
晋 中 学 院

本科毕业论文（设计）

题 目 基于 ELAN 的跨文化交际
多模态话语分析——以
“定式与偏见”主题为例

院 系 外语系

专 业 英语

姓 名 陈渝

学 号 1702111202

学习年限 2017 年 9 月至 2021 年 7 月

指导教师 要文静 副教授

申请学位 文学学士学位

2021 年 5 月 6 日

Acknowledgement

I would like to take this opportunity to express my profoundest gratitude to those individuals who have helped and supported me throughout the writing process of this thesis.

First and foremost, I would like to express my sincere gratitude to my Professor Yao Wenjing. I feel so lucky to be one of her students. She has been offering me guidance and suggestions since the beginning of writing. Without her unreserved inspirations and encouragement, this paper cannot have been accomplished.

Finally, I am grateful to my families and friends. My parents, supporting me both materially and spiritually, have done their best to create a carefree environment for me to concentrate on my paper writing. During the process of writing, my friends have also helped me by comforting and encouraging me when I run into bottleneck. Without their unconditional support, I cannot accomplish the writing smoothly.

An Analysis of Multimodal Discourse in Cross-cultural Communication Based on ELAN

-- Take the Theme of “Stereotypes and Biases” as Example

Name: Chen Yu Supervisor: Yao Wenjing

Abstract: With the development of modern information technology, the transmission of information has gradually evolved into a process of participation of various modes. Multi-modal discourse analysis has attracted wide attention of scholars at home and abroad, there are few research on dynamic text of multi-model discourse analysis text focus on the static text. Based on the teaching reform project of Jinzhong College in 2019, based on the videos of foreign language translation class students as the research objects, using the multi-modal discourse analysis theory analyze the video of the course, it takes the system functional grammar, visual grammar and auditory grammar as the theoretical basis, with the method of quantitative and qualitative combination, and analyses the verbal behavior and non-verbal behavior of the videos by software ELAN. The research analyzes the verbal behavior such as “pause”, “oral fluency”, “pronunciation” and so on. The nonverbal behaviors include “eye movement”, “head movement” and so on, and analyzes the relationship among the modes. The study finds that majority of students speak Chinglish. In the videos of students, the phenomenon of paralanguage is quite obvious. It finds that multi-modal application is beneficial to the cultivation of students’ subjective initiative to have a significant reference for foreign language teaching.

Keywords: Multi-modal Discourse Analysis; ELAN; Student speech video

基于 ELAN 的跨文化多模态话语分析

——以“定式与偏见”主题为例

学生姓名：陈渝 指导教师：要文静

摘 要：随着现代信息技术的发展，信息的传递逐渐演变为多种模态共同参与的过程。多模态话语分析引起了国内外学者的广泛关注,但多模态话语分析多集中于静态文本方面，对于动态文本的研究则相对较少。基于晋中学院 2019 级教改项目，以外语系翻译班级学生课程视频为研究语料，运用多模态话语篇分析理论分析课程的视频，以系统功能语法视觉语法和听觉语法为理论基础，采取定量和定性相结合的方法，通过视频分析软件 ELAN 对视频的语言行为和非语言行为进行分析，分析的语言行为有“停顿”、“口语流利程度”、“发音”等；分析的非语言行为有“眼睛的移动”、“头的动向”等，进而各模态之间的关系。研究发现：大多数学生的口语偏中式化。学生多模态视频中，非言语交际中的副语言现象明显。研究发现多模态应用有利于学生主观能动性的培养，对外语教学有一定借鉴意义。

关键词：多模态话语分析；ELAN；学生演讲视频

Content

Acknowledgement	i
Abstract in English	ii
Abstract in Chinese	iii
I Introduction	1
II Literature Review	2
III A Multimodal Discourse Analysis of the Theme “Stereotypes and Biases”	4
3.1 Research Purpose	4
3.2 Research Subject	5
3.3 Research Method	6
3.4 Research Procedure	7
IV The Results and Analysis of the Research	9
4.1 Discourse Analysis of Theme “Stereotypes and Biases”	10
4.1.1 <i>Verbal Analysis</i>	10
4.1.2 <i>Nonverbal Analysis</i>	14
4.2 Discourse Features of Theme “Stereotypes and Biases”	16
4.2.1 <i>Verbal Features</i>	17
4.2.2 <i>Nonverbal Features</i>	18
4.3 The Influence of “Stereotype and Biases” Subject Learning on Intercultural Consciousness ..	19
V Conclusion	20
Bibliography	23

I Introduction

In a multi-modal society, the construction of meaning depends more and more on the integration of various symbolic resources. The multi-modal nature of discourse has become an important feature of various texts used in communication in the world. With the development of discourse analysis and the deepening of discourse research in recent years, in addition to the traditional study of the main elements ---languages in discourse, a group of scholars began to focus on all communicative symbols, including language. Discourse analysis opens a new window---multi-modal discourse analysis has become an important research field of discourse analysis due to the attention and research on the various modes of discourse. With the development and wide application of modern science and technology, especially computer, multimedia and media technology, the multi-modal discourse features are becoming more and more prominent in human communication activities, and the multi-modal discourse forms such as image, sound, text, color, space, action and so on are increasing day by day. Although oral language continues to be the main mode of communication, mass media and written language in public communication will increasingly tend to image patterns and visual culture. Reading and writing ability is not only the ability to understand language and characters, but also the understanding of multi-modal discourse such as image, sound, color, space, action and so on has been put on the agenda. "Meaning is constructed through different modes and is embodied and expressed through the co-occurrence of a communication process ." (Kress and van Leeuwen , 2001) .

The study takes 5 students' speech videos as examples from the translation major in 2018 and the theme of videos are "stereotypes and biases". The multimodal discourse analysis theory is used to analyze relevant videos and make data statistics and analysis. The multimodal discourse analysis is carried out by comparing the videos of different students. Modal coordination research is a part of multimodal analysis, mainly analyzing the coordination and collaboration between different modes. "Stereotypes and biases" means that people think in the same way all the time. There are many phenomenon of "Stereotypes and biases" existing in different culture.

For example the northerners are tough, the southerners are quiet; the people is smart in Zhejiang, the people is cunning in Henan, and the people is lazy in Sichuan. But how many of these stereotypes are actually valid? On the basis of this study, one thing can be found whether they are good or bad, and just arrange them reasonably, modify them properly, and listen to other people' opinions, something unexpected will be got.

II Literature Review

In the current upsurge of multimodal discourse analysis, a group of researchers have emerged, and three theoretical methods of systematic and coherent analysis of multimodal discourse have been gradually formed. The first, represented by Gunther Kress and Van Leeuwen, constructs a theoretical framework and determines the category of analysis based on the study of functional linguistics from the perspective of social semiotics, which is usually called “systematic functional semiotics” or “system functional multimodal discourse analysis”. The second is the “multimodal interaction analysis” represented mainly by Sigrid Norris. The third is the multi-modal discourse analysis of corpus linguistics proposed by Gu Riguo, which analyzes and models the living situation of interpersonal interaction under the guidance of social behavior psychology, behavioral ecology and perceptual ecology.

The main theory used in this paper are “System Functional Semiotics” and “Systemic Functional Linguistics”, in particular, multimodal discourse analysis accepts the view that language is a social symbol and a potential for meaning from systemic functional linguistics, that symbolic systems other than language are also the source of meaning, that multimodal discourse itself is systematic, the hypothesis of pure function is accepted, that multimodal discourse is multi-functional, that is, conceptual function, interpersonal function and discourse function. Because multimodal discourse analysis is deeply influenced by systemic functional linguistics and semiotics, it can be called “systematic functional semiotics”. Halliday has made it clear that systemic functional linguistics is a functional language theory, which is essentially functional and semantic, not formal and syntactic: it studies discourse rather than sentences, and focuses on the use of language and idiomatic expression of language, not grammar. Because of this, systemic functional linguistics focuses on the

relationship between language (discourse) and social system. The relationship between discourse and the use environment, including sociocultural environment.

Systemic functional linguistics has many unique characteristics, but its integrity, functionality and applicability are the most prominent and basic characteristics. First, the integrity of the systemic functional linguistics has a strong and the language of the context, significance, vocabulary, grammar, phonology as the research object, by reflect relationship between them, including as the context of language cultural context, and as a communicative context of situational context of events, and on the basis of the development of the language class and register theory. Category theory relates the use of language to communicative purposes, while register theory relates context to meaning. Kress & Van Leeuwen(1996) believe that images are also socially symbolic in reading images—the grammar of visual design, it advocates the analysis of images by systematic functional linguistics.

The research achievements and important theories of multimodal discourse analysis are still mainly from foreign linguistic circles. For example, Lim & Ang(2006) image and speech studies based on multimodal discourse, the framework of multi-modal discourse comprehensive multi-symbol model is established. A theory of multimodal interaction analysis represented by Norris (2004), Human interaction is regarded as a social activity carried out through the intermediary role of various symbolic resources.

Currently, the study of multimodal discourse in foreign countries is in full swing, but at home is still in its infancy. In recent years, multimodal discourse has attracted the attention of domestic scholars, Good results have been achieved in theoretical research and practical application. Hu Zhuanglin (2007) first introduced the concept of multi-modal symbols, and distinguish multi-modal semiotics from multimedia semiotics; Zhu Yongsheng (2007) described the theoretical basis, content, research methods and significance of multimodal discourse analysis.

In the field of empirical research, Zhang Delu (2009) proposed a comprehensive theoretical framework for multimodal discourse analysis, the application of multimodal discourse theory to foreign language teaching is explored.

Besides, the theoretical basis of multimodal discourse analysis in a specific field

and the research and application of multimodal discourse analysis are only introduced in general, and there is no specific and in-depth research method which is basically based on the theoretical understanding and narration abroad. For the shortcomings of the study only put forward the overall evaluation of the basis. Application is systemic functional linguistics, semiotics direction no specific improvement measures. Therefore, multimodal discourse and other research results are less innovative. There is still much room for further study of the theoretical embodiment of multimodal discourse analysis. The theory of multimodal discourse analysis needs innovation and reference, methodological improvements are needed, its perfection and development need the joint efforts of several generations.

III A Multimodal Discourse Analysis of the Theme

“Stereotypes and Biases”

It briefly describes the methodology throughout the research, which involves four sections. Firstly, research purpose and research subject will be presented, then research method, including a special tool of annotation-----ELAN. Lastly, research procedure is discussed.

3.1 Research Purpose

In the development of discourse analysis theory in full swing, people find that it is impossible to analyze and study discourse comprehensively from the perspective of language alone, because a large part of the meaning of discourse is reflected by non-verbal factors. Using multimodal discourse analysis theory to analyze speaking video provides us with a more vivid and fresh perspective, understanding the modal cooperation of students. In addition, it can provide objective reference materials in detailed for foreign language teaching. And this theory has just emerged in China, from the online search results, there are few papers on this subject. Therefore, this study is going on, hoping that the study can find some modes of speech, so that they can perform better, and hope to make some contributions to future research. This is beneficial to the cultivation of students' subjective initiative and has certain reference significance for foreign language teaching. Modal coordination research is a part of

multimodal analysis, mainly analyzing the coordination and collaboration between different modes. The study is meaningful for students to learn intercultural awareness, different countries have different cultures, raising cross-cultural awareness is conducive to the cultural exchange between different countries

3.2 Research Subject

All five students made these videos after reading the text “Don’t Let Stereotype Warp Your Judgement”. The text mainly introduces the definition of stereotypes, and gives examples to illustrate. In another demonstration of the power of stereotypes to affect our vision, a number of Columbia and Barnard students were shown 30 photographs of pretty but unidentified girls, and asked to rate each in terms of ‘general liking,’ ‘intelligence,’ ‘beauty’ and so on. Two months later, the same group was shown the same photographs, this time with fictitious Irish, Italian, Jewish and “American” names attached to the pictures. Right away the ratings changed. Faces which were now seen as representing a national group went down in looks and still farther down in likability, while the ‘American’ girls suddenly looked decidedly prettier and nicer. At the same time, the disadvantages of stereotypes are mentioned. Stereotypes will interfere with our judgement which may even create contradictions.

The study takes 5 students’ speech videos as examples from the translation major in 2018. The theme of videos are “stereotypes and biases”. Stereotypes are a kind of gossip about the world, and a gossip that makes us prejudge people before they ever lay eyes on people. Hence it is not surprising that stereotypes have something to do with the dark world of prejudice. Explore most prejudices (note that the word means prejudgment) and you will find a cruel stereotype at the core of each one. From the perspective of intercultural communication, thinking patterns make people presuppose prejudice against different cultural groups, and on the basis of this prejudice, they decide the way and behavior of communication. Naturally, actions with prejudice are difficult to be just, and conflicts are likely to occur. For example, when it comes to a certain overseas country or a certain province in China, several words will appear in people’s mind, and they think that these adjectives are the cultural characteristics of these countries or provinces. Their videos are representative, and meet the following conditions: First, the video content is complete, everyone’s views will be different,

but they all show the content to be expressed through the video; Second, the video time is appropriate, because the material is different, the expression time is also different, but it is very important to control the time reasonably. Third, the video has both their appearance and material to be modified. These conditions are conducive to observing the modes that appear in the video.

3.3 Research Method

The qualitative analysis in this paper is to analyze the verbal and non-verbal behaviors in student videos. The research analyzes the verbal behavior such as “pause”, “slow speed”, “frequent speed”, “low voice” and “loud voice”. The nonverbal behaviors include “eye movement”, “head movement”, and analyzes the relationship among the modes. The quantitative analysis is the analysis and generalization of students’ verbal and non-verbal behaviors. Through data analysis, it is found that students have similarities and differences in speech. It can improve students’ cross-cultural awareness and promote foreign language learning.

In order to correctly understand the meaning of multimodal discourse, there are two essential points showing: the first is to determine the grammatical relationship between different components, and the other is to clarify the relationship between text and image. ELAN software used in this study is an annotation tool for creating, editing, visualizing and searching logos of video and audio data, which aims to provide sound technology for logos and develop and utilize multimedia clips. Although ELAN specifically provides analysis for language, sign language, posture, everyone can use it to process multimedia data, such as video and audio, for identification, analysis, and filing. Sentences, words, annotations and translations can all be annotated as media, and annotations can be added based on multiple layers, which are called layers, ELAN provide multiple different annotation views, each connected to audio or video resources. Video time can be accurate to 0.1 seconds, continuous loop play.

As following is the way that the software is used: first click on the ELAN software, then click on the file and create in turn, then find the video that needs to be studied, select and click confirm, then click on the layer and category in turn, and add the required layer name and class alias. Finally, just record the required data. The

research method is that first select five translation class students' speech video. Their videos must have their face, and the duration is not less than 3 minutes; then use ELAN software to analyze and record the video of five students; finally, the results are summarized and evaluated.

3.4 Research Procedure

Before the useful data collected, some basic preparatory work need to be done. Define type. The main research is two types, one is “language mode”, the other is “non-language mode”. Click the “Type” menu and then click “Add New layer Type” option. Put “language mode” and “non-language mode” in the blank. As shown in Figure 3.1.

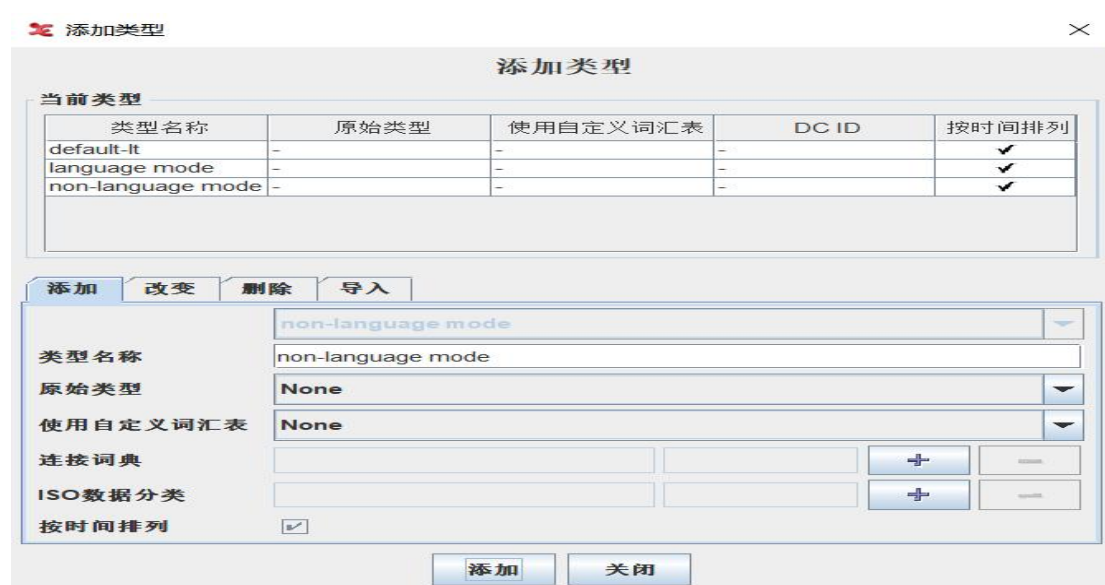


Figure 3.1 Definition of tier.

Define the tier, different patterns are defined in different types. For example, when defining “head movement” mode, click the “hierarchy” menu and click “add a new hierarchy”, and then select non-verbal mode in “layer type”, In the hierarchy name menu, select “head movement” others can be defined in the same way. As shown in Figure 3.2.

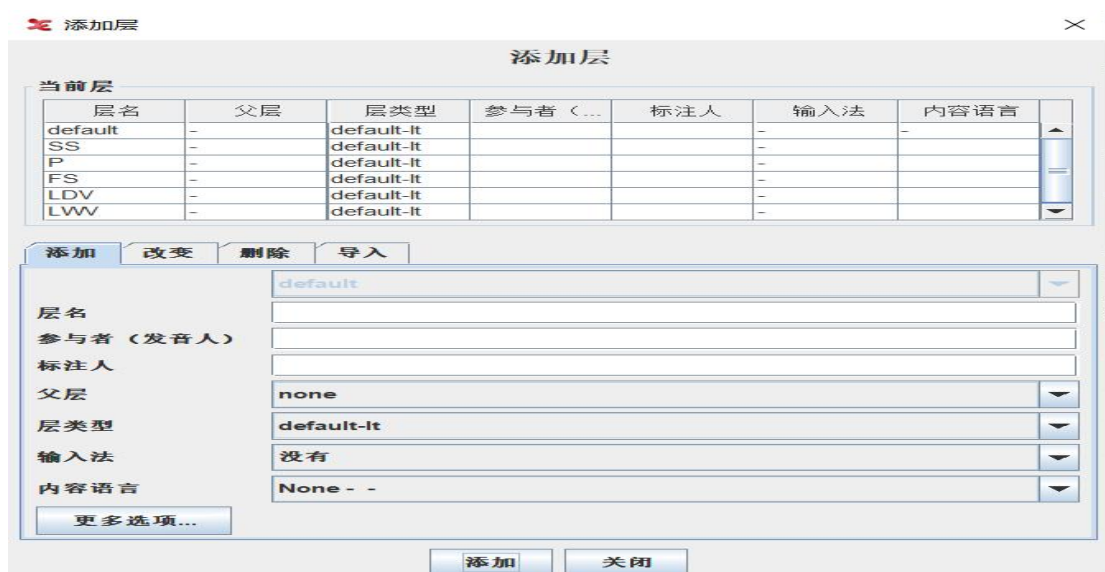


Figure 3.2 Selection of time intervals and entrance of annotation

Select time intervals and enter annotation. The time interval means the section of a complete element of mode. Click “select mode”, play the video, and pause the video when an interval is finished, and then right click on the chosen time interval of the observed tier, left click to mark the location and put the corresponding mode like “SS” which indicates a finish of annotation. As shown in Figure 3.3.

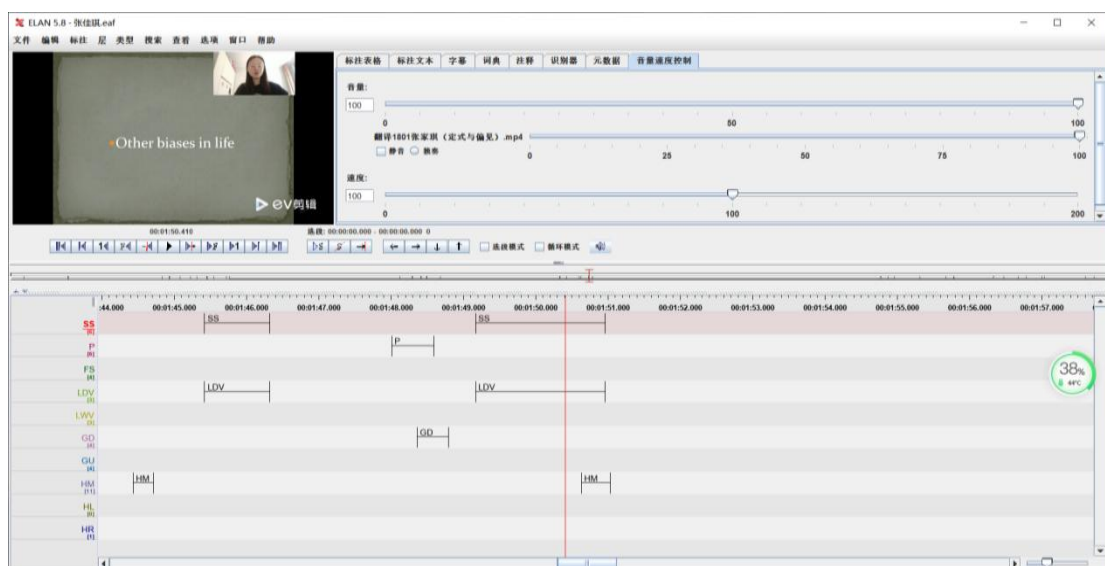


Figure 3.3 Collection of the data and analysis.

Collect the data and analyze them. After annotating all the tiers of modes, the statistical results can be collected through clicking “View” menu and selecting “Annotation Statistics” option. There are Annotation, Annotation II, Tiers, Tier Type, Participant, Annotator and Language. Click Tiers, the detailed information will be shown automatically, including number of annotations, minimal duration, maximal

duration, average duration, median duration, total annotation, annotation duration and latency. Through this, you can get some useful data that can be analyzed to enable students to perform better and more confidently in video. As shown in Figure 3.4.



Figure 3. 4 Abbreviation

The abbreviated letters stand for the following: SS --slow speed, p --pause, FS --frequent speed, LDV --loud voice, LWV --low voice, GD --glance down, GU --glance up, HM --head moving, HR --head right, HL --head left.

Verbal behaviors	Speed of speech	slow speed(SS), frequent speed(FS), pause(p),
	Voice	loud voice(LDV), low voice(LWV)
Nonverbal behaviors	Movement of head	head moving(HM), head right(HR), head left(HL)
	Movement of eyes	glance down((GD), glance up(GU)

Table 1 Frame of multimodality

ELAN is very handy as software for analyzing multimodal corpora and easy to collect data. It also has the advantage that the data can be stored for a long time and can be easily viewed. It is indispensable for the whole process of completing a paper, especially when doing data analysis.

IV The Results and Analysis of the Research

This part illustrates the results of study, including discourse analysis of theme “stereotypes and biases”, discourse hallmark of theme “stereotypes and biases” and the influence of “stereotypes and biases” subject learning on intercultural consciousness.

4.1 Discourse Analysis of Theme “Stereotypes and Biases”

This study is about the multi-modal discourse analysis of the speech video of the students in the translation class. Its theme is “stereotypes and biases”. It uses systematic functional semiotics and systemic functional linguistics to analyze the language behavior and non-verbal behavior of the students. By watching the video, you can see the number of times students use a certain language behavior or non-verbal behavior, they are used to express or why it occurs.

4.1.1 Verbal Analysis

The main analysis is paralinguage, including slow speed, pause, frequent speed, loud voice and low voice. The paralinguage here refers to the language behaviors of students in the video to explain ideas, such as pausing, speaking faster or slower, and so on. First of all, the students’ language behavior was analyzed, the videos of five students’ speeches were analyzed, including their pauses, the speed of their speech and the volume of their voice. The pause is indicated by “P”, the frequent speed of speaking is indicated by “FS”, the slow speed of speaking is indicated by “SS”, the loud voice is indicated by “LDV”, and the low voice is indicated by “LWV”.



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	标注时长...	恢复时间
SS	9	1.23	3.0	2.093333	2.07	18.84	6.489	70.647
P	13	0.29	3.16	0.940769	0.66	12.23	4.212	53.931
FS	1	1.3	1.3	1.3	1.3	1.3	0.448	168.715
LDV	3	0.9	1.37	1.19	1.3	3.57	1.23	127.325
LWV	6	1.51	3.0	2.345	2.485	14.07	4.846	70.647

Figure 4.1 Verbal behaviors of student A

According to Figure 4.1, the verbal behaviors of student A can be found: first, the number of times that the student speaks faster is only 1 time, while the number of times that the student speaks slower is nine times. The first time when the student speaks slower, she says: “There are many stereotypes in our life”, but the only time she said “People’s daily life can be seen everywhere” is because she is familiar with

this sentence. Later, it finds that there were several reasons why she spoke slowly or fast. One reason is because she was not familiar with the content, another one is because she wanted to highlight the following content, and some were because the words were not easy to speak.

Second, student A paused for 13 times, with A maximum length of 3.16 seconds and A minimum length of 0.29 seconds. Student A paused for nearly 3 seconds before the speech. Instead of the video starting to play, Student A began to speak, which left time for the listeners and was more conducive to getting information. Student A paused nine times, the reason is because he mispronounced the word or was unfamiliar with the sentence. The voice of student A gets louder 3 times, and his voice gets smaller 6 times. The main reason why the voice of student A became louder was to emphasize the content she had read. The sentence "People's daily life can be seen everywhere" became louder for the first time, while the main reason why her voice became smaller was that she was not familiar with the words or sentences. Oral English of student A is good, the voice is good, and the presentation effect of the whole video is also very good.



Figure 4.2 Verbal behaviors of student B

According to Figure 4.2, some verbal behaviors of Student B can be found: first, the number of times that Student B speaks faster is 2 times, both of them are due to the linking of the reading content, but the key content of the linking is not the key. This not only shortens the time, but also reflects the linking skills. Student B speaks slowly for 14 times, the maximum length of which is 2.87 seconds, and the minimum length is 0.98 seconds. The content read for the first time is "People's fixed and

simplified idea and impression of a particular thing”, which is to explain clearly the definition of stereotype to the audience and make it easier for them to understand. Secondly, Student B pauses for 12 times, and the maximum duration of pauses is 2.8 seconds. Her first pause is at the beginning of the video, and the following pauses are all caused by her unfamiliarity with the word, which leads to the reading of a word several times, or the pauses are caused by her unfamiliarity with the sentence. Finally, student B’s voice was increased for 4 times and decreased for 9 times. Each time, her voice decreased because she mispronounced the word or was unfamiliar with the sentence, while the reason for her voice increased was to emphasize the content she had read.



Figure 4.3 Verbal behaviors of student C

According to Figure 4.3, some verbal behaviors of Student C can be found: first, Student C speaks slowly for 4 times. The first slow speech is “background, experience and choices”, which is to tell the audience some arguments that will be explained later, and the next 3 slow speech behaviors are to emphasize the content. Student C speaks fast for 3 times, all of which are at the end of an argument to shorten the video time. Second, Student C paused 12 times, with a maximum duration of 3.8 seconds. The first pause was at the beginning of the video, 8 of which were due to word mispronunciation, and the remaining 3 were due to unfamiliarity with the sentence. Third, student C’s voice was raised 4 times, with the maximum duration of 2.05 seconds and the minimum duration of 1 second. The reason for the four times’ voice was raised was to emphasize “background, experience and choices”. Her voice trailed off twice, both times at the end of her argument.



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	标注时长...	恢复时间
SS	6	0.9	2.39	1.653333	1.75	9.92	4.75	10.47
P	6	0.27	5.17	1.458333	0.645	8.75	4.19	0.03
FS	4	0.54	3.31	1.6125	1.3	6.45	3.088	40.27
LDV	3	0.9	2.2	1.63	1.79	4.89	2.341	105.425
LWV	3	0.93	2.39	1.676667	1.71	5.03	2.409	10.47

Figure 4.4 Verbal behaviors of student D

According to Figure 4.4, some verbal behaviors of Student D can be found: first, student D slows down for 6 times. The first slowness is “Chinese and their Kungfu”, which is to highlight foreigners’ prejudice against Chinese Kungfu, and the next 5 slowness is to emphasize the content. Student D speaks fast for 4 times, all at the end of an argument to shorten the video time. Secondly, Student D paused for 6 times, with a maximum length of 5.17 seconds, which happened at the beginning of the video. The pause here was too long, and the remaining 5 times were all the time left by waiting for the next PPT page, which led to a long video. Thirdly, the loudest time of student D is 3 times, with the maximum duration of 2.2 seconds and the minimum duration of 0.9 seconds. The reason for the loudest time of student D is to emphasize that what he is talking is such as “each expectation, each expectation and all the expectations”. Her voice dropped three times, with a maximum length of 2.39 seconds and a minimum length of 0.93 seconds. All three times were due to unfamiliar sentences.



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	标注时长...	恢复时间
SS	11	0.94	3.96	2.069091	2.09	22.76	7.767	38.113
P	14	0.003	4.23	1.653786	1.28	23.153	7.901	0.03
FS	3	0.96	1.72	1.226667	1.0	3.68	1.256	47.98
LDV	3	0.56	1.31	0.963333	1.02	2.89	0.986	95.29
LWV	4	0.72	2.31	1.5375	1.56	6.15	2.099	34.27

Figure 4.5 Verbal behaviors of student E

According to Figure 4.5, some verbal behaviors of student E can be found: first, the number of students E speed slow for 11 times, the maximum length of 3.96 seconds, the minimum time of 0.94 seconds, the first time the slow speed of content: “the two groups were asked to”, described above is the content of the first part, followed by the second part, the key to highlight the contents of the second part; Student E spoke fast for 3 times. She spoke slowly throughout the whole video, and sometimes she spoke fast. Therefore, the reason for the fast speech was to shorten the video time. Second, the number of students E pauses for 14 times, her first stop in the first place of video, including 10 times due to prepare adequately for the lead to pause, the remaining three are in PPT of time, such as student E video watching is more easily, because she is too slow, but because the pause too many times and damaged the continuity of the video. Third, the number of his voice increased for 3 times and the number of his voice decreased for 4 times, all of which occurred before and after the pause. Because there were too many pauses, the size of the voice was not well controlled.

4.1.2 Nonverbal Analysis

GD	7	0.24	0.39	0.317143	0.34	2.22	0.765	58.448
GU	3	0.23	0.39	0.29	0.25	0.87	0.3	141.808
HM	16	0.29	0.53	0.38625	0.37	6.18	2.129	82.133
HL	3	0.33	0.36	0.35	0.36	1.05	0.362	110.163
HR	4	0.27	0.41	0.3275	0.315	1.31	0.451	61.018

Figure 4.6 Nonverbal behaviors of Student A

According to Figure 4.6, some nonverbal behaviors of Student A can be found: First, the number of times that Student A glances down is 7, which is what she would do when she is thinking. However, it is not recommended that the number of times that Student A glances up is 3, which is accidental. Second, the number of times her head shook was 16, related to the number of times she paused. She tilted her head to the right four times and to the left three times, which was accidental.

GD	9	0.23	0.4	0.331111	0.35	2.98	0.84	27.911
GU	-	-	-	-	-	-	-	-
HM	9	0.24	0.56	0.382222	0.37	3.44	0.97	23.975
HL	-	-	-	-	-	-	-	-
HR	-	-	-	-	-	-	-	-

Figure 4.7 Nonverbal behaviors of Student B

According to Figure 4.7, some nonverbal behaviors of student B can be found: first, the number of students to glance down for nine times, first glance down said “of people”, as if the student B a habitual behavior, sometimes glance down two consecutive times, sometimes glance down every ten seconds or so, but the number of her glance up to 0 times, both glance upward and downward glance, too many for video rendering have bad influence. Second, students move their heads 9 times. The first time when students move their heads is when stereotype is defined. The appropriate number of head moves can make the video better. Head of student B turns left or right times is 0. A proper head turns left or right can attract the attention of the audience more.

GD	22	0.28	0.61	0.465909	0.47	10.25	4.337	3.95
GU	-	-	-	-	-	-	-	-
HM	13	0.35	0.82	0.548462	0.55	7.13	3.017	4.69
HR	-	-	-	-	-	-	-	-
HL	1	0.45	0.45	0.45	0.45	0.45	0.19	69.585

Figure 4.8 Nonverbal behaviors of Student C

According to the Figure 4.8, some nonverbal behaviors of student C can be found: first, the number of students C eyes glance down to 22, the maximum length of 0.61 seconds, the shortest time of 0.28 seconds, a habit, it is as if the student C about 20 seconds will down at once, but the number of her glance down too much, give a person a sense of insecurity, and her eyes glance up the number of times of 0. Secondly, she shook her head 13 times. There are reasons: most of the head shaking happened during the pause, so she should have no idea what to do when she paused.

She tilted her head to the left once and to the right 0 times, which should be out of habit.

GD	4	0.32	0.56	0.46	0.48	1.84	0.881	29.906
GU	4	0.26	0.44	0.33	0.31	1.32	0.632	34.011
HM	11	0.28	0.61	0.427273	0.4	4.7	2.251	28.648
HL	-	-	-	-	-	-	-	-
HR	1	0.38	0.38	0.38	0.38	0.38	0.182	193.248

Figure 4.9 Nonverbal behaviors of Student D

According to Figure 4.9, some non-verbal behaviors of Student D can be found: First, student D glances up and down 4 times, which is clearly caused by his habit, but it has little impact on the presentation of the whole video. Second, student D shook his head 11 times, some of the content she read, and then his head shook with her, her head tilted to the left 0 times, head tilted to the right 1 times, which belongs to accidental behavior.

GD	6	0.49	1.46	0.861667	0.795	5.17	1.764	14.84
GU	-	-	-	-	-	-	-	-
HM	7	0.56	0.85	0.721429	0.72	5.05	1.723	28.763
HL	-	-	-	-	-	-	-	-
HR	-	-	-	-	-	-	-	-

Figure 4.10 Nonverbal behaviors of Student E

According to Figure 4.10, some non-verbal behaviors of Student E can be found: First, the number of times that Student E glances down is 6, and the number of times that Student E glances up is 0. After analysis, it is caused by his habit. Second, the number of head shaking of the student is 7, which is also related to the excessive number of pauses of the student. The number of head tilt to the left and the number of head tilt to the right are both 0. It is still good to reduce the number of head shaking.

4.2 Discourse Features of Theme “Stereotypes and Biases”

A multimodal discourse analysis was conducted on the videos of five students and recorded, and the characteristics of their verbal behaviors and non-verbal lines

were summarized. The theme of the students is the same, the theme is “stereotypes and biases”, but the selected materials are different, and the way and content of the speech are also different. The students can try their best to express the content they want to express, but due to various reasons, the students’ videos are different.

4.2.1 Verbal Features

This section focuses on the verbal features of the students. First verbal feature is obvious: each of the students at the beginning of the video will pause for a period of time, some of the most pause time is 3 seconds, the beginning of the length of the pause time also plays an important role, pause time is too long can let a person feel boring, pause time is too short, can let a person feel not to enter the state, the beginning of an appropriate pause time is very important. The number of pauses also needs to be controlled. Excessive or inappropriate pauses will have a negative impact. Student E has the largest number of pauses, which also brings a negative impact on her video. Pause time of student D was not well handled, and the pause time for changing PPT was too long. Secondly, it is found that the students have a good control over the speed of speech. Speeding up the speed can shorten the time of the video, and slowing down the speed can emphasize the key points, so that the audience can understand the key points and deepen their understanding of the video content. It is also found that there is a certain relationship between the speed of students’ speech and the size of their voice. For example, when student A speaks faster, his voice will become smaller, and when he speaks slower, his voice will become louder. There were similarities in the students’ language behaviors, and there were also big differences, but there were more similarities, such as pausing because of mispronouncing words, slowing down or making their voices louder for emphasis. The theme of this text is “stereotypes and biases”, and students use their videos to make their audience understand what “stereotypes and biases” are, or to deepen their impressions of them, and then reach the purpose of their presentations. For example, Student A gave an example of the stereotypes that people have about men and women, that men should earn money and women should stay at home to raise children, and men should not cry and women like to cry. These examples are very vivid and typical, which is enough to illustrate the theme. Student A presents these examples and

explains them to make this video more attractive and persuasive. Although student D also cited an example, it was a video about asking foreigners about China, but it did not combine with the example to give a speech, which was not vivid enough. As a result, the example and the topic are not connected in time. By analyzing the students' para-language behaviors, the number of pauses should be appropriate. Too many pauses will lead to incoherence of the content, and few times may cause slower speech speed. When explaining the example, you must control the size of the voice and the speed of the speech.

4.2.2 Nonverbal Features

This section focuses on the students' nonverbal characteristics. By observing students' speech videos, it is found that students' favorite nonverbal behaviors in speech are about head and eyes. They like to glance up or down, shake their heads, tilt their heads to the left or to the right. By analyzing video, can be found that the students appear most frequently non-verbal behavior is the shake of the head, student A shook his head 16 times, the number of shaking the head too much, the impact on the video is not good, give a person a kind of feeling of shady, proper head shaking can give a person a kind of affinity, feel like the people in the audience and interactive video. In addition, the number of students' head tilting to the right or to the left is relatively small. Student B even did not turn his head to the right or to the right once. Of course, this does not have much effect on the video. Then there is the upward glance and the downward glance of students' eyes. Student C looked down too many times, up to 22 times, giving people a very unconfident feeling. When expressing opinions, there is no interaction with the audience at all, nor can it attract the audience, which is not conducive to the expression of the content. Student D and student E can deal with their movements well, whether it is head movements or eye movements. Some students are accustomed to upward glance, while others are accustomed to downward glance. However, as long as the timing and frequency of watching are controlled, the video can be made better. By analyzing the nonverbal actions of students, each student has its own characteristics, some perform well and some do not perform well. But if you want to show better results, you must control non-verbal actions. Too much head movement will cause the audience to focus only on your head

and ignore what is being said. It is best to have an interaction with the audience in a speech, and eye contact is a good way of interaction. Do not look up or down all the time, but look at the audience and interact with it to achieve better results. When speaking about important content, it is important to make eye contact with the audience to better capture their attention and thus make them absorb what is being said. The theme of this paper is “stereotypes and biases”, and this theme is reflected in the movements used, for example, Chinese people eat with chopsticks, while Westerners eat with a spoon, using completely different movements.

4.3 The Influence of “Stereotype and Biases” Subject Learning on Intercultural Consciousness

The different historical backgrounds of Chinese and Western countries have created different cultures, and different cultures have shaped different personalities. The culture of Confucius is representative of China’s excellent traditional culture, and the idea of commonwealth that it advocates has influenced generations of Chinese people. Confucius’ ideas have permeated the ideology of the Chinese people, creating their conservative, polite and traditional personalities.

Westerners were influenced by Christianity and they sought equality, freedom and fraternity. They were not bound by traditional and feudal ideas and began to learn from other countries about advanced cultures and advanced civilizations to develop themselves. The more often they interacted with other countries, the more progress they were able to make and the more open and liberal Westerners became. In these competitive times, sometimes people should be as conservative as a bun and not express our opinions casually to prevent being hurt by despicable people. At other times people are expected to be as outspoken as a pizza and show others our strengths.

In the process of cultural exchanges between China and the West, there are many cultural conflicts. In order to change this status quo, it is necessary to increase cross-cultural awareness. Improving cross-cultural awareness is conducive to the exchange of cultures, and it is also conducive to the absorption of foreign cultures by students. In the process of learning a foreign language, students understand the foreign language culture in advance, which will promote the learning of the foreign

language.

Through the multimodal discourse analysis of the five students' videos, you can find many things that have been neglected before. When watching a video, people tend to pay attention to its verbal output or non-verbal output, but they do not think of combining the two, analyzing them together, putting forward opinions and making improvements. The theme of the student videos selected in this study is "stereotypes and biases". Although the theme of each student is the same, the video styles and effects presented by them are different. Student A mentioned that foreigners have some prejudice against Chinese people, but they don't know the fact is very different from what they think. Student C mentioned about gender. For example, women should have children and take care of children at home, while men should go out to work to support the whole family. The theme is the same, but the content is representative and easy to understand, which can deepen the understanding of different cultures. Multimodal discourse analysis of video can not only improve the ability of appreciation, but also provide an example for future generations. Stereotypes are one way in which people "define" the world in order to see it. They classify the infinite variety of human beings into a convenient handful of "types" towards whom people learn to act in stereotyped fashion. Worse yet, stereotypes get in the way of our judgment, even when people do observe the world. Hence, quite aside from the injustice which stereotypes do to others, they impoverish ourselves. Understanding the cultures of different countries can not only increase our knowledge, but also promote the cultural exchanges between different countries. Friction and integration between different cultures will surely produce unexpected effects. It is hoped that this research can improve students' intercultural awareness. For example, if a student knows that the word "nigger" is a kind of word that discriminates against black people, then the student must not mention the word in front of black people, otherwise there will be conflict. If the intercultural awareness is improved, it will not only promote the learning ability of students, but also promote the cultural exchanges between different countries.

V Conclusion

The paper is a multi-modal discourse analysis of five students' speech videos. The theme of the video is "Stereotype and Biases". The verbal and nonverbal behaviors of the five students are recorded using ELAN software. Analyze and summarize their behavior, and discover the characteristics of students' language behavior and non-verbal behavior through this. In the video, the students explained the "Stereotype and Biases" of different countries. The cultural differences between different countries can be found in videos, which improved cross-cultural awareness.

This paper conducts a multimodal discourse analysis on the speech videos of students in the translation class. From the language level, it can be found that: students like to leave a little pause time at the beginning of the video, so that the audience can get used to it. When they speak slower, their voice will go faster, when they speak faster, their voice will go lower, and of course there are cases when they are unfamiliar with the sentence, whether they speak faster or slower, their voice will go lower. Out of the five students, three students' spoken English is not very good, even have strong accent; Pronunciation of student A is the most standard, which also the most comfortable to listen to, whether it is the connection of the sentence, or the adjustment of the size of the voice is just right. Student B expressed clearly and highlighted the key points, but the speed of the speech was slow. Student D pauses for too long, which affects the expression of the content.

On a non-verbal level, students tend to move their heads and glance up or down when speaking. The quality of the video can be improved by correcting some verbal behaviors or non-verbal behaviors that have bad effects on the video, or by appropriately combining verbal behaviors with non-verbal behaviors, which will have extraordinary effects. Student C glances down too many times, up to 22 times, which is not conducive to the communication with the audience. Student E had the least small actions, and the overall performance of the video was good. In the speech video on stereotypes and prejudice, each student chose different materials, but they all explained clearly the topic of stereotypes and prejudice, which made people realize the deficiency of conventional thinking, so that they could view an event or a person in a more comprehensive way.

Through the analysis of five students' speech videos, it concludes that different

students use different way to express a theme,so the effect presented by the video also has its own characteristics. At the same time, it finds that the number of verbal behavior and non-verbal behavior used by students in speech is different, which have their own advantages and disadvantages, if it can be adjusted in time, it will produce a better effect.

There are also many shortcomings in this study. In order to obtain more reliable data, it is necessary to select more videos if only five students' speech videos are selected. In the study of students' verbal behavior and non-verbal behavior, the selection of content is limited, and some other topics have not been studied.

Bibliography

- Ang, S. H. & E. A. C. Lim. 2006. The influence of metaphors and product type on brand personality perceptions and attitudes. *Journal of Advertising*, (2): 39-53.
- Chaudron, C. 1977. A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, 27 (1): 29-46.
- 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*. London: Edward Arnold.
- Halliday, M. A. K. 1985. *An Introduction to Functional Grammar*. London: 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。)
- Kress, G & T. Van Leeuwen. 1996. *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress, G. & T. van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Edward Arnold.
- Li, Z. 2003. Social semiotic analysis of multimodal discourses. *Foreign Language Research*, (5):1-8. (李战子, 多模态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。)
- Norris, S. 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
- 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.

- Scollon, R. 1998. *Mediated Discourse Social Interaction*. London: Longman.
- Wang, W. & Y. W. 2008. Multi-modal analysis in applied linguistics research. *Foreign Language Audio-visual Teaching*, (03):8-12. (王文非、文艳, 应用语言学的研究的多模态分析方法, 《外语电化教学》2008 年第 3 期。)
- Wei, Q. 2009. On modal, medium and modality in multimodal discourse. *Foreign Language Teaching*, 30(4):54-57. (韦琴红, 论多模态话语中的模态、媒介与情态, 《外语教学》2009 年第 3 期。)
- Xin, Z. 2008. A new development of discourse analysis: multimodal discourse analysis. *Journal of Social Sciences*, (5):208-211. (辛志英, 话语分析的新发展——多模态话语分析, 《社会科学辑刊》2008 年第 5 期。)
- 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*. Beijing: Higher Education Press. (张德禄, 《多模态话语分析理论与外语教学》。北京:高等教育出版社。)
- Zhu, Y. 2007. The theoretical basis and research methods of multimodal discourse analysis. *Journal of Foreign Language*, (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。)