

Toward a Theoretical Framework of Multimodal Discourse Analysis of Criminal Courtroom Discourses

Wang Haiping

School of Foreign Studies
East China University of Political Science and Law
Shanghai, China
cathy_pingcn@163.com

Abstract—Currently, courtroom discourses are increasingly being mediated online, cases previously reserved for legal professionals in the judicial circle are now available in the public domain. Since these video-taped trials are now accessible for research and critique, a multimodal approach is needed to analyze the complexity of meaning building practices in courtroom discourses. In the current exploratory study, a tentative theoretical framework of multimodal discourse analysis of criminal courtroom discourses is designed in order to approach criminal courtroom discourses in a critical manner. Communication in the courtroom is complex, and if we need to understand courtroom interactions in all its complexity, all modes of communication should be put into consideration.

Keywords—*multimodal discourse analysis; criminal courtroom discourses; theoretical framework*

I. INTRODUCTION

Previously, legal discourse analysis mainly relies upon verbal texts, including judicial decisions, legal documents, etc. Legal discourse analysis is considered static and unidimensional. In terms of audio and video files, forensic linguistics normally transcribes the audio or video documents into verbal texts in order to carry out analytic commentaries. Up till now orthographic transcription is still the principal means of legal discourse analysis. But the orthographic transcripts of audio and video files can be inadequate and unreliable because lots of information has been lost in the process of transcribing [1]. On the other hand, in order to make legal procedures more transparent and legal education more authentic, with the help of the Internet, more courtroom discourses are mediated online via courtroom portal sites. Cases that are supposed to be open to the public are now available online in the public domain. Nevertheless, we lack a feasible way to analyze those multimodal legal documents. An applicable framework that addresses critical analysis of courtroom discourses is urged so that all modes of communication in the courtroom can be analyzed for more comprehensive meaning-building inside the courtroom.

II. MULTIMODAL DISCOURSE ANALYSIS

A. Systemic Functional Theory and Modality

Mode is a socially shaped and culturally given resource for making meaning. Image, writing, layout, music, gesture, speech moving, image, soundtrack are examples of modes used in representation and communication. [2] Modality can be viewed as semiotic resources for meaning construction, and multimodality indicates that meaning can be constructed and interpreted in different modes. Halliday (1978) argued that language is a social semiotic and hence social semiotics has become the theoretical foundation where most multimodal discourse analysis takes place.[3] Halliday's social semiotic approach was developed as systemic functional theory (SFT), which views semiotic resources as systems of meaning to fulfill various functions in human communication. This school of research on multimodality is represented by Kress, G. & van Leeuwen (1996, 2006), O'Halloran (2005, 2011, 2015, 2016), Baldry & Thibault (2006), Bateman (2008), Bednarek & Martin (2010), to just name a few. Kress, G. and van Leeuwen, T. (1996, 1998, 2001, 2002, 2006) studied the relationship between modalities and media. They described the concept of multimodality and explained the varied forms of meaning making that extend beyond language and enhance the semiotic process. They also emphasized in the age of digitalization, the different modes have technically become equal at some level of representation and hence provided the approach by which we might have "not only a unified and unifying technology, but also a unified and unifying semiotics". [4]

B. Multimodal Discourse Analysis

In light with the development of SFT, the multimodal social semiotic approach facilitates modelling and analyzing of multimodal texts, interactions and events that involve not only language, but also other semiotic resources such as images, sound, actions, etc. [5] For example, O'Halloran, K. L. (1999, 2000, 2003, 2004) has done a systematic series of researches on systemic functional perspective of multimodality, multimodality in mathematics discourse, multimodal literacy in pedagogy, software development of multimodal analysis and visualization for critical thinking and so forth.

This research was funded by Chinese MOE (Ministry of Education)
Project of Humanities and Social Sciences (Project number: 16YJC740069)

C. Software for Multimodal Discourse Analysis

In the modern digital world, computers provide us with physical conditions for storing and processing big data of multimodal information, including video, audio and textual files. Multimodal data-mining becomes promising, yet challenging. A variety of software, such as ELAN, PRAAT, MMAV (Multimodal Analysis Video), etc. enable researchers to segment and annotate static and moving images, sound streams and graphic texts. With video files, researchers can handle static and dynamic data synchronically across different modes.

For example, MMAV is a type of interactive software for annotation and analysis of vides, sample analysis and ready-made templates to facilitate teaching and learning about language, image and audio resources in videos. Visualization shows relative time for combinations of multimodal choices and adaptable concepts and frameworks for systematic analysis and interpretation for critical thinking. [6]

To summarize, multimodal discourse analysis after ten years of development, has gone beyond the scope of linguistics and has been considered interdisciplinary by nature [7]. Multimodal discourse analysis approach can be used to generate the construction of meaning beyond speech simultaneously and the establishment of a multimodal corpus can help researchers carry out quantitative research to describe the systems of multimodal data and to identify trends or patterns of them [8].

III. CRIMINAL COURTROOM DISCOURSES AND MULTIMODALITY

A. Criminal Trials and Multimodality

In Anglo-American legal system, take American criminal courts for example, the establishment of justice system in the criminal court is realized by an elaborate system of evidentiary rules and courtroom procedures. Nevertheless, an American criminal trial is much more than a sum of it evidentiary parts and it is actually a theater in which the various courtroom actors “play out the guilt or innocence of the defendant for the trier of fact to assess” [9]. Levenson (2008: 574) also argued that in reality, the outcome of a criminal case would be affected by various factors that are not technically hard evidence, for example, the quality of lawyers’ presentations, the appearance and reaction of the defendant in the courtroom and even the presence of the victim’s representativeness.

In this case, criminal courtroom discourses not only include linguistic features but also other modes of information. Multimodal discourse analysis can be used to study interactions in criminal trials including the prosecutor, the defense counsel and the judge through verbal and nonverbal communication as semiotic resources that contribute to the allocation of the power of discourse in the court [10]. In line with this research paradigm, Matoesian (2008: 195) analyzed the multimodal communicative practices, especially facial expressions and body alignment and realignment of a crucial witness in a rape trial and explored how she implemented extralinguistic resources to negotiate power and epistemological relations in

the trial with the attorney to co-construct her unique witness identity. [11]

B. Existing Criminal Discourse Analysis

However, the existing multimodal analysis of criminal courtroom discourse is quite rare. Up till now, most of the courtroom discourse research centers upon analysis of individual cases. Some probe into the relationship between courtroom discourse and power to some depth (e.g. Cotterill, 2003); some emphasize more on the equilibrium of power in the court, including the study on the impact of factors such as race, ethnicity, etc. on power manipulation in a criminal trial (e.g. Labov&Harris, 1994); studies on unequal power between females and males especially in rape trials also draw forensic linguists’ attention (e.g. Matoesian, 1993, 2001; Conley&O’ Barr, 1998). There are sporadic studies that involve multimodality, but the systematic research on criminal courtroom discourses via the multimodal-related approach is still on the way.

IV. A TENTATIVE THEORETIC FRAMEWORK

A. The Theoretical Foundation

A limitation that has been repeatedly mentioned in the field of legal discourse research pertains to the static rather than interactive analysis of the language. Besides the linguistic features as the main focus of research, a comprehensive analysis that encompasses both linguistic and nonlinguistic practices including eye-contact, facial expressions, bodily conduct and physical settings is needed for modern legal discourse research. In order to carry out further systematic research on multimodal courtroom discourse analysis, we need to construct a tentative theoretic framework for multimodal discourse analysis of the communication that occurs in courts. Since criminal courts and civil courts are distinctive from each other in terms of litigation procedures, we focus on the theoretical framework for multimodal discourse analysis of criminal trials.

As is central to social semiotic theory, each instance of text involves three strands of meaning simultaneously: a. experiential and logical meaning, which interprets our experience of the world and the logical relations herein; b. interpersonal meaning, which focuses on negotiation of social relations and expression of attitudes; c. textual or compositional meaning, which pertains to coherence of messages and their relevance to the context. [12]

B. The Tentative Framework

In order to make sense of the criminal court discourse, key modes that occur in a criminal court that contribute to different layers of meaning building should be pinned down first. Therefore, in light with the meaning-making purpose of social semiotic theory, a video clip recorded in a criminal court should contain modes such as verbal features including institutional features of courtroom discourses, intonational systems, etc. and nonverbal features that embody eye-contact (gaze), bodily conduct, facial expressions, etc. Meanwhile, the human communication occurs in a criminal court and hence the physical settings of a criminal court should also be encoded

that contribute to decorum of the court. For example, the dressing code of judges and layout of the lights, furniture in a criminal court also add some meaning to the dynamic multimodal discourse analysis of criminal trials.

To simplify this framework, at least three layers of annotation of a recorded video in a criminal court are needed:

- a. Context (physical settings of a criminal court)
- b. Verbal mode (institutional linguistic features, intonational systems...)
- c. Nonverbal mode (eye-contact; bodily conduct, facial expressions...)

All the detailed dimensions for annotation in the aforementioned three layers would vary with individual cases. Multimodal case analysis of criminal courtroom discourses is the first step that paves the way to future quantitative multimodal analysis based on the large-scale video and audio data.

On the other hand, scientifically designed computer software might ease the efforts of designing complicated coding schemes for multimodal analysis. For example, Multimodal Analysis Video developed by the Multimodal Analysis Company encompasses SFT concepts and frameworks for understanding how linguistic, visual and audio resources function to attract attention and create meaning. In another word, the library of the software can provide users with a comprehensive repertoire of system choices for annotation of streams of video images.

For example, concerning visual factors in the criminal courtroom, participants' eye contact always reveals their social distance, interaction and other important information. In a filmed session of criminal trial, with cameras in the criminal courtroom, we can imagine that there are also other potential participants, who are the online audience of the filmed video. We all have the experience that viewers will feel more engaged if the participants on the screen are gazing directly at the camera while they might feel less interested or less engaged when they see the visual participants are not looking directly at the camera. So, it becomes interesting to note whether participants in the courtroom are directly looking at the camera or not. In this case, if the judge in the courtroom is addressing the camera directly, it shows that he or she tends to engage the audience of the filmed video as observers or at least onlookers. In another word, gaze or visual address of the camera in multimodal situations functions to establish interpersonal relations between participants on the screen and the viewer. MMAV offers this system for annotation. This dimension is quite new in terms of nonverbal mode, for it offers opportunity for us to analyze to what extent participants inside the courtroom tend to engage people outside the courtroom or to what extent they would like to build connection between themselves and the camera inside the courtroom.

On the other hand, participants' intonational systems in a courtroom are important for meaning building as well. MMAV offers a variety of systems that are associated with intonation systems, including types of key (declarative, subordinate, coordinate, imperative, challenging, confirmative, etc.); information distribution (less than clause; more than clause);

information focus (head, modifier, conjunction, etc.). Therefore, if we need to analyze the intonations used by the judge in a criminal court, we can choose the afore-listed systems for video annotation. After the annotation of the video, we might also find the trend of combination of systems. For example, when a judge is declaring the opening of the court, he or she tends to use a declarative key while the defense counsel might sometimes use challenging keys while cross-examining witnesses.

Finally, spatial and temporal relation in a criminal court is also worth investigation, which is one of the system choices in MMAV. For example, in a Chinese criminal courtroom, spatial relation between the participants (the judge, the defendant, the defense counsel, the public prosecutor) is rather loose while in an American criminal courtroom, the defendant and defense counsel actually sit together. For temporal relation, confrontation between the public prosecutor and defense counsel in a Chinese criminal court strictly abides by the code of criminal procedure and the two parties speak in turns. Synchronic speech in a Chinese criminal court is rare. In an American criminal courtroom, as defense counsels have the right to cross-examine witnesses from the prosecutor party, there would be synchronic speech from both parties especially when the cross-examinee is irritated by the cross-examiner. Therefore, spatial and temporal relation in a criminal court is also an indispensable part of meaning building of the whole multimodal discourse.

So with the help of the multimodal analysis software, systems chosen for analysis can help construct a framework for each multimodal case analysis. Meanwhile, it is also important to align the theoretical foundation and software in analysis. If the software design is not compatible with the theoretical foundation for the multimodal discourse analysis, the software use will not simplify analysis, but impose more workload.

As a summary, a framework for multimodal discourse analysis of criminal court discourses is constructed upon integration of social semiotics and technical software in line with SFT. The design of the software (MMAV) is based upon social semiotics theory and hence the system choices in the software library are aligned with the theoretical foundation of multimodal discourse analysis. The advantage of combining both in a tentative framework is that it becomes possible for researchers to explore how meaning arises from combination of different choices of systems (modes). In addition, the results of analysis can also be associated to legal meanings in a larger context of criminal trials.

V. CONCLUSION

Multimodal analysis is complex when it comes to analysis of language, image, audio and video resources. Nevertheless, in the modern digitalized world, literacy is no longer confined to language only and actually, every day we are bombarded with images, videos and audio resources that play a major role in communication, which is the same with communication in the courtroom. We come to notice that court discourses are increasingly being mediated online and those cases previously reserved for legal professionals in the judicial circle are now available in the public domain. In fact, these video-taped trials

are accessible by everyone and they might serve as the first-hand research materials for forensic linguists. A multimodal approach is needed to analyze the complexity of meaning building practices in courtroom discourses. Besides, a criminal court involves dramatic conflicts and strategic maneuvering of discourse power, which needs our understanding on how multimodal communication that occurs in a criminal court is able to engage and create meanings. Communication in the courtroom is complex, and if we need to understand courtroom interactions in all its complexity, all modes of communication should be put into consideration. Furthermore, in the digital age that is presenting new changes and challenges, forensic linguists need to meet those challenges to adequately analyze the increasing forensic multimodal discourses in the public. New approaches and new software are needed and this article is an attempt to draw more attention from scholars in relevant fields. Last but not the least, with great complexity in multimodal courtroom discourses, we need to study the discourses as multimodal communication and the dimensions for analysis offered in the tentative framework of the study may shed some light upon researches on multimodal analysis of criminal courtroom discourses.

REFERENCES

- [1] Y. Gu, "Multimodal text analysis: A corpus linguistic approach to situated discourse," *Text and Talk*, Vol. 26 (2), pp 127-167, 2006.
- [2] G. Kress, What is mode?, in C. Jewitt, ed., *The Routledge Handbook of Multimodal Analysis*, London: Routledge, pp. 54-67, 2009.
- [3] M.A.K. Halliday, *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold, 1978.
- [4] G. Kress and T. van Leeuwen, *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Oxford UK: Oxford University Press, 2001, pp. 1-2.
- [5] C. Jewitt, J. Bezemer and K. O'Halloran, *Introducing Multimodality*. London, [England] New York, New York: Routledge, 2016.
- [6] K. O'Halloran, *Multimodal analysis video*. creativework. Multimodal Analysis Company, 2013.
- [7] D. Feng, D. Zhang and K. O'Halloran, "Advances and frontiers of multimodal discourse analysis", *Contemporary Linguistics*, Vol. 16 (1), pp. 88-99, 2014.
- [8] Y. Gu, "Multimodal text analysis: A corpus linguistic approach to situated discourse," *Text and Talk*, Vol. 26 (2), pp 127-167, 2006.
- [9] L.L. Levenson, "Courtroom demeanor: The theater of the courtroom", *Minnesota Law Review*, Vol. 92, pp. 573-633, 2008.
- [10] W. Guo, "On the construction of legal facts in court discourse: A multimodal perspective", *Journal of Guangdong University of Foreign Studies*, pp. 53-57, 2014 (1).
- [11] G. M. Matoesian, "You might win the battle but lose the war", *Journal of English Linguistics*, Vol. 36 (3), pp. 195-219, 2008.
- [12] S. Tan, B. A. Smith and K. J. O'Halloran, "Online leadership discourse in higher education: A digital multimodal discourse perspective", *Discourse & Communication*, Vol. 9 (5), pp. 559-584, 2015.