

Thinking Outside the “Methods Box”: New Avenues for Research in Multimodal Translation

Sari Kokkola & Anne Ketola

School of Language, Translation and Literary Studies

University of Tampere

Useat multimodaalisten tekstien kääntämistä käsittelevät tutkimukset (e.g. van Meerbergen 2009; Taylor 2003) perustuvat systeemis-funktionaaliseen kieliteoriaan. Tästä syystä ne jäsentävät kaikkia multimodaalisissa teksteissä esiintyviä moodeja peilaamalla niitä verbaalisen kielen rakenteisiin. Kyseinen ajattelutapa asettaa kielen muita moodeja tärkeämmäksi. Oma tutkimuksemme äänen roolista elokuvakääntämisessä (Kokkola) ja kuvien roolista teknisessä kääntämisessä (Ketola) on osoittanut, että ääni ja kuva ovat merkittävässä osassa käännettävän teoksen merkitysten rakentumisessa. Tästä syystä painotamme, että kuva ja ääntä tulisi tarkastella itsessään merkityksellisinä tutkimuskohteina sen sijaan, että niitä kohdeltaisiin käännettävien tekstien ominaisuuksina. Tässä artikkelissa ehdotamme, että tähän asti hyödynnettyjä metodeja voisi täydentää menetelmillä, jotka huomioivat kuvan ja äänen erityispiirteet multimodaalisessa kääntämisessä. Uusien metodien käyttöönotto vaatii tieteenalaa omaksumaansa ennakkoluulottoman näkökulman ja poistumaan “metodologiselta mukavuusalueeltaan” (Hesse-Biber & Leavy 2008: 10). Uudet menetelmät voivat kuitenkin myös rohkaista kysymään uusia kysymyksiä ja löytämään tutkimusaiheesta uusia ulottuvuuksia. Pohdimme, voisivatko tällaiset menetelmät olla askel kohti laajennettua kääntämisen käsitystä (Tymoczko 2007: 57–58), jonka omaksuminen voisi puolestaan johtaa käännöstielen rajojen uudelleenmäärittelyyn.

Keywords: epistemology, knowledge construction, multimodality, research methods, translation

1 Introduction

Multimodality is one of the central issues in the study of translation today; after all, a large number of the texts being translated are multimodal (Hirvonen & Tiittula 2010: 1). Multimodality, in essence, refers to the coexistence of more than one mode – verbal language, images, sound, etc. – within a given context (Gibbons 2012: 8). A multimodal text, such as an illustrated text or a film, hence consists of interrelated, equally important modes which all contribute to meaning making. Multimodal issues are increasingly being addressed in Translation Studies. Yet, many approaches used in research into the translation of multimodal texts (e.g. van Meerbergen 2009; Taylor 2003, 2009), draw from Systemic Functional Linguistics and treat images and sound as structurally equivalent to language, employing concepts developed in linguistics to the analysis of all modes in a multimodal text.

Translation seems to be commonly understood as a predominantly verbal activity whereby the other accompanying multimodal elements mainly form a context for the verbal and inform the translator’s decisions. We argue, however, that this is only seemingly so. Perhaps the widely held conception of the verbal nature of translation comes

from the fact that the translator rarely has the opportunity to manipulate the image or sound of a multimodal text. Thus, the idea of translation being mainly verbal seems obvious and natural – a given – and as such easily remains unproblematized in a research setting.

Indeed, image and sound do not physically change if several different translations are co-presented with the same multimodal text. And yet, most translation professionals with experience of translating multimodal materials would readily agree that the multimodal whole appears different in the presence of different translations. We can, for example, imagine a painting and, say, ten different translations for its name. Each translation would make the artwork appear somewhat, or even radically, different. In other words, translation affects the entire experience of the artwork. What is conveyed is not the verbal message, but the entire multimodal experience. It is thus rather the relationship between a) the visual, the aural and the verbal, and b) the relationship between the reader/viewer/experiencer and the multimodal whole that is modified with the presence of different translations.

In this article, in the light of examples from our research – analysing the role of sound in film translation (Kokkola) and the role of images in technical translation (Ketola) – we suggest that the idea of translation being a mostly verbal activity is not natural but rather naturalized, or constituted, and as such can be questioned, unsettled, or subverted. In order for research to be genuinely multimodally oriented, it cannot employ methods that confine image and sound to a role subservient to the verbal. We therefore propose that research methods not dependent on the linguistic-based description of modes are needed to complement the existing methods. We suggest that doing so would require the discipline to step outside of its “methods comfort zone” and think outside its “methods box”, as expressed by Hesse-Biber and Leavy (2008: 10).

Further, we ask if adopting new methods addressing the specificities of image and sound in multimodal translation requires us to rethink the traditional definition of translation and, eventually, reassess the borders of Translation Studies. As Tymoczko (2007: 54) remarks, the ordinary definition of translation as “transferring a text from one language into a text in another language” becomes problematic when we acknowledge that even the seemingly simple concept of *text* is an open one. This notion is especially important for the study of multimodal texts: What exactly do we transfer when the meaning of the source text is created in the interplay of various modes? Tymoczko calls for an enlarged definition of translation which is able to accommodate the diversity of ideas about translation:

Unlike many earlier scholars who have attempted to define translation and who have sought closure on the question, I am interested in exploring the openness of the definition and the implications of that openness for the emerging international discipline. (ibid. 57–58)

In this article, we too set out to expand the definition of translation by acknowledging sound and image as essential elements of multimodal source texts, as well as to tentatively reflect on the implications of this for the discipline.

2 A Reflexive Perspective on Methods of Multimodal Translation

We call for a reflexive approach to methods employed in research on multimodal translation, particularly regarding the analysis of image and sound. A reflexive approach goes far beyond choosing a research method suitable for a particular research setting. Reflexivity refers to the capacity of the researcher to see his/her work as part of the “big picture” of knowledge construction within the discipline. Reflexivity is, thus, not interpretation of data, but rather, as Alvesson and Sköldbberg (2000: 6) describe it, “interpretation of interpretation”. This entails looking at one’s own research, or research conducted within a discipline, from alternative perspectives and understanding the fact that researcher involvement always has an effect on how the phenomenon under study is perceived. Each study on translation reflects the researcher’s stance towards the phenomenon and, taken together with other research, contributes to our understanding of how translation is conceptualized, what constitutes knowledge within Translation Studies, and how the borders of the discipline are defined. Reflexive thought in research acknowledges the fact that these issues are under constant negotiation and encompass broad considerations such as analysis of mechanisms of social power and control that shape Translation Studies as a global discipline (see e.g. Tymoczko 2007; Susam-Sarajeva 2000).

A reflexive mindset essentially includes critical examination and rethinking of taken-for-granted ideas within a discipline and a general openness to the idea that the dominant way of thinking within a discipline is not a stable state of affairs, but rather something constituted over time, and as such something that can potentially change. Explicit reflection on research frameworks increases the epistemological self-awareness of Translation Studies through enhanced understanding of how knowledge is produced within the discipline. Reflexivity allows to rethink, renew, and improve existing approaches and to introduce new ones in order to keep the conceptual and methodological basis of the discipline evolving.

3 Existing Approaches to Visual and Audio Phenomena in Translation Studies

The most influential approach to multimodality within Translation Studies seems to be the so-called *social semiotic multimodal analysis* (Jewitt 2009: 29) based on Halliday’s (e.g. 1978) theories of Social Semiotics and Systemic Functional Linguistics. This approach offers a view of multimodality in which “common semiotic principles operate in

and across different modes” (Kress & van Leeuwen 2001: 2) and therefore treats all modes as identical in their meaning-making strategies.

The social semiotic approach to multimodality posits that all modes of representation, including images, must fulfil the three metafunctions Halliday (1978) proposed for language: the ideational, the interpersonal and the textual (Kress & Van Leeuwen [1996] 2006). An example of applying the metafunction approach in Translation Studies is Van Meerbergen’s (2009) research into the Swedish translations of the Dutch *Nijntje* picture-books. For instance, Van Meerbergen analyses interpersonal interaction between the character portrayed in the image and the viewer of the image by examining the gaze vectors (direct lines formed between the participants’ eyes) involved, as proposed by Kress and van Leeuwen ([1996] 2006: 116–129). The analysis reveals that direct contact between the protagonist of the book and the viewer is frequently created through gaze vectors, and that at times the Swedish translator has reinforced the interpersonal interaction by adding the Swedish personal noun “du” (you) used to address the viewer, hence establishing a more direct personal contact (Van Meerbergen 2009: 11).

The metafunction analysis seems excellently suited for the particular type of data. Yet, in order to complement the validity of research findings arrived at using a particular method, the method has to allow for replicability and comparability of analyses using different data, in this case, different types of illustrated texts. It is reasonable to ask how one could go about analysing interaction within the interpersonal metafunction in different types of illustrated texts, for instance, where the images do not depict human (or human-like) characters and where, consequently, there are no gaze vectors involved. If the issue of interpersonal interaction cannot be accessed, is it even a relevant question to consider when working with particular type of data? This is one of the examples that has lead us to doubt if methods developed for linguistics may always be applied to the analysis of other modes in a consistent, coherent and productive manner.

Another example of how the social semiotic approach has been used within the discipline is *the multimodal transcription and text analysis*, originally developed by Baldry and Thibault (2005) and applied to the context of Translation Studies by Taylor (2003), which is the most extensive attempt so far in Translation Studies to explain the role of the different modes of a multimodal text. The analysis model seeks to identify the components of multimodal messages by presenting a breakdown of the visual, verbal and audio aspects of a multimodal text in the form of a block of columns (for illustration, see Taylor 2003: 196). The model consists of six columns representing the following elements: 1) time, 2) visual frame as a still image, 3) a detailed description of the contents of the visual frame, 4) description of kinesic action (movements, gestures etc.), 5) soundtrack (mention of the presence of dialogue, music, sounds), and 6) subtitles.

A closer inspection of the model reveals that, despite its high level of detail, the model seems to privilege certain modes over others. The analysis focuses on the relationship between the visual and the verbal (or the image and the subtitle), and the analysis of sound phenomena other than speech seems to stop at itemization, i.e. the presence of sound phenomena such as “silence”, “tempo”, “volume” or “pause” is mentioned in the transcription itself, but it is not made explicit how the translator actually analyzes them. Might it be that models such as multimodal transcription are better suited for studying certain modes while downplaying others? Might the limited attention paid to sound in Translation Studies be due to the fact that Translation Studies lacks methods for analyzing sound phenomena in their own right?

Both of these approaches, in our interpretation, privilege the verbal mode over others. The metafunction analysis of images views visual representation as structurally equivalent to verbal language. Multimodal transcription, on the other hand, produces precise verbal descriptions of individual modes but does not consider how they combine to make meaning. To sum up, we propose that employing these approaches within Translation Studies supports an idea of translation as a predominantly verbal activity and possibly even constitutes a methodological restriction that prevents us from developing more profound understandings of the role of other modes than the verbal in multimodal translation. We ask whether adopting approaches focusing on the specificities of sound and image developed within other disciplines, e.g. Film Studies, might render these modes more visible and more explicitly discussable and analyzable from a translational perspective.

4 Reflection on Our Own Research

In our research dealing with translation of multimodal materials – the role of images in technical texts and the role of sound in films – we have encountered questions about multimodality that the existing methods in the field do not seem to be able to fully answer. In the following sections we reflect on the methods we have set out to test in our own research efforts.

4.1 Examining Images in Technical Translation

Ketola’s research interest lies in considering the effect word–image interaction has upon the translator. Her research examines the translations of an illustrated technical text describing the operating principles of two ore beneficiation devices, translated from English to Finnish by eight Master’s level translation students. Ketola’s research employs one possible method of inquiring into the effect of word–image interaction on the translator’s choices, namely choice network analysis (Campbell 2000).

Choice network analysis sets out to compare the translations of the same source text by multiple translators in order to determine the different ways in which the same text may be interpreted. The method works on the premise that collecting different translation solutions into a network-like flowchart allows us to identify the similarities as well as the differences between them, in other words, to identify the range of possible solutions that were available to the translators when translating a specific source text item (Hale & Campbell 2002: 18). Comparing the translations of the same illustrated text segment hence reveals the range of the possible ways in which a particular group of translators extracted information from the combination of words and images. Ketola's study attempts to assess if the translation solutions provided by the translation students seem to be based on purely verbal information or on a negotiation of meaning from two different modes, in other words, considering if the images may affect the way in which the verbal text is translated.

An example of the results of the analysis, yet to be published, can be drawn from the choice network representing the translations of the phrase "The magnetic particles are separated from the rest of the stream as they adhere to the drum surface in the area of the magnet." Of particular interest in the network are the translations for the word *magnet*. The word is in the singular form in the verbal source text since it refers to the magnet unit of the separator as a whole. Yet, in the image of the source text, the magnet unit is depicted as clearly consisting of four individual magnets, coloured with eye-catching bright red. The choice network of the phrase displays that all seven translators who maintained the prepositional phrase translated *magnet* in the plural form (one translator omitted it altogether). It hence seems that the most obvious option the multimodal source text offered for the translators was in fact referring to the magnet in the plural form. The reason for this could well be that the visual source text so clearly depicts four individual magnets. The example hence illustrates how the visual source text may affect the choice between singular and plural forms of nouns.

Choice network analysis, too, has significant limitations that need to be acknowledged. The method is not able to access the reasons behind the translation solutions – which, admittedly, is an inherent limitation of all product-based studies investigating the cognitive processing during translation. Moreover, the method is not able to assess the level of conscious attention behind the translation choices – even though, evidently, the same holds for other methods used to study what goes on in the mind of the translator, such as eye tracking or keyboard logging. In the above example, for instance, there is no way of knowing if the decision to accommodate the verbal information with the visual took place consciously or unconsciously.

4.2 Phenomenological Analysis in Studying the Role of Sound in Film Translation

Only limited attention has been paid to the role of sound in Translation Studies so far. Kokkola’s research on the role of sound in film subtitling proposes phenomenologically informed audiovisual analysis based on theories developed within Film Studies (e.g. Sobchack 1992; Chion 1994) as a method that renders sound more explicitly analyzable in the context of film translation. Simply put, phenomenology is a study of embodied, lived experience. The analysis requires a “phenomenological attitude”, a particularly open way of looking at things, as free of preconceptions as possible. Philosopher Don Ihde describes phenomenology as follows:

[...] as a radical philosophy, phenomenology necessarily departs from familiar ways of doing things and accepted ways of thinking. It overturns many presuppositions ordinarily taken for granted and seeks to establish a new perspective from which to view things. (Ihde 1986: 17)

Phenomenology, in its capacity of being a study of possibilities or “a probing for what is genuinely discoverable and potentially there but not often seen” (ibid. 26) is particularly well suited for exploring the role of sound and moving image because, compared to words and still images, which are more objectlike, sound and movement are ephemeral and eventlike and thus often escape the researcher’s attention. The phenomenological approach starts from casting aside preconceptions about the phenomenon under study and initially *equalizes* all phenomena. Phenomenological study of film translation, then, looks at all modes of a film as equals. The analyst employs a reflective mode of looking and listening whereby any one of the modes has the potential to emerge as the most salient instead of being guided by the verbally-driven translational thought which privileges verbal language over other modes.

The phenomenological approach to film translation is based on the idea that the meaning of a film is created cross-modally through the viewer’s engagement with the film. Simply put, this means that, when co-presented, modes “transform” each other at the moment of the viewer’s encounter with the film in the sense that sound changes what we “see” and image changes what we “hear” (e.g. Chion 1994). For example, the same film scene might evoke very different meanings when co-presented with different kinds of music. Phenomenological film analysis is interested in effects – the way the film as a multimodal whole *appears* as a result of the interaction of the modes.

Kokkola’s research on film subtitling examining four films by Aki Kaurismäki and their English and German subtitled versions shows that subtitles, too, play a role in cross-modal perception. Thus, the presence of different sets of subtitles for the same film or the presence of subtitles compared with none at all influence the overall experience of the film. For example, if a foreign language scene, which is unsubtitled in the original film, is subtitled for some language versions, the affective effect of the scene can

change in a profound way. From a phenomenological perspective, the task of the film translator, then, is to identify the cross-modal effects and render them in translation.

By connecting film analysis to the embodied, lived experience, the phenomenological method offers a more nuanced, inclusive and rich description of the role of the different modes in film translation than models such as multimodal transcription that seek to produce generalizable data, but seem to downplay certain modes in the process. The phenomenological method is quintessentially relational and allows for analysis of the effects the modes create together and emphasizes the importance of understanding the film as a whole. Thus, it provides an in-depth understanding of how the different modes relate to each other, while models such as multimodal transcription are more concerned with the surface structure relations between the modes and are dealing with isolated fragments of a film in a rather atomistic fashion. However, the phenomenological method examines individual films and does not produce generalizable data. Moreover, analysis and interpretation of data as well as presentation of the results can be a complex and time consuming process, and only a limited number of films can be included in one study. Due to its openness to novelty, the phenomenological method can help identify new issues and ideas as they are emerging in a discipline and can, thus, contribute to the development of new theories and the discipline in general.

5 Learning to Think Multimodally

This article has argued for multimodal approaches to translation that acknowledge the essential role of sound and image in multimodal meaning making. We encourage future research to base itself on a multimodally-driven translational thought instead of the verbally-driven thought employed so far. We believe this might constitute a step towards the enlarged concept of translation discussed earlier and, eventually, even lead to the reassessment of the borders of the discipline.

We propose that introducing research methods not dependent on the linguistic-based description of modes may complement the existing methods, for instance, by relating the surface structure of the multimodal text with its perceived effect. We believe that interdisciplinary effort is required for this purpose and hence advocate the use of methods that have originally been developed within other disciplines for analyzing modes other than the verbal. We believe that adopting innovative methods may advance our understanding of multimodal phenomena in the context of translation and even provide an impetus for asking new questions and discovering new realities. Yet, we wish to stress the importance of critical methodological reflection – identifying the opportunities and limitations of each method, including the ones we are proposing – for greater understanding of knowledge construction within Translation Studies.

No single approach exhausts the phenomenon it addresses. We can, for example, examine perception of multimodal materials in the context of translation from the perspective of cognitive science or phenomenology, but neither approach captures the phenomenon in its totality or renders useless the questions asked by the other. Although both approaches address the same phenomenon, the subjective experience of multimodality addressed by phenomenology is qualitatively different from cognitive scientific accounts of multimodal experience seeking to explain the mechanisms of cross-modal perception. Thus, instead of treating methods as oppositional, it might be more illuminating to see them as complementary, contributing to the big picture of multimodality in Translation Studies.

Further, we also wish to emphasize the need to establish disciplinary practices that are themselves multimodal. An example of such a practice is the multimodal presentation of research. For instance, traditional print media reduces sound to verbal descriptions and sometimes visual presentations and is therefore not the ideal channel for publishing research on sound. We believe that the best way to present the analysis of film sound and experience would be to include video clips preserving the sound’s original context, which – technically speaking – would currently be perfectly possible with digital multi-format publishing practices. However, issues such as copyright permissions might complicate the publishing process.

To engage in critical reflection on particular Translation Studies methods addressing multimodality is to self-reflexively look at the engagement of Translation Studies with multimodal materials at a more general level as well. We can either take existing Translation Studies research as “given”, without problematizing its way of looking at the role of multimodality, or we can adopt a self-reflexive attitude towards knowledge construction in Translation Studies by looking at existing research on multimodality in order to examine its mode of engagement with multimodality. Each research method makes translation appear in a particular way instead of revealing “the truth” about the phenomenon. Thus, Translation Studies as a discipline can develop by self-reflexively turning in on itself and identifying the mechanisms of its own knowledge construction.

To examine translation in multimodal terms is not merely to further illustrate multimodal data by engaging in more detailed description of them, but to bring about a new articulation on them: a new order – a reconfiguration – of the modes, an active constitution of new thought which makes explicit what has previously been present only implicitly in current theory and practice of translation and largely learned by experience. This article is an invitation to think about multimodality differently, to create an opening of a particular kind towards how we conceive of its role in Translation Studies and to acknowledge the fact that the topic can be broached in more ways than one, and many possible worlds of multimodal translation are yet to be discovered.

References

- Alvesson, Mats & Kaj Sköldbberg (2000). *Reflexive Methodology: New Vistas for Qualitative Research*. London: Sage.
- Baldry, Anthony & Paul J. Thibault (2005). *Multimodal Transcription and Text Analysis*. London: Equinox.
- Campbell, Stuart (2000). Choice Network Analysis in Translation Research. In: *Intercultural faultlines. Research Models in Translation Studies I*, 29–56. Ed. Maeve Olohan. Manchester: St. Jerome Publishing.
- Chion, Michel (1994). *Audio-Vision: Sound on Screen* (Trans. by Claudia Gorbman). New York: Columbia University Press.
- Gibbons, Alison (2012). *Multimodality, Cognition and Experimental Literature*. New York: Routledge.
- Hale, Sandra & Stuart Campbell (2002). The Interaction Between Text Difficulty and Translation Accuracy. *Babel* 48(1), 14–33.
- Halliday, Michael A. K. (1978). *Language as a Social Semiotic*. London: Edward Arnold.
- Hesse-Biber, Sharlene Nagy & Patricia Leavy (Eds.) (2008). *Handbook of Emergent Methods*. New York: Guilford Press.
- Hirvonen, Maija & Liisa Tiittula (2010). A method for analysing multimodal research material: audio description in focus. *MikaEL, Electronic proceedings of the KäTu symposium on translation and interpreting studies 4*: 1–12. [cited 15.9.2015]. Available at: http://www.sktl.fi/@Bin/40698/Hirvonen%26Tiittula_MikaEL2010.pdf
- Ihde, Don (1986). *Experimental Phenomenology: An Introduction*. Albany: State University of New York Press.
- Jewitt, Carey (2009). An introduction to multimodality. In: *The Routledge Handbook of Multimodal Analysis*, 14–39. Ed. Carey Jewitt. London: Routledge.
- Kress, Gunther & Theo van Leeuwen (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Oxford: Oxford University Press.
- Kress, Gunther & Theo van Leeuwen ([1996] 2006). *Reading Images. The Grammar of Visual Design*. London & New York: Routledge.
- Sobchack, Vivian (1992). *The Address of the Eye: A Phenomenology of Film Experience*. Princeton, NJ: Princeton University Press.
- Susam-Sarajeva, Şebnem. (2000). A ‘Multilingual’ and ‘International’ Translation Studies? In: *Crosscultural Transgressions: Research Models in Translation Studies II. Historical and Ideological Issues*, 193–207. Ed. Theo Hermans. Manchester: St. Jerome.
- Taylor, Christopher J. (2003). Multimodal transcription in the analysis, translation and subtitling of Italian films. *The Translator* 9(2), 191–205.
- Taylor, Christopher J. (2009). Pedagogical tools for the training of subtitlers. In: *Audiovisual Translation: Language Transfer on Screen*, 214–228. Eds. Jorge Díaz Cintas & Gunilla Anderman. London: Palgrave Macmillan.
- Tymoczko, Maria (2007). *Enlarging Translation, Empowering Translators*. Manchester: St. Jerome.
- van Meerbergen, Sara (2009). Dutch Picture Books in Swedish Translation. Towards a Model for Multimodal Analysis. *Translation and the (Trans)formation of Identities. Selected Papers of the CETRA Research Seminar in Translation Studies*. [cited 15.9.2015]. Available at: <https://www.arts.kuleuven.be/cetra/papers/files/van-meerbergen.pdf>