknowledges in the first stage of their viewing (the perception phase sketched in Section 3.1). It is only then that the viewer can begin actually to *read* the denotative elements of the poster, such as the vegetables, the net, the colours and the “Panzani” label, which will direct their understanding towards the concept of *italianity*.

This image immediately delivers a series of discontinuous signs. ... A second sign is nearly as obvious; its signifier is *the union of* the tomato, the pepper and the tricolor tones (yellow, green, red) of the poster; its signified is Italy, or rather italianity; this sign is *in a relation of redundancy* with the connoted sign of the linguistic message (the Italian assonance of the name Panzani) (Barthes, 1964, p. 41; my translation and emphasis).

The first important aspect to note is the compositional nature of these “signs”: the abovementioned signifier, for instance, is a *combination* of visual elements (tomato, red pepper and mushroom), on the one hand, and the described colours, on the other. Furthermore, the entire sign is in a redundancy relation with the Italian assonance of the brand. Taken as a whole, these elements lead the reader towards the signified, that is, the idea of “italianity”.

All of these stimuli, made manifest through relative salience, become meaningful to the reader as a result of the steering effect caused by their co-occurrence: “Human cognition tends to be geared to the maximisation of the *cumulative relevance of the inputs* it processes” (Sperber & Wilson, 1995, p. 261; my emphasis). Every image is polysemous and implies, underlying their signifiers, a “floating chain” of signifieds, and the reader can select some of these and ignore others (Barthes, 1964, p. 44). But the relative strength of a given reading (be it the “obviousness of the sign” and the “choice of signifieds” that

Barthes mentions, or the “fraction of the net being activated” in Kintsch’s terms) results in fact from the *physical proximity of the salient signifiers* and from the *mental proximity of their signifieds* (that is, their shared features or redundant semantic aspects). It is interesting that this makes the semantic understanding process almost comparable to the basic perceptual grouping principles of Gestaltian *Prägnanz*, such as the laws of proximity, similarity, closure or continuity described by Todorović (2008).

3.3 The integration phase: grounding the inferred assumptions

Eventually, an *integration phase* is needed so that the audience can stabilise their inferences from the confrontation of co-occurring saliences (described in Section 3.2). The audience does this by selecting those of the suggested assumptions which are the most manifest and need to be treated as being the most relevant. In reception forms that build more obviously over time, such as film-viewing, the integration phase can be conceived of as a means of *grounding* the most relevant inferred assumptions made thus far, so that in turn they become part and parcel of the background knowledge for a new, subsequent perception and/or construction phase. This is presented schematically in Figure 1 below.

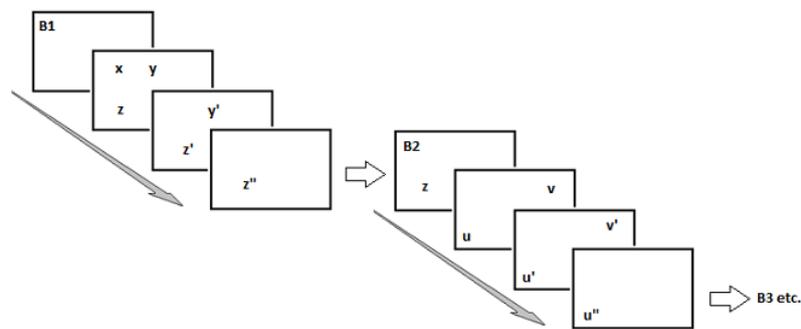


Figure 1: The construction–integration process. *B1* is the initial background defined during the perception stage. The possible readings are *x*, *y* and *z*, based on the co-occurring salient elements present in the document; “*z*” is the most relevant reading, that is, the one the audience will keep in mind—or *integrate* (as shown by the thicker *arrow*)—as a new starting point (*B2*) for further construction.

This is at the heart of Kintsch’s notion of “cycle”. The reception process is a continuous, recursive movement which constantly builds on prior stages of knowledge and where *integration* acts as a springboard for subsequent *construction* phases. In fact, as comprehension and learning are strongly intertwined, the whole process closely relates to the concept of *spiral learning* in the adjacent field of education:

- (c) A spiral curriculum is one in which there is an *iterative revisiting of topics*, subjects or themes throughout the course. A spiral curriculum is not simply the repetition of a topic taught. It requires also the deepening of it, with *each successive encounter building on the previous one* (Harden & Stamper, 1999, p. 141; my emphasis).

4. Redundancy reading in practice

Let us now turn briefly to reviewing some concrete examples of this redundancy system and analysing the reception process in deferred multimodal communication situations that might involve different translation scenarios. Because of space limitations, I will concentrate, on the one hand, on logos and advertisements, which are not necessarily adapted in their entirety for new audiences and, on the other, on feature films, which are usually fully translated. Furthermore, I analyse only parts of the reading process in each of these cases, hoping nevertheless that these examples will shed light on the entire redundancy system as well as on the translation process, which will be addressed in the discussion Section 6.

4.1 Company logos and advertisements

As Barthes (1964) noted, the role of images is most easily studied in advertising due to the obvious intentionality of that particular form of communication (Barthes, 1964, p. 40). With company logos, the interplay between image and text is often strongly built on redundancy. The bookshop logo presented on the left in Figure 2 contains the stylized picture that visually represents both the letter “S” and two books on top of each other. This creates an endophoric double redundancy with the Finnish name of the company: the initial letter of its name, “S(uomalainen)”, and the “book” image, also part of the compound word “kirja(kauppa)”. Once these relations are known to the audience, the visual logo could actually stand alone as a symbolic representation of the company. However, one has to notice that the redundancy in this first case is *endophoric* and, to be precise, only *partial*.



Figure 2: Two company logos – *Suomalainen kirjakauppa* (“The Finnish Bookshop”) and *Target Advertising*.

The same type of observation may be made regarding the company name “Target”, because the *endophoric* link between the image of the sight and the word “target” is that the latter is actually *what is viewed through* the sight. Furthermore, by the addition of the word “advertising” to the reading, it becomes clear that it is in the *exophoric* general semantic context of “how one uses sights”—as in the activated part of the associative net where targets, sights and advertising meet—that the sight eventually becomes what allows the viewers to picture themselves as the agent, deciding on the actual impact point of their advertising action. This example reveals how the combination of text and image can create meanings that go far beyond the mere addition of the stimuli’s separate senses.

In Figure 3, a mere glimpse at the document—the perception phase—will convey to the viewer that (a) this probably is an advertisement (for several reasons, such as the size of the text elements—which are brands—and the clear salience of the object); (b) this advertisement probably concerns a luxury product for women (because of the general aesthetics due to the colours, the dimmed lighting, the woman’s face, the glowing

necklace and the pearl-shaped flask); and, because of this communication genre, (c) that beyond the basic commercial message, the document probably contains another message that somehow relates to the product.

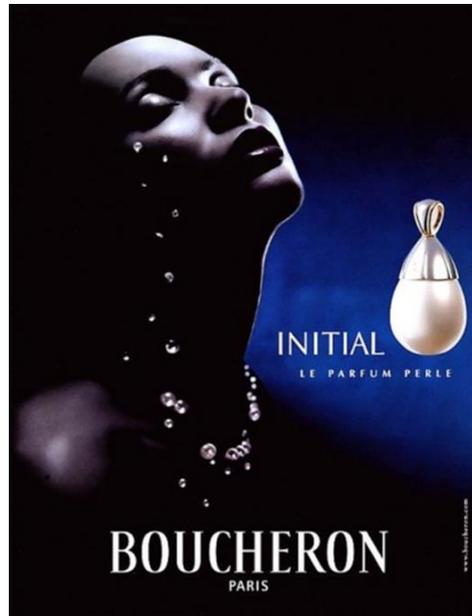


Figure 3: A French advertisement – *Boucheron (Paris)*, “Initial, *Le parfum perle*”.

The world depicted in the *Boucheron* advertisement—or the story it conveys—is presented to the viewer to be discovered through the combination of salient clues. The audience re-creates this second message—in a construction phase—by linking all the relevant textual and visual elements, thus activating an associative net where all of these features become cohesive:

- the product name “Initial”;
- its description as “*Le parfum perle*”;
- the flask shape;
- the pearls of a necklace in suspension;
- the dark black and blue background colours;
- the light emanating from above;
- the feminine face;
- the nudity of the shoulder.

In this case, the viewers might feel as though they are witnessing an underwater scene, deep under the surface, somewhere in the original purity of a mystical primeval sea. Perhaps this scene is populated by mermaids or, even better, the pale and perfect womanly figure herself is a pearl. The strength of the “sea” connotation is produced by the redundancy of the shared ground of significance of all the above-mentioned stimuli combined, as a whole, that is, as a *cumulative* signifier. None of the triggers alone can achieve this, and none does: the idea of the “sea” is completely exophoric and constructed in the viewers’ mind. One could formulate this particular process otherwise:

the exophoric redundant element is a variable (x) whose value (x = “sea”) is calculated through the shared features of the specified trigger stimuli (the listed elements above in their interrelation).

There are many types of actual semantic relation between these stimuli. For instance, the flask *has the shape of* a pearl (formal resemblance), the necklace *is made of* pearls (constitution), pearls *originate from* the sea (agency), mermaids *are known to be* feminine creatures living underwater (contextual association), etc. But what these activated nodes actually do is that they all point to the same semantic context. As Kintsch states, “[t]he immediate associates and semantic neighbors of a node constitute its core meaning”, which, as one recalls, is defined as the “momentary, ... activated semantic and experiential context of a concept” (Kintsch, 1988, p. 165).

One could argue that the pearls in Figure 3 seem to be falling like tears from the women’s right eye. But if one takes the “activated semantic context” where this pearly sub-event appears, a far more likely explanation of their state of suspension would be that it is the necklace that is breaking up into bubble-like orbs, which float to the surface. This type of detail illustrates how a single stimulus can be truly coerced by its co-occurring stimuli to enter an associative net or a core meaning in the most relevant manner.

4.2 Feature films

For feature films, one needs to consider the inherent sequentiality of the document as well as the temporality of reception that goes with it, which makes the reading process somewhat more complex. Furthermore, the sequenced images of feature films are not still pictures: instead, they are predominately moving scenes (although presented in distinct shots). These scenes cannot be visualised side by side (as they would in graphic novels, for instance). Rather, they pass before the viewer at a pace defined by the communicator(s).¹ Finally, it is important to note that the almost limitless world of sounds is an inherent part of the filmic document. One could also add that the subtitles in themselves contribute a whole new layer to the viewing of a film.

Figure 4 presents an apparently trivial cut-edit. In the first shot, a husband tells his wife about his situation in the army during World War I. In the English subtitles of the DVD, he states: “If I desert, the gendarmes will come and get me.” Immediately following his utterance, a brief shot from a black-and-white film that resembles a documentary is inserted into the dialogue, showing a hooded man being executed and letting the spectator hear guns being fired.



Figure 4: A feature film excerpt—Jean-Pierre Jeunet, 2004. *Un long dimanche de fiançailles/A very long engagement*, Warner Bros.

This film excerpt resembles the advertisement analysed earlier in that it exemplifies how salient and co-present stimuli are used to construct meaning. Of course, the “co-presence” of semantic triggers (or their “physical proximity”) is, in this cinematic case, also induced

by their temporal proximity. In other words, the list of relevant stimuli that enables the viewer to infer the idea that the husband *fears being executed* rather than merely *being apprehended* by the gendarmes comes from both shots. It is only through the montage of these that the viewer can retrieve the assumption that the husband foresees himself being executed, which is what prevents him from sleeping. The activated associative net of cohesive features builds up over time:

- night time, silence (with background music);
- characters in bed, wearing pyjamas, but still awake;
- man speaks to his wife;
- his utterance: “*Tu comprends, si je déserte, je serai repris par les gendarmes*”;
- subtitles: If I desert, the gendarmes will come and get me;
- --- Cut-edit ---
- day time, black-and-white image (but music continues);
- hooded man tied to an execution pole;
- sound of synchronised shooting;
- smoke;
- man falls down (with sound of his falling).

The viewer’s integration of the completely implicit and exophoric overall idea of the husband’s facing execution and his fear that it might happen is achieved through the multiple linking of all these salient features: the “I” of the utterance (or its subtitled correspondent) is linked both to the husband speaking and to the hooded character being executed; the mentioned “gendarmes” are instantiated by the unseen firing squad, the latter being suggested only by the sound of the shooting and the image of its impact on the tied-up man, etc. The unity of the scene, which allows the viewers to create these associations (or, in fact, which coerces them to do so, as the guaranteed relevance of the scene is assumed), is produced essentially by the temporal proximity of the two distinct but collated shots. These are combined with the use of a single, continuous musical background. Thus, any alteration of this net of cohesive redundancy could have an impact on the overall perception one has of the scene.

5. Two types of redundancy

The documents briefly analysed above display how salient triggers or, more precisely, *trigger compounds* (groups of stimuli, inputs or signifiers) steer the reader towards some core meaning (most relevant assumptions, associative nets or signifieds). These meanings can be either endophoric (explicit within the text) or exophoric (implicit). In other words, the difference lies in the location of the redundant reference.

First, a set of triggers might essentially serve to refine *one single endophoric reference*. During monomodal communication, visual, auditory or verbal triggers can be repeated for narrative, stylistic or memorizing reasons, as in speeches by orators. Take, for instance, Martin Luther King’s speech in Washington in 1963:

I have a dream that one day this nation will rise up, live out the true meaning of its creed ...

I have a dream that one day on the red hills of Georgia [the] sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood. I have a dream that one day even the state of Mississippi ... will be transformed into an oasis of freedom and justice.

I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream ...

(Martin Luther King, Jr, 1963, pp. 4–5; my emphasis)

Triggers may also be reiterated in slightly different forms (cf. distinctive viewing angles of one single character in comic-strip panels); or, in text linguistics, they can be syntactically substituted in different ways (cf. the referential chain relations described by Halliday & Hasan, 1976). Nevertheless, the reference is the same and it is explicit. In multimodal situations, the only difference is that these triggers appear in various modes, such as the case of the symbolic logo combined with the bookshop name in Section 4.1. Moreover, the feature film mentioned in Section 4.2 also contained several monologue utterances describing an event that were immediately followed by a distinct and purely audiovisual scene that reiterated the same event in other modes.

Even though this first type of redundancy essentially serves as a cohesion-building tool, it is important to consider that each new trigger may alter the endophoric reference. This alteration depends on the predication the trigger belongs to. The referential chains, which are predominantly anaphoric, constantly refer to the last mention of the referent, which evolves in time within the narrative. For instance, at each mention of the word “dream” in Martin Luther King’s speech, its signified becomes clearer to the audience.

Secondly, and perhaps more crucially for creating meaning, we find cases of *partial redundancy of exophoric reference*. This reference is also created through multiple triggers, perceived here as clues by the audience because of their relative salience and spatiotemporal proximity. Nonetheless, in this instance, the entity (idea, location, object, character or event) prompted by the triggers is not explicitly stated in the document. Instead, a shared semantic feature of the trigger compound defines the value of the exophoric variable.

In brief, when salient elements act as multiple signifiers, they converge towards a common exophoric signified. They activate an associative net, which becomes the core meaning of the message (or at least one of them). In multimodal situations, these triggers are simply in different modes, as we witnessed in the *Target advertising* logo (Fig. 2 in Section 4.1), the *Boucheron* advertisement (Fig. 3 in Section 4.1) and the excerpt from *A very long engagement* (Fig. 4 in Section 4.2), which are all occurrences of this second type of redundancy.

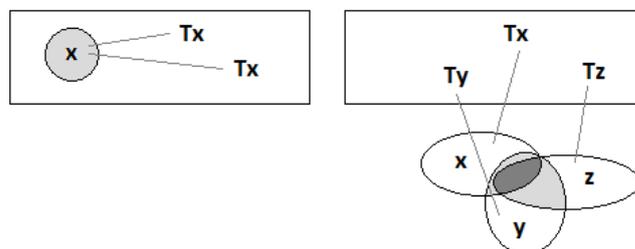


Figure 5: “Tx” stands for “Trigger for referent x”. The darkened regions represent endophoric redundancy (on the left) and exophoric redundancy (on the right).

Needless to say, these two types of redundancy can be combined in communication: the word “*perle*” in the advertisement analysed in Section 4.1 had intratextual counterparts in the shape of the perfume flask and in the necklace. In the light of Figure 5, one could

say that those particular T_x triggers therefore have an endophoric reference x , but one which can then be linked to other T_y - or T_z -triggered exophoric references such as the mermaid-woman, the connotations of “Initial”, etc.

6. Discussion

From the viewpoint of translation, matters become slightly more complicated. As a mediator in communication, translators are positioned between the original communicator and the targeted audience of the translation. In cases of deferred communication, more precisely, they need to assume the role of the intended receiver of the communicator’s “ostensive communication artefact” and then adopt the role of a second communicator who addresses the target audience. In both aspects of their work, translators deal with the presumption of optimal relevance.

Knowing that the communicator has selected the most relevant stimuli for the original audience and made them salient in order to diminish the processing effort of the addressee, professional translators will always strive to determine with care those particular salient stimuli and invest their best efforts in building their interpretations primarily on them. The idea that “each act of understanding is unique and so are, consequently, translations and interpretations” (Muñoz Martín, 2010, pp. 175–176) is, of course, difficult to contradict; but I would argue that the redundancy system presented in this article can precisely be considered as an efficient means of restricting the possibilities of random individual readings. The translator is well aware that the original intended interpretation will be strengthened by finding in-document arguments (within the trigger compounds suggested by the communicator). The integrated “mental representation” is in this sense truly “prompted by the source text segment” (Ketola, 2016, p. 71; my emphasis).

At the output end of their work, nevertheless, translators are confronted by all types of constraint (linguistic, formal, spatial, temporal, etc.), which might lead to different strategic choices. One very concrete example of linguistic constraint, for example, would be the absence, in the Finnish language, of gender-specific personal pronouns in the third-person singular. Because of such “underspecification”, the translator has to find ad hoc ways to disambiguate an anaphoric pronoun referring to a preceding sentence of an English source text: for example, where both “he” and “she” would find an antecedent.

In audiovisual translation (especially in subtitling and audio-description), a generally accepted policy is to reduce as much redundancy as possible in order to gain space (or time) to express the most relevant information. In short, all that is said need not be translated, unless it is necessary to the plot (Vertanen, 2012, p. 153). One might argue that the “plot” is not a sufficient criterion, and that the necessity of finding a way to relay the joy of discovering what is implicitly signified by the document is of equal importance.

Typically, subtitles correspond to textual elements or utterances of their source document and in this sense they share the same role as the original linguistic triggers in a communicative situation. Problems usually arise when these linguistic triggers cannot be translated. When faced with meaning-free trademarks in their work, translators are discharged of their duties, whereas for the examples analysed in section 4.1, where brands and/or logos have to be left untranslated even though they have intrinsic meanings, the situation might appear more frustrating. For example, in Figure 2, the “S”-shape of the book logo is perceived as meaningful only to those who understand “Suomalainen kirjakauppa” and, similarly, the relation between the series of letters “T-a-r-g-e-t” and “A-d-v-e-r-t-i-s-i-n-g” and the visual representation of the sight will remain nebulous to anyone who does not comprehend written English. This would still be the case even if the word “Advertising” were actually translated.

The linguistic elements in the advertisement of Figure 3 that probably would not be translated would be the proper nouns, such as the trademark “Boucheron” and the name

of the perfume itself, “Initial”. Neither of these words would necessarily be familiar to, say, a Finn. Still, the difference between the two is that the first name has no specific meaning, whereas the latter does.² Furthermore, the Finnish translator would probably need to translate the slogan “Initial—The pearl perfume” so that it explains the relation to the flask shape in the advertisement’s visuals (since “*perle*” is to be translated), but they would not be able to create the link to the primeval sea (since “Initial” would have to remain in its original form, which is not transparent to the readers of the translation). This would weaken the overall cohesion of the translated poster, owing to the inevitable impoverishment of the original trigger compound. As a result, the idea of “the sea” would remain, but the relative strength of its presence would be lost in translation. This occurs because while the original endophoric redundancy remains unaltered, the exophoric one is diminished. The degree of exophoricity is therefore different between the original and the translation, and consequently a change occurs in the nature of the reception.

Another type of impact on reception could be observed if the translator were to spell out a meaning that should have been constructed by the viewers themselves, as in the original communication. In the feature film sequence we discussed in Section 4.2 (Figure 4), this would have been the case if the subtitler had used the verbal phrase “execute me” instead of “come and get me” when talking about the gendarmes’ action. In a way, from the viewers’ perspective, the challenge of constructing meaning through the incentives suggested by the communicator would not only have been flawed, but the very nature of the redundancy would have changed. This is because the underlined explicitness of the endophoric redundancy would again have relegated the exophoric redundancy to the periphery of the scene’s meaning.

7. Conclusions

In this article, we observed that a message consists of different types of salient trigger and that the communicator presents these to their audience as clues. Because the audience trusts that the communication is relevant and knows that repetition is the easiest way for the communicator to emphasize an assumption, the abovementioned saliences are perceived as cumulative signifiers or trigger compounds. These cumulative cues lead to a shared semantic ground, which is to be found either in endophora or in exophora. In exophora, the semantically shared subparts of the multiply-triggered exophoric references constitute core meanings necessary to the interpretation of the document. Discovering those core meanings and integrating them into one’s prior knowledge is precisely what reading is all about. One links the triggered references in a construction phase, then recursively verifies that the created associative nets are in line with previously integrated knowledge; this is why reading is considered to be a cyclical process.

We assume that translators working with multimodal texts need to master this complex redundancy system, since they share the roles of both the original text addressees and the communicators who address their own audiences. Such special expertise means being able to manipulate meanings that are triggered by any mode. As we have seen, the relative importance of each mode can vary according to the different trigger combinations. From a methodological point of view, this means that the study of multimodal communication and its translation could benefit from a unified approach of multimodal meaning-making, avoiding any *a priori* that would prioritize visual, aural, spoken or textual triggers in any manner.

Furthermore, I believe that analysing multimodal translation within the framework of Relevance Theory and the Construction-Integration Model, and especially through the prism of exophoric redundancy, might offer new insights into some classic issues in Translation Studies, such as *translation strategies* (cohesion change; use of synonyms, hyponyms or superordinate terms; omission...), or into more controversial issues, such as the alleged tendency of *explicitation* in translated texts. This is so because studying

translation-inherent explicitation through multimodal corpora might at least permit the researcher to eliminate some of the language-specific biases or interferences implicated in the translation process (see Becher, 2010; Klaudy, 2008).

Ultimately, the use of exophoric semantic redundancy is a primary creative asset for producing implicit—but nonetheless central—meanings in multimodal communication. Professional translators know that it is important not to overlook the impact of seemingly trivial changes in individual triggers, associative nets or degrees of exophoricity through inauspicious explicitations. Doing so is important in order to transmit not only the facts or the plot of a document to their audience, but also the whole positive experience of meaning construction.

References

- Barthes, R. (1964). Rhétorique de l'image, *Communications*, 4(1), 40–51. doi:10.3406/comm.1964.1027
- Becher, V. (2010). Towards a more rigorous treatment of the explicitation hypothesis in translation studies. *trans-kom*, 3, 1–25. Retrieved from http://www.trans-kom.eu/bd03nr01/trans-kom_03_01_01_Becher_Explicitation.20100531.pdf
- Forceville, C. (2014). Relevance theory as a model for analysing visual and multimodal communication. In D. Machin (Ed.), *Visual communication* (pp. 51–70). Berlin: Mouton de Gruyter.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Harden, R. M., & Stamper, N. (1999). What is a spiral curriculum?, *Medical Teacher*, 21(2), 141–143. doi:10.1080/01421599979752
- Ketola, A. (2016). Towards a multimodally oriented theory of translation: A cognitive framework for the translation of illustrated technical texts, *Translation Studies*, 9(1), 67–81. doi:10.1080/14781700.2015.1086670
- King, M. L. Jr. (1963). “I have a dream”, Speech by the Rev. Martin Luther king at the “March on Washington”. Retrieved from <https://www.archives.gov/files/press/exhibits/dream-speech.pdf>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model, *Psychological Review*, 95(2), 163–182. doi:10.1037/0033-295X.95.2.163
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Klaudy, K. (2008). Explicitation. In M. Baker & G. Saldanha (Eds.), *Encyclopedia of translation studies* (pp. 80–85). London: Routledge
- Kleiber, G. (2001). *L'anaphore associative*. Paris: Presses universitaires de France.
- Lautenbacher, O. P. (2014). La redondance, principe moteur de la cohésion du film sous-titré: Étude de cas, *Parallèles*, 26, 53–68.
- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction, *Journal of Sociolinguistics*, 20(3), 336–366. doi:10.1111/josl.1_12177
- Muñoz Martín, R. (2010). On paradigms and cognitive translatology. In G. M. Shreve & E. Angelone (Eds.), *Translation and cognition* (pp. 169–187). Amsterdam: John Benjamins.
- Seliger, M. (2008). *Katujen galleriat: Ulkomainnonnan visuaalista retoriikkaa Helsingissä vuosina 2004–2005*. Jyväskylä: Gummerus.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
- Sperber, D., & Wilson, D. (1986/1995). *Relevance theory: Communication and Cognition*. Oxford: Blackwell.
- Todorović, D. (2008) Gestalt principles. *Scholarpedia*, 3(12), 5345. Retrieved from http://www.scholarpedia.org/article/Gestalt_principles
- Vertanen, E. (2012). Ruututeksti tiedon ja tunteiden tulkkina. In R. Oittinen & P. Mäkinen (Eds.), *Alussa oli käännös* (pp. 131–153). Tampere: Tampere University Press.
- Visual Cognition Laboratory, Kansas State University. (n.d.). Recognizing the gist of a scene. Retrieved from <https://www.k-state.edu/psych/vcl/basic-research/scene-gist.html>

-
- ¹ Of course, there are means of pre-visualizing motion shots and their sequentiality in *storyboards*, for instance.
 - ² Incidentally, even “Boucheron” might convey some connotations for, say, speakers of French, but in this particular communication act these are irrelevant and hence unwanted. Any connotation the communicator would *want* to share would have more than one trigger stimulus.