Understanding meaning-making through technology use—a multimodal layer perspective

Karoline Schnaider

Department of Education, Umeå University, karoline.schnaider@umu.se

Abstract

Previous research has highlighted that technology implementation in schools may lead to increased complexity, as digital hardware and software offer a variety of possibilities for sign-making activities. Moreover, recent studies argue that since classroom practices are facilitated increasingly by screen-based activities, digital technology opens a multitude of ways to represent meaning, as an abundance of sign systems becomes available for communication through various digital visual user interfaces (DVUIs) (Jewitt, 2017). In addition, research indicates that technology implementation has a strong impact on school practice (Säljö, 2013) and that knowledge on how to take advantage of technology in learning settings from a more comprehensive perspective is needed (Bezemer & Kress, 2016). To gain a more comprehensive picture of technology use in educational environments, the main goal of the thesis is to explore the use of hardware and software by teachers and students in sign-making activities from a multimodal layer perspective. The main aim of this paper is, in particular, to discuss how multimodal methodology can be used to explain detailed aspects of technology use in networked learning (NL) settings. Concerning the various means used in school and their affordances in semiotic mediation (Norman, 2007; Wartofsky, 1979), all are considered in relation to the users and results of use. From a technology perspective, the multimodal layers, therefore, include things-to-things, things-to-human/human-to-things and human-to-human connections (Bonderup Dohn, Cranmer, Sime, de Laat and Ryberg, 2018; Goodyear, Carvalho & Bonderup Dohn, 2014) and focus on technologies, communication resources (i.e. sign systems), representations and activities. The technologies and their functions are therefore regarded as important. In addition, the multimodal layers relate to the semiotic properties of technology, how they inhere and prompt sign systems in different ways as interpreted by the actors and are reshaped into modes of representation in different activities. The conclusion is that multimodal methodology, particularly the multimodal layer approach, seems to be beneficial to unpack the relationships and connections between the means used and the actors in NL environments via its coherent approach. A greater understanding of the detailed aspects of technology use in teaching and learning may also be obtained if the existing multimodal layers are accounted for and connected. Insights can guide stakeholders on how to integrate technology in future practices and inform technology choices in relation to specific activities.

Keywords

Multimodality, multimodal layers, digital technology, hardware, software, use, activities, representations.

Introduction

This introductory part of the text will describe the theoretical frameworks that form the basis for the thesis. Thereafter, methodological considerations will be presented, as well as examples of data and results offered by the multimodal layer approach. The text ends with a discussion and conclusions.

The problem the author tries to find solutions to is the inattention in contemporary research to a comprehensive understanding of *meaning-making through technology use* needed to truly identify the many aspects that work together in a learning environment. Hence, the purpose of the thesis is to give detailed insights on the use of technology by teachers and students in sign-making activities in educational settings; or, more specifically, to provide information on technology use based on the development of a framework called multimodal layers that can provide insight into the complexity of networked learning (NL) settings and help inform stakeholders on how to choose and use technology for meaning-making. The research question that guides the thesis is: From a

multimodal layer perspective, how are digital hardware and software used by teachers and students in sign-making activities? The empirical studies are intended to inform and develop the framework of multimodal layers by applying it in investigations and by offering understandings of meaning-making through technology use by a comprehensive approach that encompasses all salient components in the use of DVUIs. This paper discusses some approaches that will be used in case study research strategies, such as data collection instruments and methods for analysis in relation to the proposed framework.

From these starting points, the project includes a meaning-making and a technology focus. The focus is on understanding meaning-making through the use of technologies by teachers and students to guide future strategies for technology selection and use. From an NL perspective, the multimodal layer approach is used to connect all components defined as salient in an NL setting by its comprehensive focus on things-to-things, things-to-human/human-to-things and human-to-human connections (Bonderup Dohn, Cranmer, Sime, de Laat & Ryberg, 2018). Hence, the multimodal layer framework will be used to understand the conflations between technologies and sign-systems (DVUIs) and between teachers and students and DVUIs (Bezemer & Kress, 2016), and add to existing NL understandings by illuminating the connections by focusing equally on meaning-making and technologies. The multimodal layer approach will contribute with a detailed perspective that considers not only how various components are connected, but also how differences between technologies encompassed by hardware and software are jointly configured in use from their functions and visual properties (such as sign systems) and how teachers and students make meaning in use from their features. From that point of view, and in line with NL research, *meaning-making through technology use* is regarded here as a circular activity—from the technologies' prompting of functions and sign systems conflated with humans' representations, to the signs of meaning made. These contributions to understanding the intricacy of NL settings will be further explained below.

Theoretical alignments

To illuminate the details connected to meaning-making through technology use, a set of theoretical frameworks will guide data collection and analysis and be briefly accounted for in this section.

The overall sociocultural ontology used in the thesis can explain the foundations of meaning-making through technology use in school settings. The initial ideas focus the intimacy between two focal components: humans and artifacts. Mainly from the philosophical idea that artifacts are vital to us—so important that they evoke emotions, memories and stories and are used as a means for work, thinking, communication and learning (Säljö, 2013; Vygotsky, 1978; Wartofsky, 1979). Following Wartofsky (1979), both physical (objects) and intellectual production (speech, writing and other sign systems) and the activities through which they come into life are examples of artifacts and how meaning is made in the world. Such ontological arguments can be applied when understanding technology use in today's educational environments, for instance, to view technologies as important objects, how they are important, and the multiple sign systems they bring about in meaning-making practices (Selander & Kress, 2017). On one hand, these claims emphasize the non-dualism between humans and the artifacts they use to make meaning. The claims also put the focus on technologies and sign systems as vital for human existence, both fixed into physical artifacts and communicated in real-time activities. Hence, both technologies and sign systems have major impacts on learning environments through their connections and conflations with humans and through their various affordances and constraints on meaning-making activities between people (Gibson, 1986; Norman, 2007; Kress, 2017; Selander & Kress, 2017).

Understanding the link to and importance of artifacts in human meaning-making as these theories emphasize is significant. The theories help us explain why instrumental technology use proven by recent studies to exist in school and associated with tech takeover (Kervin & Mantei, 2017) or gift wrapping (Fischer, 2018) might hinder meaningful technology use. Second, these pediments have implications for how technology implementation and use in learning settings must be understood, because the use of technologies and sign systems (DVUIs) is not a random enterprise. Choices based on understandings of which one, how they were manifested and designed and what limitations and possibilities humans find in them for sign-making activities is crucial for future successful use. Third, since technology use is not an arbitrary practice, and according to these founding arguments, all "artifacts" are regarded important and are connected within themselves and to humans, why has research to date partially regarded some technologies (e.g., software) and sign systems (e.g., language) as more important (Schnaider, Gu & Rantatalo, 2020, unpublished manuscript) and almost completely disregarded others? In particular, how are technologies and sign systems jointly configured in use? Hence, studies need to take all salient components into consideration when using DVUIs from a detailed and connected perspective, such as hardware and software and the sign systems they inhere and evoke in humans, represented in different activities. Stressed by, for instance, Ravelli & van Leeuwen (2018), it is vital to understand how various types of digital hardware function, how they prompt sign systems and how they configure the software and thereby change how

software is displayed. The impact is on the sign-making activities and what meaning can subsequently be made. From these starting points, an NL perspective that focuses on a non-dualistic view of meaning-making through technology use that includes all aspects that work together in and around DVUIs is important, here encompassed by the multimodal layer approach.

The multimodal layer approach

The thesis is intended to unravel detailed aspects of meaning-making through technology use, perspectives highly sought in research on IT and learning (The Swedish research council, 2019). In the epistemology of social semiotics and multimodality, various detailed aspects that integrate the non-dualism of humans and artifacts are taken seriously (Bezemer & Kress, 2016; Jewitt, 2017) and will be the theoretical lens though which knowledge can be gained in the thesis. The multimodal layer framework puts focus on the connection between technologies, sign systems and human's representations in activities, and relate technology landscapes to the adoption of strategies to access detailed components in meaning-making practices (Jewitt, 2008; 2009), highly needed to properly understand technology use in classroom settings. With the increasing entry of digital technologies in school, there have been a gradual replacement of books with digital technologies. This change means that the world went from being described (through language) to being visualized (by screens and other modes in combination) (Kress, 2010). As such, DVUIs are perceived and interpreted from their features, emphasizing the visual.

Although the aforementioned theories stemming from the sociocultural field of knowledge have the power to explain phenomena and properly illuminate the relevant object of knowledge—meaning-making through technology use—these theories lack insights on how to apply them easily in combination to comprehensively understand the technologies, prevailing sign systems and human representations in activities. These theories also fail to clarify how the components of DVUIs are linked and how they can be explained from their various connections in representations and activities. Hence, there is a knowledge gap in both research and practice on how to analyze technology use from a comprehensive perspective that encompass the technologies and the meaning-making practices, something that the multimodal layer perspective attempts to cover.

The concept of multimodal layers is derived from the notion that the digital world is multimodal per se (Jewitt, 2017) and gets extended to explain meaning-making through the use of technology by teachers and students, not only taking into account the many modes of representation (Kress, 2017). The framework is developed through and applied to empirical studies to provide a detailed perspective of these phenomena. At the same time, it is intended to unravel the complexity and existing opportunities that digital technologies bring to learning settings by connecting all salient components related to the specificity of the activity. As such, the multimodal layer approach strive to explain technology use by including five components for analysis; *technologies*, technologies' *functions*, technologies' *semiotic properties*, the *modes of representation* and *activities* (Bezemer & Kress, 2016; Ravelli & van Leeuwen, 2018), additionally explained below.

The layers

In this section, the five layers are briefly explained and exemplified:

- *Technologies*: As important objects (Wartofsky, 1979), all digital technologies, hardware and software together, compose the DVUI that teachers and students encounter in meaning-making practices. Based on things-to-things connections, hardware and software are important to regard both separately and jointly, especially from a visual point of view in relation to sign-systems. This means that the thesis is delimited to regard only visual entities in the use of hardware and software and does not consider the hardcore technical functions that work behind the scenes, or the fact that hardware also needs software to operate and vice versa. This delimitation is also of relevance to the division into hardware and software made in the thesis—they are visually separated and connected from how they display different functions and semiotic properties (sign systems see below), and the displaying of software changes with combinations of hardware devices (Bezemer & Kress, 2016; O'Halloran, 2013; Ravelli & van Leeuwen, 2018). A simple example of that is the visual difference between using a smartphone and a laptop with an Outlook calendar. Some of those differences exist in the visual functions of the hardware, such as the accessories, screen and physical/virtual keyboard, as well as their semiotic properties such as the size and shape. Regarding software, the functions for adding activities on a smartphone or a laptop address—for instance, the functions of touch and navigation with the cursor and semiotic properties of colors, lines, spaces and framing—are altered between and within these technologies, impinging on the interplay between them and, in the end, the meaning made (Ibid.) Meaning-making through technology use will, therefore, be investigated at the intersection of and configuration between hardware and software as constituitive of DVUIs from a visual point of view.

- *Functions*: This aspect is linked to technologies' technical functionalities as visually perceived. For instance, that the Enter key is used to confirm requests or to make spaces, a new post is added in an Outlook calendar by double clicking in the framed space of a specific date or that the slide bar zoom-function in a program can alter the size of what is displayed.
- *Semiotic:* These properties are related to technologies' inherent sign systems (modalities/semiotic resources, Kress, 2017) and are intimately, visually connected to the layer functions (cf. things-to-things). For instance, the function of the hardware screen to display simultaneously determines the size, framing and space of what is displayed, hence the semiotic properties.
- *Representations:* From a non-dualistic perspective, humans make representations from their various conflations with the world as signs of what kind of meaning is made. In technology use, these representations are made when prompted features gets interpreted by individuals and come about in activities, explicating the affordances of the technologies (Gibson, 1986; Norman, 2007; Kress, 2017). Hence, humans interpret what they experience and perceive as important in handling technology and sign systems, integrating these actions and impressions into cognitive processing (Kress, 2010). Some of the thoughts, emotions, etc. are expressed externally by different outputs that refers directly to the context in which the meaning was made (Kress, 2010; Vygotsky, 1978). Hence, to understand meaning-making through technology use, it is reasonable to ask which of these prompted features led to the specific representation. Representations are made visible to humans through actions and semiotic resources/modalities, such as colors, lines, spaces, shape and size, speech, gestures and images (Kress, 2017).
- *Activities:* Representations are arranged together with technologies and sign systems in higher-order actions, called activities, and configured in various ways depending on the context, subject and discourse (Norris, 2014). In terms of discourses, a math lesson would have a specific set of material and intellectual resources, while a lesson in physical education would involve others.

Methodological considerations—theoretical frameworks and case study strategies

This section of the paper provides a summary of research strategies, data collection instruments and methods linking them to theoretical frameworks in a coherent methodological structure. The description is intended to clarify how the detailed aspects of NL can be elucidated from multimodal strategies.

To get comprehensive insights on meaning-making through technology use from multimodal layers, the actors, technologies and contexts are of great importance. Since the main point of the thesis is to describe, explain and understand through in-depth investigation, the case study research strategy will inclusively approach the object of knowledge from two important case study ideas: the *contemporary phenomena* of technology use and *real-life settings,* such as school. Empirical studies will be focused on three cases constituted by different primary schools and several units of analysis—one student and one teacher in each school (Aspers, 2011).

Case study research consists of exploring *how* phenomena work, and the main focus of this thesis is to explore how technologies are used and form implications for the researcher during data gathering. Hence, the researcher does not set out to control or influence the object of knowledge but instead places the researchers in the study context with methodological tools, such as video recordings, observations and interviews (Norris, 2015); associating these aspects with a case study strategy allows the formation of comprehensive descriptions of people and objects in their settings (Aspers, 2011; Jensen & Sandström, 2016). Triangulation is common in a case study strategy, involving the use of several data sources, data collection methods and theoretical perspectives, here linked within the framework of multimodal layers.

The social semiotic multimodal approach emphasizes a range of factors, such as technologies, sign systems, representations and activities. Its methodology, therefore, extends and reshapes what is possible in retrieving and analyzing data, as it embraces all aspects (Jewitt, 2009). In line with case studies and research aims, the goal is to get a picture of the possible variations that exist at the intersection of various layers while developing the multimodal layer framework. In recognition of variations, it may be possible to detect trends in use; for example, if specific layers are addressed frequently or specific technologies or sign systems are present in particular activities and representations. Thus, from the notions of layers, the main focus of the case study approach is to gather qualitative data and, using mapping methods, pair teachers' and students' actions and activities in the use of technologies to the specificity of the technologies and sign systems used.

Sampling and data gathering instruments

For starters, something brief must be mentioned about the sampling procedure. Purposive sampling will be made from 1:1 classrooms and teachers and students who are experienced in managing digital technology on a daily basis. From the idea that various types of hardware matter and to get a broad representation of different technologies, classrooms that are representative of different types of 1:1 technology (hardware) are important. This means that data collection is based on the use of hardware (laptops, tablets and smartphones) as an independent variable and on different settings and subjects. Due to the huge amount of data that the proposed research strategy can provide, three students and three teachers will be selected as representative users of the appointed hardware technology.

Video recording will be used as a data collection instrument because it can capture detailed aspects associated with NL environments. The visual mode of video recording can depict interplay with technology and, together with the functions that allow the viewer to review the material repeatedly, allow the researcher to obtain multimodal data. It is important to make decisions on whether to use two or more cameras before the actual recording; for example, one that takes in the whole environment and one that targets something specific—the placing of the cameras is of great importance (Jewitt, 2009; Kress, Jewitt, Ogborn & Tsatsarelis, 2014; Norris, Geenen, Metten & Matelau, 2015).

Video recordings can be linked to other data successfully from observations and interviews (Bezemer, 2015). Although everything that happens in a classroom cannot be captured, transcribed or analyzed, systematic observation notes from multimodal layers can be used to clarify and add information to the video recordings. In addition, semi-structured interviews can establish teachers' and students' thoughts on the specificity of the used technologies in relation to what representations are made in activities (Aspers, 2011). The interviews can be done individually with informants while video recordings are reviewed. This method requires several pre-selected episodes from which the interview guide can identify the use of technologies based on multimodal layers. For example, how are the technologies in use configured based on the goals of the activity? What functions of the technologies (hardware and software) were important and how did they support the user? What role did the semiotic property of ... play in representation ...?

Data analysis

In line with the case study strategy, data analysis will focus on how technologies are used by teachers and students from the theoretical framework of multimodal layers and through methods for mapping and pairing data from the content of the text-based transcripts (Silverman, 2006). From a parallel coding process of data obtained from the three cases, transcripts of video recordings with recorded sequences, transcribed observation notes and interviews will be analyzed in qualitative data analysis software from the following categories: settings, users, hardware/software, technologies' functions, technologies' semiotic properties, representations and activities. The advantage of such a procedure would be a comprehensive recognition of detailed aspects salient in technology-rich educational settings.

The following concepts can shed light on the categories for analysis:

- Configuration, foregrounding and backgrounding in interplay (Norris, 2014, adapted to technology use) can help analyze the importance of the technologies, their functions and semiotic properties and how they come about in interplay in specific representations and activities.
- Semiotic resources and modalities (Björkvall, 2009; Kress, 2017; van Leeuwen, 2005) help analyze the semiotic properties of technologies as well as the representations made by individuals.
- Production and consumption (Kress & van Leeuwen, 2001; Kress, 2010) help analyze overall activities in learning design and sign-making activities.

Expected findings

This section exemplifies how a mapping method could be used to analyze data and how data can be paired based on the idea of multimodal layers (see table 1 below). Fictitious data exemplifies the layers, the mapping and pairing research design and data-gathering instruments based on an image, connecting two students' uses of technologies in an NL environment.

Table 1: Example of multimodal mapping

Transcriptions:	**Snapshot:**  Picture retrieved from: https://pxhere.com/en/photo/619706	**Video recording:** Students talking about the story while pointing at the screen while inserting an image into the essay.	**Observation notes:** Students used the semiotic property (the size of the screen) and the zoom function to highlight a part of the essay while connecting it to the image.	**Interviews:** Students particularly emphasized that the actions that took place in discussions involved the screen display. The colors were important in highlighting parts of the text to get a cohesive story.
Multimodal layers:	**Participants/ actors:** Student to student.	**Technologies:** Desktop computer, animation software.	**Functions and semiotic properties:** Using the screen to display joint content, keyboard for typing of text/navigation with hotkeys and zoom. Inserting and drawing functions in software, such as the toolbar, to change colors and layout. Semiotic properties: Size (screen), framing and lines (software).	**Activities and representations:** Production of a multimodal essay in sign-making activities. Modalities: Speech, gestures, images, layout and text. Semiotic resources: colors, framing and space.

Results

The goal of mapping data and pairing the multimodal layers that are salient in NL settings is to understand meaning-making through technology use by teachers and students from a comprehensive, detailed perspective. Hypothetically and with a little imagination, mapping from a video-recorded snapshot of interplay in a specific learning environment can illuminate a collaborative activity involving students' coproduction of a multimodal essay. The students engage with DVUIs from the functional advantages of the desktop computer screen and a software animation program to display shared representations, especially with the help of the zoom function. These functions are foregrounded in use with the software insertion functions. The semiotic properties connect the hardware and the software in use, such as the screen's framing properties and the zoom function that resizes what is displayed, to the other semiotic properties of the software, such as lines and spaces. To represent their thinking in signs, students used the modalities of speech and gestures to negotiate the visual semiotic resources represented in the layout, such as colors and text. These functions/properties and representations can be seen as a higher level of configuration and are therefore more important in relation to this specific activity.

Discussion and conclusions

The case study research strategy and the data collection instruments described above enable in-depth, detailed data on multimodal layers. Proposed strategies and instruments will most likely render an enormous amount of data that is difficult to handle, therefore, making the research difficult to manage from time to time. On the other hand, the limited number of participating informants can resolve this challenge to a certain extent. Due to small sample sizes and the situated nature of the obtained data, the generalizability and transferability will be

deliberately downplayed (Aspers, 2011; Jensen & Sandström, 2016). By selecting appropriate data collection instruments, such approaches can still affect reliability, even if they are considered and built into the research through the development of detailed data coding schemes and categories for analysis based on multimodal layers. Validity can, on the other hand, be discussed in terms of a coherent methodological approach offered by theoretical frameworks, strategies, instruments and methods tried and frequently used in multimodal research and renewed in this research by the multimodal layer approach. Suggested methods can also help the researchers stay close to the goals of the research in data processing through iterative processes (Cutajar, 2018). However, as the multimodal layer approach is used to clarify the complexity of NL settings, a strong focus on a specific set of theoretical perspectives may prevent or restrict a researcher from thinking differently and lead to a one-sided focus on certain starting points (Eriksson Barajas, Forsberg & Wengström, 2013).

From the above mapping method, it becomes apparent that by mapping and pairing the multimodal layers in analysis, one can gain a comprehensive understanding of things-to-things, things-to-human/human-to-things and human-to-human connections in activities from a visual point of view (Bonderup Dohn, Cranmer, Sime, de Laat and Ryberg, 2018; Goodyear, Carvalho & Bonderup Dohn, 2014). This would be a contribution to NL research and a perspective sought in the thesis in order to understand meaning-making through technology use. By paying equal attention to both meaning-making and technologies, as portrayed in the example above, details are unraveled and connected by focus on students' choices regarding technologies and sign systems constituitive of DVUIs based on their need to represent meaning in sign-making activities (Jewitt, 2009). From a non-dualistic perspective that regard all "artifacts" as equally important, it becomes apparent that by mapping and pairing layers, certain functions and semiotic properties in these specific technologies supports and is considered more important in relation to certain representations and activities (Norris, 2014). Moreover, there will probably be a rich variety of combinations when technology use is mapped and paired in the proposed way. The multimodal layer perspective will undoubtedly also highlight certain patterns or trends in use. On the other hand, how the detailed aspects provided by this framework through combinations of pairing can provide a clearer understanding of technology use remains to be elaborated and developed, one aim of the thesis. Hence, further investigation is needed into how new knowledge can be obtained on meaning-making through technology use by teachers and students in sign-making activities from a multimodal layer perspective and how such a framework can be used more successfully. New knowledge might also be able to guide the future use of technologies in NL settings. If these insights reach stakeholders, they may also prompt questions, such as: If other technologies and functions had been used instead, how would they support representations and activities differently? From these ideas, sign-making activities can benefit from a learning design in which the multimodal layers are considered initially and throughout the process in relation to desirable results. Hence, increased awareness of the importance of various technologies and sign systems to meaning-making processes can guide choices of which to use and how to use them.

The author of this paper suggests that there is no middle way to fully understand the interconnected nature of meaning-making through technology use and the intricacy of the rhizomatic networks that educational environments entail. Therefore, if a multimodal layer approach is applied in research and practice, new and comprehensive findings can be made and new strategies for use can be applied in education, insights that can possibly also inform design of new technologies.

References

Aspers, P. (2011). Entografiska metoder. Stockholm: Liber.

Bezemer, J. (2015). Multimodal transcription: A case study. In: S. Norris & C. D. Maier (Eds.), *Interactions, images and text. A reader in multimodality*. Boston: Walter De Gruyter.

Bezemer, J. & Kress, G. (2016). *Multimodality, learning and communication. A social semiotic frame*. London: Routledge.

Björkvall, A. (2009). *Den visuella texten. Multimodal analys i praktiken*. Stockholm: Hallgren & Fallgren.

Bonderup Dohn, N., Cranmer, S., Sime, J-A., de Laat, M. & Ryberg, T. (2018). *Networked Learning: Reflections and Challenges*. Cham: Springer International Publishing.

Cutajar, M. (2018). Variation in Students' Perceptions of Others for Learning. In, N. Bonderup Dohn, S. Cranmer, J-A. Sime, M. de Laat & T. Ryberg (Eds.), *Networked Learning: Reflections and Challenges*. Cham: Springer International Publishing, pp. 79-94.

Eriksson Barajas, K., Forsberg, C., & Wengström, Y. (2013). Systematiska litteraturstudier i utbildningsvetenskap, vägledning för examensarbeten och vetenskapliga artiklar. Stockholm: Natur & Kultur.

Fischer, G. (2018). *From Renaissance Scholars to Renaissance Communities: Learning and Education in the 21st Century*, Retrieved 2019-12-04, from l3d.cs.colorado.edu.

Gibson, J. J. (1986). *The ecological approach to visual perception. New York*: Psychology press.

Goodyear, P., Carvalho, L., & Bonderup Dohn, N. (2014). *Design for networked learning: framing relations between participants' activities and the physical setting*. Proceedings of the 9th International Conference on Networked Learning, Edinburgh, Scottland.
Jensen, T. & Sandström, J. (2016). *Fallstudier*. Lund: Studentlitteratur.
Jewitt, C. (2008). Multimodality and Literacy in School Classroom. *Review of Research in Education*, 32(1), 241-267.

Jewitt, C. (2009). Technology, liceracy and learning. A multimodal approach. Oxon: Routledge.
Jewitt (Eds.). (2017). *The Routledge Handbook of Multimodal Analysis*. New York and London: Routledge.
Kervin, L. & Mantei, J. (2017). Children creating multimodal stories about a familiar environment, The Reading Teacher, 70(6), 721-728, doi: 10.1002/trtr.1556.
Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
Kress, G. (2010). *Multimodality. A social semiotic approach to contemporary communication*. London: Routledge.
Kress, G., Jewitt, C., Ogborn, J. & Tsatsarelis, C. (2014*). Multimodal teaching and learning. The rhetorics of the science classroom*. London: Bloomsbury.
Kress, G. (2017). What is mode? In: C. Jewitt (Eds.), *The Routledge of multimodal analysis* (p. 60-75). London: Routledge.
Norman, D.A. (2007). *The design of future things*. New York: Basic books.
Norris, S. (2014*). Analyzing multimodal interaction: a methodological framework*. London: Routledge.
Norris, S. (2015). Developing multimodal (inter)action analysis: A personal account. In: S. Norris & C. D. Maier (Eds.), *Interactions, images and text. A reader in multimodality*. Boston: Walter De Gruyter, pp. 127-136.
Norris, S. Geenen, J., Metten, T. & Matelau, J. (2015). Studying social actors: Some thought about ethics. In: S. Norris & C. D. Maier (Eds.), *Interactions, images and text. A reader in multimodality*. Boston: Walter De Gruyter, pp. 127-136.
O'Halloran, K. (2013), Multimodal analysis and digital technology. In Montagna, E. (ed), *Readings in Intersemiosis and Multimedia*. Israel; IBIS Editions, pp. 35-53.
Ravelli, L.J. & van Leeuwen, T. (2018). Modality in the digital age. *Visual Communication*, 17(3), 277–297. doi: 10.1177/1470357218764436
Schnaider, K., Gu, L., & Rantatalo, O. (2020). *Understanding technology use from multimodal layers: a research review*. Unpublished manuscript. Department of Education: Umeå University.
Selander, S., & Kress, G. (2017). *Designs för lärande*. Lund: Studentlitteratur.
Silverman, A. (2006). *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*. London: Sage.
Säljö, R. (2013). *Lärande och kulturella redskap. Om lärprocesser and det kollektiva minnet*. Lund: Studentlitteratur.
The Swedish Research Council. (2019). *Forskningsöversikt 2019. Utbildningsvetenskap*. Retrieved 2019-06-05, from https://www.vr.se/analys-och-uppdrag/om-svensk-forskning/forskningsoversikter-2019.html
van Leeuwen, Theo. (2005). *Introducing social semiotics*. London: Routledge.
Vygotskij, L. (1978). *Mind in Society: The Development of Higher Psycological Processes*, edited by M. Cole, V. John-Steiner, S. Scribner and E. Souberman. Cambridge, Mass.: Harvard University Press.
Wartofsky, M.W. (1979). *Models. Representation and scientific understanding*. Dordrecht: D. Reidel Publishing company.