晋 中 学 院

本科毕业论文（设计）

题　　目<u>基于 **ELAN** 的跨文化交际多模态话</u>

<u>语分析—以"移民与认同"主题为例</u>

院　　系<u>　　　外语系　　　　　</u>

专　　业<u>　　　英语　　　　　　</u>

姓　　名<u>　　　高宇凤　　　　　</u>

学　　号<u>　　1702111205　　　</u>

学习年限<u>2017 年 9 月至 2021 年 6 月</u>

指导教师<u>　　要文静　副教授　　</u>

申请学位<u>　　文学学士学位　　　</u>

2021 年 5 月 6 日

Acknowledgments

I am greatly indebted to all people who support and encourage me during the production of this paper. I would like to take this opportunity to express my genuine appreciation to them, without whom the completion of the paper would have been impossible.

First and foremost, I would show my deep gratitude to my supervisor, Miss Yao Wenjing, for her spiring guidance and insightful instruction. When I encounter any difficulty in research, she always arranges a meeting or offer timely and valuable suggestions despite the busy schedule. Without her patience and revision, this work could not have been brought to its present form.

My sincere thanks also go to other teachers who instruct me and cultivate me knowledge during my studies in Jinzhong University. Their professional lectures, academic enthusiasm and scholarly achievements have great impact on my study and life.

Finally, I would extend my gratitude to my family, my roommates and friends for their understanding, constant assistance and encouragement on me to complete this paper.

The Multimodal Discourse Analysis of Intercultural Communication Based on ELAN
---Take the Theme of "Immigration and Identity" as Example
Name: Gao Yufeng   Supervisor: Yao Wenjing

Abstract: Multimodal discourse is a language phenomenon which auditory, visual, tactile or other senses are used to communicate with each other by combining various means and semiotic resources, such as language, images, actions, etc. According to the theories of multimodal discourse, the paper attempts to combine quantitative and qualitative research methodology and take the theme video of "Immigration and Identity" as example in the Intercultural Communication Course for Translation Majors of 2018 in Jinzhong University. It explores the features and relationship between visual and auditory modalities in communication by ELAN based on the teaching reform project of Jinzhong University in 2019. The major findings are as follows: Students utilize visual modality such as text and image, as well as auditory modality such as pause and speed. Different modalities have different features and distribution. The relationship between visual and auditory modalities is complementary and cooperative. The paper provides a new perspective for the cultivation of students' cross-cultural communicative competence. It includes selecting appropriate modalities or modal combinations to enhance the multimodal awareness and makes use of the relationship among modalities to cultivate intercultural communication competence.

Keywords: Multimodal discourse analysis; intercultural communication; visual modality; auditory modality

基于 ELAN 的跨文化交际多模态话语分析
——以"移民与认同"主题为例

学生姓名：高宇凤　指导教师：要文静

摘　要：　多模态话语是运用听觉、视觉、触觉等多种感觉，通过语言、图像、动作等多种手段和符号资源进行交际的现象。本文基于多模态话语分析理论，采用定量和定性相结合的方法，基于晋中学院 2019 年教改项目以晋中学院 2018 级翻译专业学生跨文化交际课程"移民与认同"主题视频为例，借助 ELAN 多模态话语分析软件，研究听觉和视觉两种主要模态的话语特征以及两种模态间的关系。通过研究发现：学生在跨文化交际中使用了视觉模态如幻灯片中的文本、图像等以及听觉模态如停顿、语速等；不同的模态具有不同的模态特征和分布；听觉模态和视觉模态之间的关系主要为互补和协同关系。本研究为学生跨文化交际能力培养提供新视角，学生根据不同的交际阶段选择合适的模态或者模态组合提高模态使用意识，学会利用模态间的关系来培养跨文化交际能力。

关键词：多模态话语分析；　跨文化交际；　听觉模态；　视觉模态

Contents

Ⅰ **Introduction**

With the collision and communication of multiculturalism in the era of big data, the trade between countries is becoming increasingly frequent and cross-cultural communication has become an indispensable way for countries to conduct trade. At the same time, the way of information transmission has gradually expanded from a single language to multiple modal forms such as language, action, and image. The meaning of discourse serves as information transmission in the communication, which is conveyed through not only a single print-based text or spoken language mode, but also through multiple modes.

Multimodal Discourse Analysis (henceforth MDA) has begun to attract domestic and foreign research attention. Multimodal discourse refers to the phenomenon of applying various senses such as hearing, sighting and touching to communicate through various means and symbolic resources such as language, image, sound and action. It is a kind of discourse analysis perspective which treats all the communicative modes as meaning-making resources (Jewitt, 2006). Classroom teaching under the guidance of multimodal discourse analysis theory has become a contemporary development trend, which has emerged and highly been when applied in English classrooms. Nowadays, educational technology and electronic products are flooding into classrooms. The new English curriculum standard also points out that the discourse in language teaching is often presented in the form of multimodality, including in speaking and writing, audio and video, etc (Ministry of Education of the People's Republic of China, 2018). Cross-cultural communication refers to the communication between people with different cultural backgrounds. The discourse carries language knowledge and cultural background, and also conveys cultural connotations, thinking modes and value orientations. It has received attention and research in recent years, but there are relatively few multimodal discourse analyses on cross-cultural communication.

The paper, combined quantitative and qualitative research methodology, takes the theme video of "Immigration and Identity" in the Intercultural Communication Course for Translation Majors of 2018 in Jinzhong University as an example, and explores the features and relationship between visual and auditory modalities in

communication by Elan software. That is to say, it aims to analyze what modes are used to facilitate communication and how students make full use of different modes to learn the connotation meaning on the basis of the theory of multimodal discourse analysis.

Ⅱ **Literature Review**

This part aims to provide a theoretical and analytical basis of multimodal analysis from different perspectives, which mainly include the perspective of social approach and the perspective of linguistic approach. In the context, the research is grounded on Halliday's Systemic Functional Linguistics, Kress and van Leeuwen's Visual Grammar and framework for MDA of Zhang Delu. Besides, there are substantial researches of multimodality on teaching and learning abroad and home, in both theorical and practical application.

According to social semioticians, symbol meaning exists strongly and universally in any sign system, in various visual, auditory, behavioral and other codes, which are made up of society and must be dealt with as social practice. Symbol resources are not only the product of cultural history, but also the cognitive resources that are used to create meaning when they are produced and interpreted images and other information. With the advent of the image age, especially the mass production of images in media, education and children literature, multi-modal discourse plays an indispensable role.

Halliday (1978) puts forward the most suitable theorical basis for multimodal discourse analysis which is known as the theory of Systemic Functional Linguistics. Halliday views language as a form of donging rather than a form of knowing,

i.e., social demand on language has helped to shape its structure. Language interacts to create meanings along with three basic perspectives of meaning or meta-functions: the ideational, the interpersonal and the textual, which is known as Meta-functional Principle proposed by Halliday. Systemic Functional Linguistics puts complex relations among various modes into the same theoretical framework for researches, which can be used to study various disciplines related to language, such as language teaching, sociolinguistic, and discourse analysis.

Following Halliday's SFL, Kress and van Leeuwen (1996, 2006) introduce visual grammar and develop a descriptive framework of multimodality as a tool for visual analysis. In their book Reading Images (Kress and van Leeuwen, 1996, 2006), Visual Grammar interprets language as a socially semiotic system. Therefore, based on the theory of SFL, three similar meta-functions of an image are proposed, namely, representational, interactional and compositional functions to describe semantic relationships between elements. Kress and van Leeuwen's (1996, 2006) work lays a solid foundation for expanding theory of SFL across a range of modes. By extending three meta-function from linguistics to Visual Grammar, they turn the concept of grammar into a more flexible one, which can be used to analyze verbal and non-verbal modalities in the English classroom of junior middle schools.

Li (2003) drew into the multimodal discourse framework which was based on systemic functional linguistic and held accountable for it. What's more, the significance of multimodal discourse analysis was also discussed by Li. In English teaching, the social semiotic analysis of multimodal discourse has positive significance. Teaching with pictures should fully understand the difference and connection between image as a social symbol and language as a social symbol, as well as the difference in resources used between image and language.

In the rapid development of discourse analysis theory, people have found that it is no longer possible to conduct in-depth and comprehensive analysis and research on discourse from the perspective of language alone, because a large part of the meaning of discourse is embodied by non-verbal factors. Communication does not make use of only one sense, but two or more senses are used at the same time. For example, using PPT for class is visual and auditory at the same time; simulation and explanation are auditory. The discourse produced in this way of communication is multimodal discourse. Since 1990s, no matter at abroad or at home, various scholars have conducted researches on multimodal teaching and learning when multimodality gradually becoming one of the hot topics in the world of linguistics.

The New London Group (1996) first proposes the term "multiliteracies" and applies multimodality in language teaching. In the digital world and in the background of new technology, they propose four teaching methods: overt

instruction, critical framing, transformed practice and situated practice in order to meet the needs of communication. Walsh (2010) emphasizes that teachers need to encourage students in effective multimodal literacy learning with complex modes of images, sound, music, film and graphics. He discusses how to use the multimodal literacy in classroom practice. Christine (2016) explores multiliteracies and multimodal discourses in the foreign language classroom, and states that the ambition of the multiliteracies framework is to go beyond verbal communication and multimodality as a mean to be encompassed other types of semiotic resources. These studies have discussed multimodality from different aspects, such as language teaching and second language acquisition, which has further instructed and expanded multimodal theory.

Compared with the foreign researches, domestic studies of multimodality do not start until the early 2000s and are still at the infant stage (Zhang, 2009a). Zhang (2009) proposes a theoretical framework for Multimodal Discourse Analysis and explores a series of researches on different modalities and their inter-semiotic relations. According to Zhang, multimodal discourse system can be classified into four levels: context level, content level, culture level and expressive level. Context of culture plays a crucial and critical role in the context of situation. Context of situation serves as the decisive factor in the choices of discourse meaning, which includes three variable (tenor of discourse, mode of discourse and field of discourse) and decides three kind of meanings: ideational, interpersonal and textual meanings. He also emphasizes the significance of the inter-semiotic relations in Multimodal Discourse Analysis, which are named complementary and non-complementary relationships on the level of form. Zhang (2017) analyzes the necessity and feasibility of applying multi-modal teaching practice to cross-cultural communication courses. Zhang (2018) uses multimodal video discourse analysis to study the cultivation of college students' cross-cultural sensitivity.

From the above all, scholars at home and abroad have conducted theoretical research on multimodal discourse, but relatively few studies have been conducted in the field of cross-cultural communication under the perspective of multimodal discourse analysis. There is a lot of room to make further improvement in theoretical

and practical research. Therefore, the paper, based on the theory of multimodal discourse analysis, attempt to analyze the characteristics of multimodal discourse and use the relationship between the modalities to achieve meaning construction, in order to improve students' cross-cultural communication competence.

Ⅲ Research Methodology

This paper aims to explore the features of visual modality and auditory modality, analyzes their relationship between two modalities and how they work together to cultivate student's intercultural communication competence in the English classroom. Hence, for the research methodology, this part consists of research purpose, research subject, research methods and research procedure.

3.1 Research purpose

In order to attempt to explore the intercultural communication from the perspective of multimodal discourse analysis, this paper takes efforts to delve into two main modalities (visual and auditory modes) used by five translation major students in the intercultural communication course with the theme video of "Immigration and Identity" as example.

Immigration has become a global phenomenon. In the process of cross-cultural communication, immigrants and local residents will inevitably produce cultural collision and integration, and therefore cultural identity will appear. According to American psychoanalyst Eriksson (1968), cultural identity refers to the affirmative recognition of the most meaningful things of the nation formed by members of a group living together for a long time in the national community. It is a value affirmation of the spiritual existence of people, and it is mainly through the characteristics of the nation itself. Customs and lifestyles have been passed down to this day in the form of collective unconsciousness, integrating people's various identities, thereby preventing possible cultural conflicts between different identities.

Collier and Thomas believes that intercultural communication refers to the communication between people who have differences in language and cultural background. The identity among people is the key logo of cross-cultural communication, and intercultural communication is defined as the process in which

people of different cultures identify and advocate their respective cultures in cross-cultural encounters. The success or failure of cross-cultural communication depends on many factors. Only by deepening the understanding of micro-cultural knowledge of different cultures, language habits, social culture, customs, etc., can cross-cultural communication be effectively and smoothly realized. Students make a video of cross-cultural communication on immigration and identity, which is the material of the research.

In the research on the theoretical basis presented in Part Two, the paper aims to explore what features exist in the modalities and how students apply visual and auditory modalities to take intercultural communication. With the help of ELAN, the research combines quantitative and qualitative analysis and investigates the relationship between visual and auditory modalities to discuss the cultivation of students' intercultural communication competence.

3.2 Research subject

The research of this paper is on the basis of the construction and application research of the project-based flipped classroom teaching resource library based on VR. The subject of this research is the theme video of translation major students of 2018 in Jinzhong University about "Immigration and Identity" in the Intercultural Communication Course. Five students are chosen as the subjects of the research. The theme is excerpted from *Mrs. Spring Fragrant and Other Writings* written by Sui Sin Far that is the pen name of a Chinese-American writer Edith Maude Eaton. In this first-person autobiographical work, the author recalls the discrimination she received as a child and later as an adult, and expresses the confusion about her identity as a member of an ethnic minority, which is thought-provoking. Students reflect on the theme of immigration and identity and make a video about it.

For convenience, five chosen students in the videos are respectively referred to as Student A, Student B, Student C, Student D and Student E. One male student and four female students are included in five. The characteristics underlying the selection of them are: they all speak fluently English; their video contents are relatively complete; the time of all videos are roughly four to five minutes; and they can smoothly express their own ideas about the theme. In addition, the reasons why

select them as participants are as follows: First, they are all translation majors. They focus on cultivating cultural awareness, cultural comparison awareness and cross-cultural communication competence in English language learning and English-Chinese translation. Secondly, students indirectly perform their competence though making the theme video that mainly make use of visual and auditory modality, which make it possible to take the multimodal discourse analysis of intercultural communication.

3.3 Research methods

The research employs quantitative and qualitative methods to make the research more reliable. All of videos are analyzed through the annotation of ELAN (EDUICO Linguistic Annotator).

ELAN is a multimodal annotation tool developed by the Max Planck Institute that allows to create, annotate, visualize and search annotations for video and audio data. ELAN can be designed for the analysis of language, sign language, media corpora. The paper uses the advanced version is ELAN 5.8, which can be downloaded from ELAN official website https://tla.mpi.nl/tools/tla-tools/elan/ . There are three main steps involved in the annotation process: define tier types and tiers, select time interval and enter annotations. "Type" and "tier" are used to interpret at the first place that are two key terms of ELAN annotation. A tier refers to a set of annotations which share the same characteristics. Types are what tiers are assigned to. For example, one tier includes pure language while one type includes pure language and paralanguage which belong to verbal modes.

Hence, what needs to do first is to categorize the modes adopted by students before annotating students' theme videos with ELAN. Through observing videos carefully and based on the multimodal media system presented above, the paper is mainly designed to discuss visual modes and auditory modes. For visual modes, three mode parameters are taken into consideration, including text, image and overstriking words. For auditory modes, speed, pause and expression style are contained.

For the sake of convenience, modes need to be coded and are named after their abbreviations. The coding scheme is designed in Table 3.1.

Table 3.1 Coding scheme

Code	Meaning
1. [I]	Text
2. [I]	Image
[T &I]	Text & Image
3. [OW]	Overstriking words
4. [S]	Speed
[S1]	Slow speed
[S2]	Quick speed
5. [P]	Pause
6. [ES]	Expression Style
[ESN]	Expression-Style Narration
[ESM]	Expression-Style Monologue

3.4 Research procedure

For research procedure, there are three steps to obtain data of multimodality with the aid of ELAN to take a discussion. The main procedures are as follow:

The first step is to define the type(Figure 3.1). There are mainly two types visual modality and auditory modality. Click "Add New Tier Type" option and then add the name of "Visual modality" and "Auditory modality" into the blank.
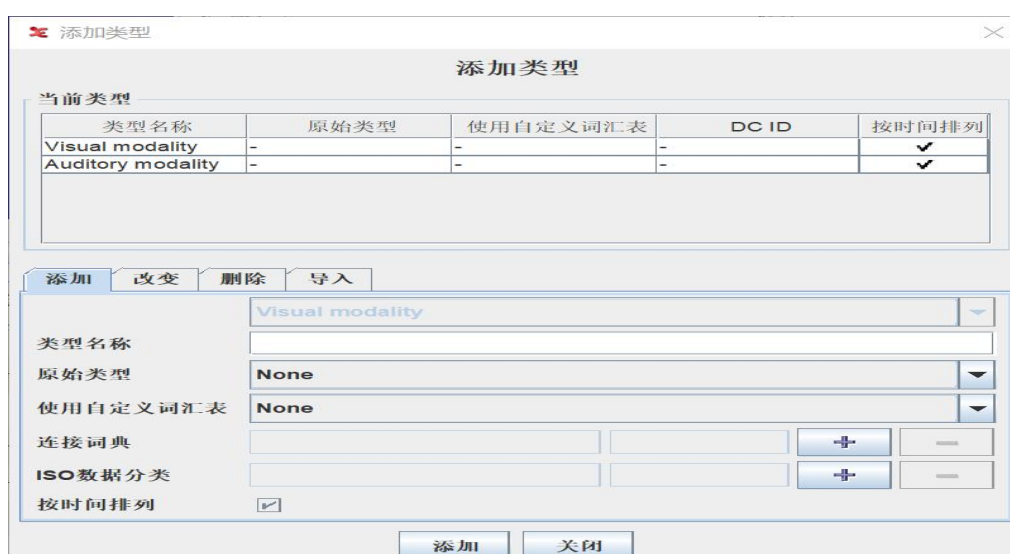

Figure 3.1 Types of ELAN

The second step is to define the tier (Figure 3.2). There are six tires under different types which are defined as text, image, overstriking words, speed, pause and expression style. For example, when defining the tier of "Speed", click "Add New Tier" and choose "Auditory modality" in the "tier type" so as to code "Speed" in the "Tier Name" menu.
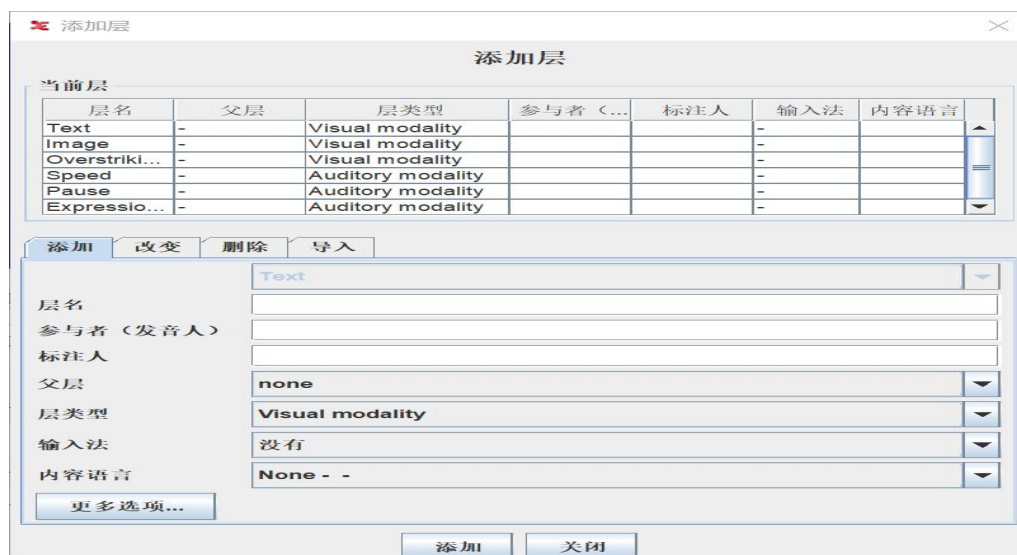
Figure 3.2 Tier of ELAN

The third step is to select time intervals and enter annotations (Figure 3.3). The time interval means the section of a complete element of mode. One could click the "Selection mode", play the video and pause it when an interval is finished. Then right click on the chosen time interval of the observed tier, select "New Annotation Here" and put the corresponding mode code which indicates a finish of annotation.

Figure 3.2 Example of Annotation

Ⅳ The Results and Analysis of the Research

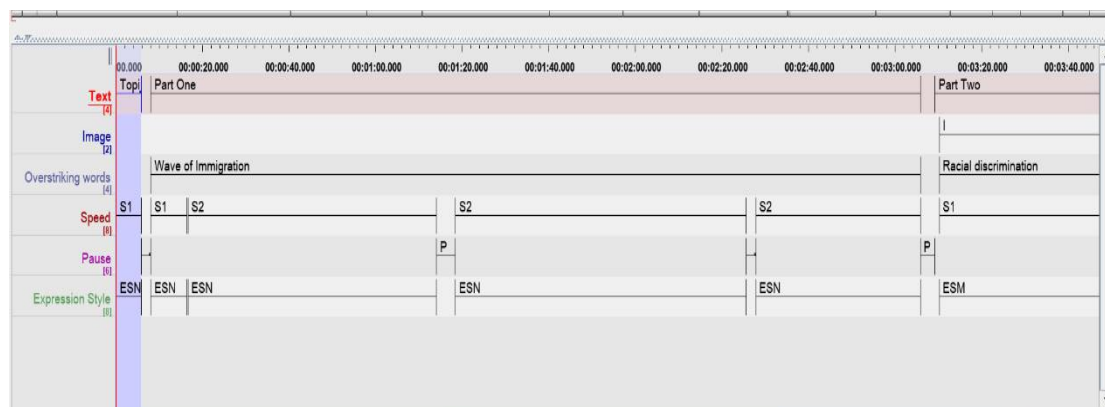According to the researcher methodology proposed in the Part Three and based on the theoretical basis presented in Part Two, the paper annotates different modes for the theme videos of translation major students, analyses the data of the research with the aid of ELAN and makes a discussion through the observation results. Therefore, in this part, the research aims to explore translation major students apply what visual and auditory modes to facilitate intercultural communication, and discuss the relationship between visual and auditory modalities, including complementary relationship and cooperative relationship.

4.1 Modality features

After video segmentation, transcription, and annotation are completed by ELAN, the visual and auditory modes of the five theme videos are analyzed quantitatively and qualitatively. The detailed analysis of visual modes will be presented in the following part.

4.1.1 *Features of visual modality*

According to Kress and van Leeuwen (2006), visual design, in order to achieve communicational requirements, can function as a full system of communication like all semiotic modes. Therefore, based on the Visual Grammar, the research sets up three modality parameters to explore the feature of visual modality, which includes text, image and overstriking words. The duration of visual modality will be shown as the Figure 4.1.
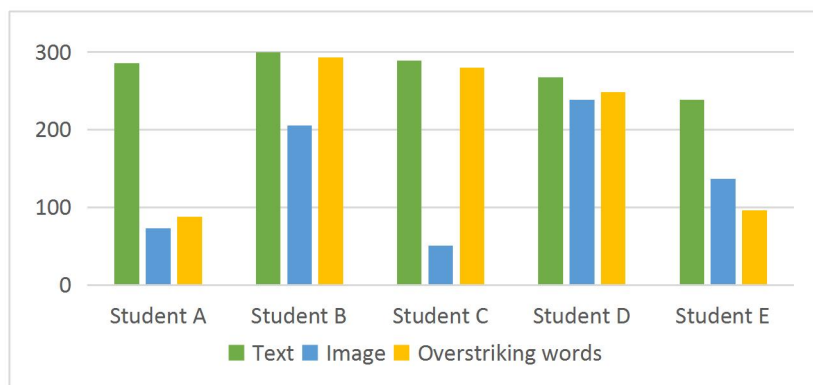

Figure 4.1 The duration of visual modality

From the statistical results of Figure 4.1, five students all adopt three mode parameters, among which text occupy the most and overstriking words is second. Power Point is

the dominant presentation tool to facilitate intercultural communication. The exquisite design and appropriate arrangement of different semiotic resources of Power Point also contribute to the success of five students to take conversation. Images help to suggest or explain the text. Overstriking words actually reminds the thesis of the speaker at different stages, helping the viewer to better understand and grasp the key points. Students use the words of text to complete the main task of information transmission, while pictures and overstriking words are assisting with the completion of the intercultural communication.

The first mode parameter is text that is applied by all five students to demonstrate topic, definition, reason and reflection about the theme immigration and identity. Although the topic is the same, the text of the PPT has a clear difference in the making the whole content. The text, which is divided into three parts, not just include some information from the books and internet, but also covers students' own reflections after the students' careful consideration and selection. In order to achieve successful communication better, they use different manners, such as color, size and bold. The pictures below are two extracts of text from Student A and Student C.

Figure 4.2 The example of Student A

Figure 4.3 The example of Student C

Figure 4.2 show the definition of immigration. Figure 4.3 shows three waves of the immigration. They are the students' generalization through collecting information from books and the Internet. With the assist of the text on the PPT, the purpose of understanding the development and meaning of immigration can be achieved. In these pictures, students present the main idea of the theme immigration. At the same

time, important sentences, key words and phrases are bold in different colors to highlight students' verbal language. Thus, audience can receive the input more clearly and have a better understanding of the content through the attractive PPT.

Table 4.1 The duration of image and overstriking words

	Tiers	Number of Annotation (times)	Average Duration (seconds)	Annotation Duration Percentage (%)
Student A	Image	2	36.752	25
	Overstriking words	6	14.707	30
Student B	Image	6	29.251	66
	Overstriking words	7	45.240	94
Student C	Image	2	25.289	17
	Overstriking words	4	69.879	92
Student D	Image	3	79.594	86
	Overstriking words	4	62.120	89
Student E	Image	3	45.652	55
	Overstriking words	5	19.315	38

Although text accounts for a large percentage, overstriking words and image play considerable roles with unique features in the intercultural communication. As we can know from Table 4.1, the duration of image and overstriking words account for a significant proportion. The annotation duration of image is up to 25% at least and 89% at most. The annotation duration of overstriking words takes up from 30% to 94%. These two semiotic modes cannot be underestimated, which help text modes to express meaning.

Overstriking words refers to the type of words which is thicker and larger than other words on the same page of PPT labeling in different colors. Overstriking words not only helps the views to better understand and grasp the key points, but also remind the thesis of the speaker at different stages. Also, with the function of making

abstract concepts into intuitive and concrete ones, images can provide direct stimulus, capture viewer' attention and promote memorizing. Take Figure 4.4 and 4.5 as examples.

Figure 4.4 The example of Student B

Figure 4.5 The example of Student C

For the Figure 4.4, Student B shares her own opinions with us about what we should do faced with the problems of immigration. After displaying the advantages and disadvantages of immigration, Student B introduce the topic of culture identity and these vivid images helps to figure out six aspects about the identity. Figure 4.5 is selected from Student C. When Student C talks about racial discrimination caused by immigration, overstriking words "racial discrimination" performs as generalization while images, being enlarged with spoken words, sever as supplement.

4.1.2 *Features of auditory modality*

After observing five theme videos, it finds that these students all have a good command of English, appropriate vocabulary and articulate pronunciation in speech. Therefore, the research has the analysis of auditory modes, including speed, pause and expression style.

In the intercultural communication, language is an important channel of information and plays a dominant role in the conversation. Communication always takes place in a certain context and abstract meaning of language is realized in the real situation of communication. Context consists of tone of discourse and expression styles. Each component of context and other modalities are interrelated and restricted, which jointly construct the discourse meaning of theme video.

In terms of the tone of discourse, it includes phonological features such as speed, pause, intonation and stress. The variety of phonological features can effectively highlight the key points of intercultural communication and enhance the attention of effective information. Therefore, pause and speed as two main kinds of phonological features will be illustrated and discussed in the following part.

Table 4.2 The annotation of auditory modality

	Tiers	Number of Annotation (times)	Minimal Duration (seconds)	Maximal Duration (seconds)	Average Duration (seconds)
Student A	Speed	10	3.700	80.800	27.551
	Pause	6	0.426	5.173	3.017
	Expression Style	7	6.400	80.800	38.423
Student B	Speed	6	26.900	64.077	48.474
	Pause	6	1.916	2.738	2.297
	Expression Style	7	16.646	64.038	41.528
Student C	Speed	8	5.963	67.246	34.693
	Pause	6	2.161	4.599	3.146
	Expression Style	8	5.963	67.246	34.693
Student D	Speed	6	5.389	133.676	44.606
	Pause	4	0.246	1.362	0.829
	Expression Style	6	5.389	133.676	44.606
Student E	Speed	7	5.201	72.257	27.981
	Pause	5	1.059	4.242	2.077
	Expression Style	7	5.201	72.257	27.981

The above table shows the annotation of auditory modality used by five students. The characteristics of pause in the theme videos can be quantitatively analyzed by annotating pauses of the videos. The detailed annotation about pause is shown in the Table 4.2. Pause refers to a period of time during which someone stops talking in communication. Pause has the function of adjusting breath, showing tone, and highlighting key points in oral expression. For example, pauses occur twice in the Figure 4.6. In order to explain vividly national identity, Student E plays a short video about the strong cohesion of China and produce an emotional pause which give the audiences time to reflect and feel the spirit of national identity. Next, Student E shares the ideas of the video with us to arise emotional resonance. Then,

Student E move on "The differences between China and Western countries" by means of a logical pause, which makes the purpose of language clear and the meaning fully expressed. Proper pause, with clear sentence structure and distinct meaning, can enhance the sense of rhythm of the language, enhance the clarity and expression of the language, and also attract the attention and thinking of the audience.
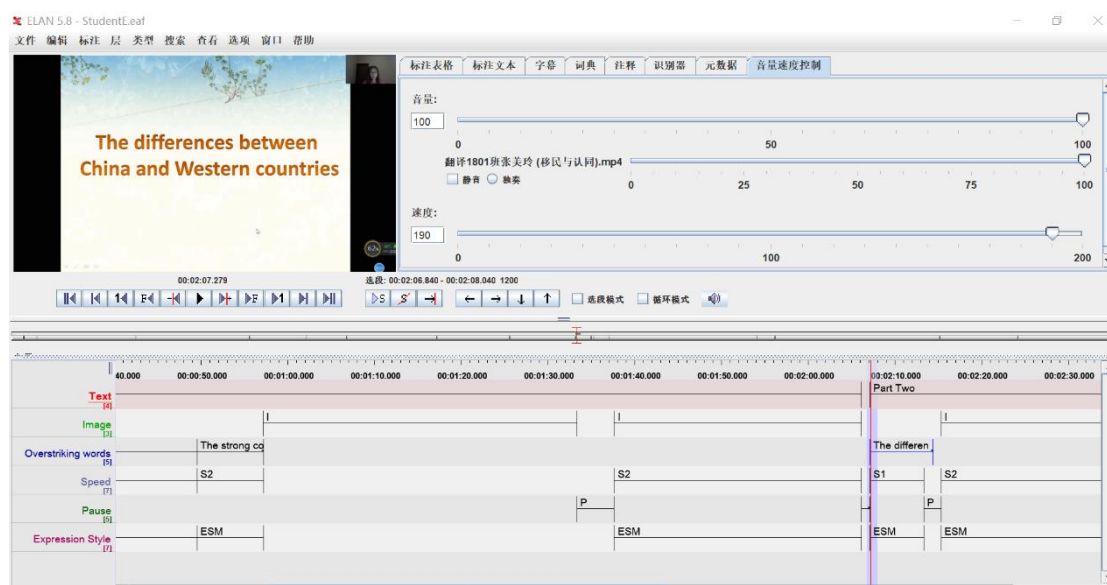

Figure 4.6 The pause of Student E

Speed is one kind of paralanguage modes which can be defined as the vocabulary rate of words or human language symbols expressing meaning in a unit of time. According to the observation with the help of ELAN, the duration of speed are quantified and shown in Table 4.3. At the level of speech speed, the speed ratio of topic videos is close to 1:2. In other words, the five videos adopt an organic combination of fast and slow speech rate interaction, with slow speed as the mainstay, conveying dynamically cross-cultural information and video themes.

Table 4.3 The distribution of speed

	Student A	Student B	Student C	Student D	Student E	Average
Slow speed	61.61%	62.61%	38.09%	78.70%	46.70%	57.54%
Quick speed	31.72%	30.52%	54.03%	17.98%	32.15%	33.28%

In terms of expression style, it can be classified into dialogue, monologue and narration. Dialogue refers to the speech act of two or more people. monologue refers to the inner speech of the video participant; narration, which is voice-over, is narrated by a third party other than the video participant. For the theme videos, dialogue happens between students and audience in the virtual scene where intercultural communication takes place. Thus, the paper mainly analyzes and discusses the distribution of monologue and narration. Table 4.4 makes a comparison and presents the distribution of expression style used by five students.

Table 4.4 The distribution of expression style

	Student A	Student B	Student C	Student D	Student E	Average
Narration	41.56%	79.46%	85.16%	78.70%	63.48%	69.67%
Monologue	49.55%	13.63%	6.96%	17.98%	15.37%	20.70%

As shown in Table 4.4, undoubtedly all students apply two types of expression style in the intercultural communication and it is obvious that narration is the most frequently used by students. In virtue of narration, students display the definition and development of immigration; the reasons and problems about immigration; and the recognition of identity. As a voice-over, the narration can clearly narrate and introduce the content of the video and retain the integrity of the image without destroying the visual sense of the video. In virtue of narration, students express their own reflections for the phenomenon of immigration and the effect of identity. Students use monologues to bring in strong emotions and realize the interactive meaning of words.

4.2 Modality Relationship

After the above discussion and analysis, this part aims to discover the relationship between visual and auditory modes, and find how they cooperate to realize meaning. There are two relationships among modes: complementary and cooperative relationship.

4.2.1 *Complementary relationship*

Complementary relationship, a term named by Zhang Delu（2009）, means that

when one mode cannot exactly express the meaning alone, other modes are needed for assistance and association so as to help to construct the intentional meaning of communication. That is to say, the modes sometimes complement each other and also the modes sometimes strengthen each other.

In the different stages of communication, for visual modality, text mode is the main mode while image and overstriking words are subordinate to it. At the same time, auditory modality that is not usually designed by students can strengthen visual modality. For example, in Student D's video, the text of PPT show the content of communication. Overstriking words "HOW" is marked in red and bold which summarizes the topic of the slide and underline the key point. Image shows some foreigners dancing with dragons that is a unique celebration to China, which embodies indirectly the result of cultural identity. All of them are used to strengthen the text mode. Student D use slow speed and monologue to express her opinions. There are short pauses in the elaboration of each suggestion. These auditory modes can also strengthen the visual modes.

Figure 4.7 The example of Student D

Different cultural features, collision and comparison will make individuals have a reaction to maintain their identity in the face of different elements and cultures. Identity is formed through interaction with others. In the process of intercultural communication, there will be the need for cultural identity, which will inevitably face the problem of cultural identity of mother tongue and cultural identity of target language. In the videos, the speakers make use of the target language as the main modality to communicate as well as consciously and unconsciously convey the

cultural value of the nation of his or her own. During this process, the visual modality such as image, overstriking words and text modes are complementary to each other while the auditory modality functions as strengthening. Representational meaning can be realized through all of these modes so that hearers could master the main ideas. Also, to achieve the interactive meaning, the auditory modality such as speed, pause and text mode are complementary to each other while the visual modality severs as reinforcement.

4.2.2 *Cooperative relationship*

Cooperative relationship means that different modalities jointly express the overall meaning of communication, that is to say, each mode is indispensable and no one mode make contribution to another one.

In the whole process of communication, the cooperation and combination between visual and auditory modalities make communication specific and convenient. When different modalities are used to express the same meaning, the modalities can form a mutually cooperative relationship in which different modalities have different roles. For instance, from the beginning to the end of the video, sound and images are played at the same time. Both images and sounds are necessary, otherwise it is difficult to complete cross-cultural communication. As shown in Figure 4.5, when student A presents the figure about the reasons of immigration, she uses narration and quick speed to make an explanation.
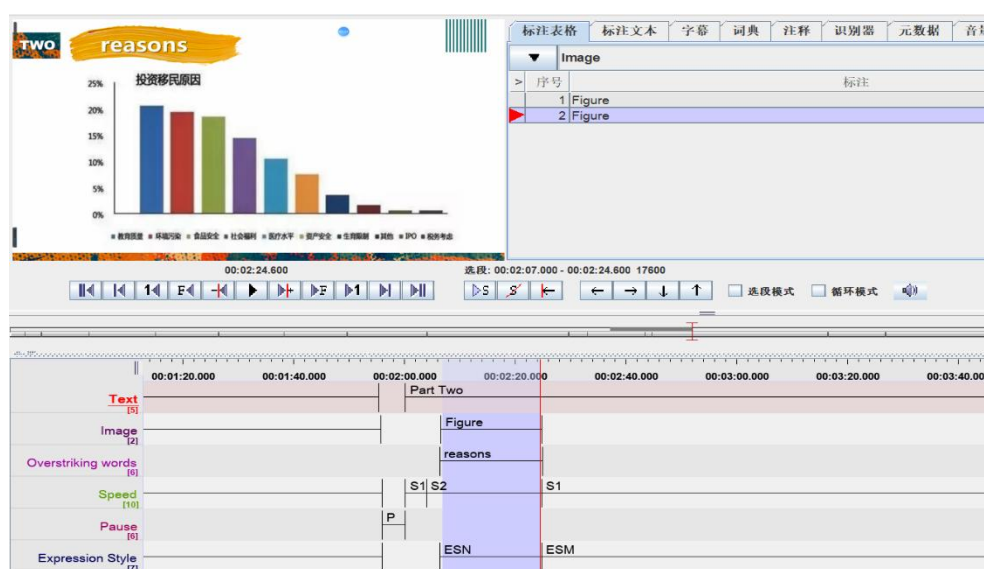

Figure 4.5 The example of Student A

Language learning needs to be compatible with the multicultural interaction and coordinated development. In the process of foreign language learning, learners realize the cultural identity of the target language on the basis of the identification of the native language culture and values, which is a complex, dynamic and multi-dimensional process. Similarly, in the process of intercultural communication with the help of multimodalities, students can fully create opportunities for two or more modes of expression interaction and identification. In other words, students can express meaning by giving full play to the cooperative relationship between auditory modes and visual modes, and then complete intercultural communication. It has to be noted that speakers unconsciously use some unnecessary modes which waste time and energy to some extent. It is important to find whether there are some useless modes or not during the communication.

V Conclusion

In order to make multimodal discourse analysis for intercultural communication, the paper selects five students' theme videos as research subject that focuses on the theme of immigration and identity and employs ELAN as research tool. With qualitative approach and quantitative approach, the paper explores the feature and relationship of visual and auditory modalities employed by five students on the basis of theoretical foundations mentioned multimodal discourse analysis theories. The major conclusions are summarized and presented as follows:

Students, consciously or unconsciously, are fond of utilizing visual and auditory modalities that involve various modes to facilitate and achieve the purpose of intercultural communication. Visual modality can be composed of text, image and overstriking words. With regard to auditory modality, speed, pause and expression style are applied by five students. Some of them have subsection, for example, speed includes slow speed and quick speed, and expression styles are composed of narration and monologue. Students adopt various modes to explain and support the topic of intercultural communication.

Multimodal discourse analysis can be efficiently applied among students in intercultural communication. It can be found that the features of visual and auditory

modalities are observed through analyzing the distribution of different modes and every mode plays a significant role in the communication. In general, text mode as one main visual mode is used to support the spoken language to achieve meaning. Image and overstriking words are used to assist text mode to realize the aims of communication. When it comes to the feature of auditory modality, pause occurs to emphasize key points, change the topic and attract hearer's attention. Also, students are fond of slow speed and narration to explain abstract meaning such as definition, reasons and problems. Students take advantage of monologue as summary.

In the process of communication, the relationships between visual modality and auditory modality are is complementary and cooperative. Students combine visual modality and auditory modality to finish the whole communication. Different modes have different roles to express meaning so that complementary and cooperative relationship occur in the modality or among modalities. For example, the combination of text and image in which image severs as assistance to help text explain abstract information; the cooperation of text and pause in which pause happens when changing the situation of communication.

In the process of intercultural communication, students choose appropriate modalities or modal combinations according to different communication stages, and make rational use of the relationship between modalities, which not only helps to realize the meaning of discourse but also cultivate students' intercultural communicative competence. However, it has to be acknowledged that there are some limitations about the research at present. The research scope is not big enough and extensive. It is hard for me to collect data and the research may be affected by manual annotation in the annotation process. The effect and result of intercultural communication cannot be investigated from the perspective of hearers. Therefore, in further studies, more samples should be included and more intelligent annotation software should be designed. In order to promote the development of multimodal analysis theory, more research fields like psychology should be involved into multimodal research.

Bibliography

Christine, S. 2016. Multiliteracies and multimodal discourse in the foreign language classroom. *Multiliteracies in world education*. New York: Routledge.

Gu, Y, 2007. On multimedia Learning and Multimodal learning. *CAFLE*, (2):3-13 (顾曰国，多媒体、多模态学习剖析，《外语电化教学》2007 年第 2 期。)

Halliday, M. A. K. 1978. *Language as social semiotic: the social interpretation of language and meaning*. Baltimore: University Park Press.

Halliday, M.A.K & Hasan, R. 1985. *Language, context and text: aspects of language in a social-semiotic perspective*. Geelong: Deakin University Press.

Halliday, M.A.K. & Matthiessen, C.M.I.M. 1999. *Construing experience through meaning: A language-based approach to cognition*. London: Cassell.

Halliday, M. A. K. 2004. *An Introduction to Functional Grammar* (Third Edition). London: Edward Arnold.

Hu, Z. 2007. Multimodality in social semiotics. *Language Teaching and Research*, (1)1-10. (胡壮麟，社会符号学研究中的多模态化，《语言教学与研究》2007 年第 1 期。)

Hu, Z., Y. Zhu & Z. Li. 2008. *Introduction to Systemic Functional Linguistics*. Beijing: Peking University Press. 胡壮麟，朱永生，李战子，《系统功能语言学概论》，北京大学出版社，2008 年)

Jewitt, C. 2006. *Technology, literacy and learning: a multimodal approach*. London: Routledge.

Kress, G. 2010. *Multimodality: a social semiotic approach to contemporary communication*. London: Routledge.

Kress, G. & van Leeuwen. 2001. Multimodal Discourse: The Modes and Media of Contemporary Communication. New York: Oxford University Press.

Kress, G. & T. van Leeuwen. 1996/2006. *Reading Images: The grammar of visual design*. London: Arnold.

Li, Z. 2003. Social semiotic approach to multimodal discourse. *Foreign Language Research*, (5):1-8 (李战子，多模式话语的社会符号学分析，《外语研究》2003 年第 5 期。)

Li, Z. & Lu, D. 2012. Multimodal Semiotics: Theoretical basis, Research approaches and development prospects. *Foreign Language Research*, (02):1-8. (李战子，陆丹云，多模态符号学：理论基础，研究途径与发展前景，《外语研究》 2012 年第 2 期。)

O'Halloran, K. 2004. *Multimodal Discourse Analysis*: *Systemic-Functional Perspectives*. London: Continuum.

Royce, T. 2002. Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, (2): 191-205.

van Leeuwen, T. 2005. *Introducing social semiotics*. London: Routledge.

van Leeuwen, T. & Jewitt, C. 2001. *Handbook of visual analysis*. London: Sage.

Walsh, M. 2010. Multimodal literacy: what does it mean for classroom practice? *Australian Journal of Language and Literacy,* (3), 211-239.

Zhang, D. 2009. On a synthetic theoretical framework on multimodal discourse analysis. *Foreign Languages in China*, (1): 24-30. (张德禄，多模态话语分析综合理论框架探索，《中国外语》2009 年第 1 期。)

Zhang, D. 2015. *Multimodal Discourse Analysis and Foreign Language Teaching* (In Chinese). Beijing: Higher Education Press. (张德禄，《多模态话语分析理论外语教学》，北京高等教育出版社。)

Zhang, R. 2017. The teaching practice of intercultural Communication course under the framework of multi-modal discourse. *English Teaching*, (9):13-16(张睿思，多模态话语框架下的跨文化交际课程教学实践，《英语教师》2017 年第 9 期。)

Zhu, Y. 2007. Theory and methodology of multimodal discourse analysis. *Foreign Language Research*, (5):82-86 (朱永生，多模态话语分析的理论基础与研究方法，《外语学刊》2007 年第 5 期。)

Appendix

Figure 1 The annotation statistics of Student A

Figure 2 The statistics variables of Student A's speed

Figure 3 The statistics variables of Student A's expression style

Figure 4 The annotation statistics of Student B

统计参数

标注	命中	频率	平均时长	时间比例	恢复时间
S1	4	0.01303334908...	48.8805	0.63707661980...	4.3
S2	2	0.00651667454...	47.66	0.31058470862...	33.7

Figure 5 The statistics variables of Student B's speed

统计参数

标注	命中	频率	平均时长	时间比例	恢复时间
ESM	2	0.00651667454...	43.897	0.28606246232...	133.6
ESN	5	0.01629168635...	40.580799999...	0.66112966553...	4.3

Figure 6 The statistics variables of Student B's expression style

标注统计　　　　　　　　　　　　　　　　　　　　　　　　— □ ✕

统计

标注 ｜ 标注 II ｜ 层 ｜ 层类型 ｜ 参与者（发音人）｜ 标注人 ｜ 语言

统计参数

层	标注数量	最小时长	最大时长	平均时长	中间时长	总标注时长	标注...	恢复...
Text	4	5.963	177.851	72.158752	52.41	288.635	95.806	0.0
Image	2	6.022	44.556	25.289	25.289	50.578	16.788	190.249
Overstriking words	4	20.978	177.851	69.878754	40.343	279.515	92.779	8.137
Speed	8	5.963	67.246	34.692749	37.11	277.542	92.124	0.0
Pause	6	2.161	4.599	3.145833	2.758	18.875	6.265	5.963
Expression Style	8	5.963	67.246	34.692749	37.11	277.542	92.124	0.0

保存　　关闭

Figure 7 The annotation statistics of Student C

统计参数

标注	命中	频率	平均时长	时间比例	恢复时间
S1	5	0.01681146945...	22.9518	0.38585348468...	0.0
S2	3	0.01008688167...	54.260999999...	0.54732428652...	16.603

Figure 8 The statistics variables of Student C's speed

统计参数

标注	命中	频率	平均时长	时间比例	恢复时间
ESM	2	0.00672458778...	32.1385	0.21611816445...	190.249
ESN	6	0.02017376334...	35.544166666...	0.71705960674...	0.0

Figure 3 The statistics variables of Student C's expression style

Figure 10 The annotation statistics of Student D

Figure 11 The statistics variables of Student D's speed

Figure 12 The statistics variables of Student D's expression style

Figure 13 The annotation statistics of Student E

统计参数					
标注	命中	频率	平均时长	时间比例	恢复时间
S1	5	0.02061379640...	23.201	0.47826069031...	0.0
S2	2	0.00824551856...	39.9305	0.32924767888...	49.597

Figure 14 The statistics variables of Student E's speed

统计参数					
标注	命中	频率	平均时长	时间比例	恢复时间
ESM	2	0.00824551856...	19.088	0.15739045828...	0.0
ESN	5	0.02061379640...	31.538	0.65011791091...	6.306

Figure 15 The statistics variables of Student E's expression style