A Multimodal Discourse Analysis of Tourism Promotional Video *Beijing*

Jiao Yanwei, Hanita Hassan
Universiti Teknologi Malaysia, Malaysia
Email: jiao@graduate.utm.my, m-hanita@utm.my

Abstract
An increasing number of people can now afford domestic and international travel thanks to economic growth, which gives rise to the fact that tourism is now a significant business in many nations. As a result, research on tourism discourse has become a trend for linguistic studies in many countries. Tourism discourse, as a specialized type of discourse, provides information to tourists and persuades potential tourists to choose tourist destinations. With the development of science and technology, the communication between human beings is realized through many modalities instead of a single modality, text, image, music, sound, and color combined to construct the meanings of communication. As a result, people get tourism information through various channels, among which tourism promotional video is the most typical one combining verbal and other semiotic elements. The purpose of this study is to investigate the multimodal features of the tourism promotional video *Beijing* and explore how verbal and visual elements are combined to represent the destination image of Beijing from the perspective of multimodal discourse analysis.
Keywords: Multimodal Discourse Analysis, Tourism Promotional Video, Verbal and Visual Element

Introduction
As a semiotic industry, tourism is committed to the "production, commodification, and representation of culture and cultural difference"(Thurlow and Jaworski, 2011). Meanwhile, tourism is an information-intense industry (Chung and Koo, 2015). It plays a vital role in people's life and serves as an essential part of a nation's economic development and international communication. Since more and more people are longing for experiencing exotic scenery and culture, international tourism has become a trend. For a nation to draw in foreign investment, tourism helps to develop its national image (Salim et al., 2012). Consequently, there is a growing trend of the publication of tourism research articles concerning tourism industries in different countries worldwide (Li et al., 2018).

With the growing role of the internet, tourism has been combined with online communication, which leads to the emerging of some new communication types in tourism (Lowry, 2017). People can explore specific sceneries, local customs, buildings, and local folks in a tourist destination by browsing webpages and watching online videos. As a result, tourism promotion is essential in attracting potential visitors and influencing their decision to choose

the destination they travel to. Additionally, the discourse about tourism promotion has a significant impact on how travelers choose their tourist destinations. (Isti'anah, 2020) Therefore, language is an essential consideration in all areas of tourist promotion., just as Hitchcock and Dann (1998) assert that the language of tourism uses images, pamphlets, and other media to entice millions of people into traveling and then seeks to manipulate their thoughts and behavior. In the era of the internet, language has emerged as a critical medium for disseminating information on tourism through various media, including official travel websites, blogs, and promotional videos, etc. According to Li et al (2020), social media provides great platforms for tourists to get tourism information in either textual or verbal language. Since language serves as a conduit between potential visitors and tourist sites, language and tourism are closely intertwined in tourism promotion.

Because it is utilized for tourism-related reasons, tourism discourse has long been viewed as an ESP. (Irimiea, 2019). Tourism discourse is a specialized language used by tourism staff (Cappelli, 2006) and has its vocabulary and phrases with high frequency (Francesconi, 2014b). Tourism discourse is characterized by unique syntactic and semantic elements that serve various pragmatic tasks (Lam, 2007). Similarly, Cesiri (2011:2) claims that tourism discourse serves as a specialized language for its role as a medium through the discourse in brochures, flyers, promotional videos, and tourism websites and its components like images, texts, and music. As a specialized discourse, tourism discourses are formed according to certain grammatical norms and use specialist vocabularies (Dann, 1996:2). As one type of discourse, the analysis of tourism discourse has emerged as a branch of discourse analysis.

This study will focus on the multimodal discourse analysis of the tourism promotional video (TPV). With various methods of promoting tourist destinations are used, promotional videos are the most effective by appealing to both the visual and auditory senses of the audience (Aoki and Fujimoto, 2021). As reported by Lee, et al (2011), tourism videos are effective in promoting tourist destinations, for they enable potential tourists to visualize what tourist destination looks like through short videos and make tourists fancy about travelling to the tourist destinations. There are generally five destination images in most tourism promotional videos: nature, scenery activities, culture, and people (Gudlaugsson and Magnu'sson, 2010). The beautiful sceneries displayed in promotional videos stimulate people's desire to travel to a particular place to enjoy the sights there (Josiam et al., 2012).

As a significant way of promoting tourist destinations, TPVs function as one type of advertisement to convince potential tourists to choose a particular destination and purchase there. Advertisement videos not only contain more emotional content but also easily represent reality, thus leading to more entertainment for customers (Fong et al., 2017). What's more, since videos process 75% of full information in the brain (Bhosle, Chaudhary and Sharma, 2012), more and more TPVs are produced to attract visitors.

On May 16, 2018, the TPV *Beijing* was released on Youtube, which is 3-mins-10-seconds long. The TPV displays every aspect of the ancient and modern capital of China. As TPVs contain both linguistic and other semiotic modes. The study aims to examine how ideational function and representational meaning are realized in the TPV *Beijing*.

Literature Review
Systemic Functional Linguistics
The nature of language is tightly related with its functions in social context (Halliday, 1970, p. 7), and language can be seen as a social meaning system as well as a meaning potential (Halliday, 1978, p. 123). When one wants to express a meaning, he or she should intentionally

choose the language from the system. According to Halliday (1970; 1985), language has three meta-functions, i.e. ideational function, interpersonal function and textual function. In this study, only ideational function is discussed.

Ideational Function

Ideational function consists of experiential function and logical function. Experiential function refers to language's function of expressing people's experiences in the world, while logical function is the function of expressing the logic between two or more meaning units. In this study, only experiential function is adopted. As experiential function is realized through transitivity and voice, this study focuses on the transitivity analysis of the TPV. According to Halliday (1964), transitivity refers to "a network of systems whose point of origin is the major clause". The transitivity systems are about "the type of process expressed in the clause with animate and inanimate participants in the process, diverse attributes and circumstances of the process" (Halliday, 1967). Transitivity is a meaning system, which categorize human's activities into different processes and indicate the participant and circumstantial elements in the processes. In general, there are six processes in transitivity system, namely, "material process, mental process, relational process, behavioral process, verbal process and existential process" (Halliday, 1994, p. 143).

Visual Grammar

In terms of discourse analysis, it is not enough to analyze discourse just from the perspective of the verbal language, because a part of the meaning in discourse is represented by non-verbal language, such as tone, tonality, gesture, facial expression, circumstances, etc. According to Kress and Van Leeuwen (2006), both language and visual communication can realize "the same more fundamental and far-reaching systems of meaning that constitute our cultures" through their own forms independently (p. 17). Just like the word classes and clause structures in language, the colors and compositional structures in visual communication can also make meanings. Based on Halliday's three metafunctions of language, Kress, G. and Van Leeuwen (2006) concluded that visual images can also make meanings. The meanings of visual images are "representational meaning, interpersonal meaning and compositional meaning, representational meaning corresponding to ideational function, interpersonal meaning corresponding to interpersonal function, and compositional meaning corresponding to textual function". In this study, only representational meaning is analyzed.

Representational Meaning

Just like the ideational function of language in SFL, in representational process, the "participants" in visual images can also depict the experiential world and represent human's mental activities. According to Kress and Leeuwen (2006), in visual grammar, representational meaning is realized by narrative representation and conceptual representation (p. 59).

In narrative processe, "participants are connected by a vector" (Kress and Leeuwen, 2006, p. 59), while there is no vector in conceptual processes. Vectors represent different processes, such as action process, reactional action, speech process and mental process, and conversion process. In this study, only action process and reactional action are discussed. In action process, there are Actor, vector and Goal. The Actor is the participant which produce the vector, while the vector directs to the Goal. When there is only Actor in the process, this action is called non-transactional action process (Kress and Van Leeuwen, 2006, p. 63). When

there are Actor and Goal, the process is transactional action process. What's more, when the action only concludes the Goal in the action, it is called Event. When the vector in the process is formed by eyeline or glance, the process is reactional process in which there are Reactors and Phenomena. Likewise, reactional process can be transactional and non-transactional. "In transactional process, there are both Reactors and Phenomena, while in non-transactional process, there is no Phenomena" (Kress and Van Leeuwen, 2006, p. 64).

For conceptual process, there is no vector linking visual elements together, but "it represents participants in terms of their more generalized and more or less stable and timeless essence in terms of class, or structure or meaning" (Kress and van Leeuwen, 2006: 79), and it can be realized through classificational process, analytical process and symbolic process. In classificational process, "participants are related to each other through taxonomy, one set of participants being subordinate and the other set superordinate". In analytical process, "the relation between participants is part and whole, including Carrier (the whole) and Possessive Attributes (the parts)" (Kress and Van Leeuwen, 2006, p. 88). The symbolic process is concerned with what a participant means or is.

Previous Studies on Multimodal Discourse Analysis
Multimodal discourse analysis is widely used in analyzing different types of discourse. Some scholars conduct multimodal discourse analysis of books. Guijarro and Sanz (2008) examine children's picture books and the intersemiotic relationship between visual and textual meaning and their realizations in a variety of linguistic and visual modes, using Kress and van Leeuwen's social semiotic analysis and Halliday's linguistics as a framework. Chovanec (2019) uses multimodal critical discourse analysis to explain how a succession of photos with textual captions may form a self-contained story based on a common stereotyped script.

What's more, some scholars conduct research on pedagogical discourses. Bower and Hedberg (2010) use multimodal discourse analysis approach to examine the relationships between collaborative, content-related, and technological discourses. Hong (2012) proposes a novel teaching model for college English listening and speaking based on the multimodal discourse analysis and state that the multimodal teaching paradigm outperforms the conventional model in enhancing students' intrinsic desire for English learning and promoting students' practical capacity to communicate in English. Lim (2019) examines teachers' use of gestures in the classroom for meaning construction via systemic functional multimodal discourse analysis, recognizing that teachers' use of a variety of semiotic resources is good for students' classroom experiences. Shin et al (2020) investigate a sixth-grade L2 writer's writing of digital multimodal texts, as well as how he learns about language and images through a systemic functional approach to multimodal discourse analysis and sociosemiotic ethnography. They conclude that multimodal composition pedagogy may enable L2 learners to achieve multimodal communicative competence. Twiner *et al* (2021) use sociocultural discourse analysis and multimodal analysis to examine how instructors and students make meaning, concluding that verbal engagement in combination with other modalities is more effective.

In addition, some scholars conduct multimodal discourse analysis of tourism discourses. Ansori and Taopan (2019) examine the ideational meaning and representational meaning of the tourism promotional video of Indonesia. Arfin *et al* (2014) adopt a multimodal discourse analysis approach to explore multimodal recourses used in tourism websites to portray the

diversity of Malaysia. Francesconi (2014) discusses the meaning of tourism language through analyzing the three metafunctions of the texts and visual images. Houliston, Ivanov and Webster (2021) examine nationalism in Official Tourism Websites of Balkan Countries from the perspective of multimodal analysis and discuss the salient features.

Compared with previous studies, this study focuses on the dynamic tourism discourse about China's tourist destination Beijing by analyzing the ideational function and representational meaning of the TPV *Beijing*. What's more, the implication of the result can be a reference for the multimodal tourism promotion of China's tourist destinations.

Methods

This study adopts a qualitative research method combining theoretical research to analyze the verbal texts and images of the TPV *Beijing* from the perspective of multimodal discourse analysis. Qualitative research investigates and comprehends the meaning that individuals or groups attach to a social or human issue (Creswell, 2014, p. 32). Thus, qualitative research is widely used in tourism studies.

In qualitative research, data collection includes setting the boundaries for the study and collecting information documents and visual materials (Creswell, 2014, p. 239). This study investigates the verbal texts of the TPV based on the ideational metafunction in Systemic Functional Linguistics (Halliday, 1994), and it analyzes the visual images based on the representational meaning in visual grammar (Kress and van Leeuwen, 1996).

Methodology

This study will collect verbal and visual data from the TPV *Beijing*, for Beijing ranks first among the "Top 10 Most Popular Tourist Destinations in China" based on the statistics on the *Top China Travel* website. On May 16, 2018, the TPV *Beijing* was released on Youtube, which is 3-mins-10-seconds long. The TPV displays every aspect of the ancient and modern capital of China. Therefore, the TPV *Beijing* is selected to explore the cooperation of both verbal and visual elements in representing the destination image. The verbal data, i.e., subtitles and captions, will be obtained through Nvivo by transcribing the subtitles of the video. At the same time, the visual data are obtained the frames of the video obtained from the software Premier.

Data Analysis

Qualitative research documents the details of events and identifies the meaning of the conduct for the participants and the viewers. Though qualitative research focuses on qualitas, the range and frequency of actions and meaning perspectives are also essential for good qualitative research (Papakitsou, 2020). Therefore, this study focuses on the qualitative data and considers the frequency of some actions and meaning perspectives. The current study employs NVIVO 12 to analyze data, for NVIVO enables the researcher to manage both the verbal and visual data. NVIVO 12 is applicable for this study because of its functions in organizing and categorizing different types of data into groups by coding and calculating the frequency.

Results and Discussions

Analysis of the Verbal Mode of the TPV Beijing

The main form of verbal language in this TPV is the caption. After transcription, there are seven sentences whose content covers history, culture, landscapes, and climate.

Table 1

The Distribution of Six Processes Within Transitivity

Proportion Process Type	Number of Processes	Percentage
Material Process	1	12.5%
Mental Process	2	25%
Relational Process	5	62.5%
Verbal Process	0	0
Behavioral Process	0	0
Existential Process	0	0
Total	8	100%

Based Table 1, 12.5% of the clauses in this video is material process, 25% mental process, and 62.5% relational process, which indicates that relational process is dominant in this tourism promotional video.

(1) Material Process

Material process refers to the process of doing, in which there are Actor, Process and Goal (Halliday, 1994: 109). The clause can be in either active or passive voice. Among the eight sentences, three are in the type of material process. The sentence below is analyzed as an example.

a. The city of Jiyan *was constructed* three millennia before, five dynasties the capital beheld.

The city of Jiyan	was constructed	three millennia before, five dynasties the…
Goal	Process	

The italicized part in the sentences above show the material process, from which the history of Beijing is displayed. Example a shows the long history of Beijing as the capital of five dynasties in China.

(2) Mental Process

A Mental Process is a process of feeling, thinking, and perceiving and the two participants involved in the mental process are called "Senser" and "Phenomenon" (Halliday, 1994: 113).

a. You may touch the palm print of history, the age of civilization, the gateway of innovation

You	may touch the palm print of history, the age of civilization, the gateway…
Senser	Phenomenon

b. You shall taste prevailing context of literacy, lingering fragrance of history in seasons of sun and snow.

You	shall taste prevailing context of literacy, lingering fragrance of history in…
Senser	Phenomenon

The examples above are about visitors' perceptions of Beijing's culture from various perspectives. Example b states the civilization and innovation of Beijing, which indicates Beijing is not only ancient but also innovative. Example c conveys the information that the atmosphere of Beijing is surrounded by the fragrance of literature and delicacies.

(3) Relational Process

The Relational process defines the relation between two participants involved in a clause. There are two main types of relational process, namely attributive and identifying relation processes. The ones that appear in this video belong to identifying relation process, which involves "identified" and "Identifier". According to Halliday (2004), identifying clause consists of "Token" and "Value".

d. The good nature of people is inborn.

The good nature of people	is	inborn
Identified	Process	Identifier
Token		Value

e. It's the City of Future.

It	is	the City of Future
Identified	Process	Identifier
Token		Value

f. It's the City of Dreams.

It	is	the City of Dreams
Identified	Process	Identifier
Token		Value

g. It's the City of resplendence.

It	is	the City of resplendence.
Identified	Process	Identifier
Token		Value

h. It's the City of eternity.

It	is	the City of eternity.
Identified	Process	Identifier
Token		Value

Example d shows the education of people, which indicates the education people received. Example e demonstrates that Beijing is a promising city. Example f shows that Beijing is a place for people to realize their dreams and make tourists here enjoy their dream. Example g

suggests that Beijing is a city with glorious culture and history. Example h forecasts that the history and culture of Beijing will be eternal. All the sentences describe a beautiful and glamorous Beijing, which makes potential tourists dream of going.

Analysis of the Visual Mode of the TPV *Beijing*
The place of visual communication is referred to as the "semiotic landscape" by Kress and Leeuwen (1996: p. 35). The semiotic landscape is influenced by social, cultural, and economic factors, which leads to the linguistic and cultural diversity of visual communication. For TPV Beijing, 171 frames were selected from the video by Premiere for the analysis of representational meaning of the visual data, which adopts both qualitative and quantitative approaches in data analysis.

Any semiotic mode, as well as visual mode, can represent the objects in the world. Corresponding to the ideational metafunction of language proposed by Halliday, the representational meaning of visual images has a similar function of representing the objective and inner world of human beings and the relations of objects. The representational meaning of visual modes can be constructed through the narrative and conceptual processes.

Table 2
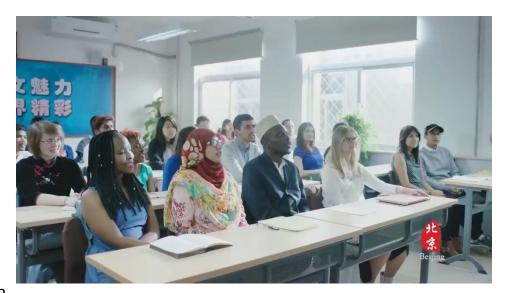*The Distribution of Narrative Process and Conceptual Process*

Representational Process		Number	Percentage
Narrative Process	Action process	50	29.2%
	Reactional process	15	8.8%
	Circumstance	25	14.6%
Sub-total		90	52.6%
Conceptual Process	Classification process	7	4.1%
	Analytical process	29	16.9%
	Symbolic process	45	26.3%
Sub-total		81	47.4%
Total		171	100%

From the figures in Table 2, in the TPV *Beijing*, the portion of the narrative process is 52.6%, while the conceptual is 47.4%. It indicates that most of the video tries to represent what participants are doing and the rest of the video shows the relation, class, or meaning of the participants. In the narrative process, the action process occupies 29.2%, the reactional process accounts for 8.8%, and circumstance takes up 14.6%. For the conceptual process, the symbolic process takes up the most significant proportion, followed by the analytical process and the classification process.

Frame 1

In frame 1, the process is an action process, with the girl dressed in white as the Actor, the vermilion wall as the Goal, and touching as the process. The girl stretches her arm to touch the vermillion wall, which is the embodiment of Chinese culture, implying the girl feels the culture of Beijing. This action is transactional.

Frame 2

Frame 2 shows a Reactional Process, with the international students as the Reactor and the direction of the students' glance as the vector. As the students are the reactors of the glance, and the direction of their glance forms the vector of the process, this process is non-transactional.

Frame 3

Frame 3 displays a conceptual process, the residence as Superordinate and the blocks of villas and apartments as Subordinate. All the blocks of villas and apartments are arrayed orderly. Based on the function of these buildings, they all belong to the residence, which is the Superordinate in this process, and all the buildings are the subordinates of the residence. Their relation is covert, and the process is a classificational process.

Conclusion

Regarding ideational function, three processes in transitivity system are used in the TPV *Beijing*, with relational process most frequently used followed by mental process and material process. It indicates that the TPV *Beijing* tries to tell audiences the features of Beijing and show what kind of city Beijing is. In terms of representational meaning, both narrative process and conceptual process are used to represent the image of Beijing, with action process and analytical process dominating other processes. It shows the TPV *Beijing* intends to display what people can do in the city and what the city looks like. The narrative process represents the history, culture, and high-quality life of Beijing, while the conceptual process represents the beautiful and splendid scenery of Beijing.

Limitations and Future Studies

Although this has made some progress, there are still some limitations which should be pointed out for other researchers to make their future research in relative field more efficient. Firstly, the perspectives of the current study is limited, for it only analyzes the ideational function and representational meaning of the video. To explore the meaning construction of TPVs, interpersonal and textual function of Systemic Functional Linguistics and interactive and compositional meaning of Visual Grammar should also be considered. Secondly, the data collected for analysis in this study is not enough to explore how verbal and visual modes construct meanings to represent the image of Beijing. More TPVs about Beijing or other tourist destinations should be included to find out the common multimodal features of TPVs.

References

Ansori, M., and Taopan, L. L. (2019) 'a Multimodal Discourse of Promotional Video Wonderful Indonesia', *Elite Journal*, 6(1), pp. 1–18. Available at: http://journal.uin-alauddin.ac.id/index.php/elite/article/view/7875.

Aoki, S., and Fujimoto, T. (2021) 'Examination of Tourism Promotion Using Online Videos', in *Lecture Notes in Networks and Systems*. doi: 10.1007/978-3-030-65796-3_30.

Bower, M., and Hedberg, J. G. (2010) 'A quantitative multimodal discourse analysis of teaching and learning in a web-conferencing environment - The efficacy of student-centred learning designs', *Computers and Education*, 54(2), pp. 462–478. doi: 10.1016/j.compedu.2009.08.030.

Chung, N., and Koo, C. (2015) 'The use of social media in travel information search', in *Telematics and Informatics*. doi: 10.1016/j.tele.2014.08.005.

Francesconi, S. (2014a) 'Reading tourism texts: A multimodal analysis', *Reading Tourism Texts: A Multimodal Analysis*, 50, pp. 1–184. doi: 10.1016/j.annals.2014.11.007.

Francesconi, S. (2014b) 'Reading Tourism Texts: A Multimodal Analysis', *Reading Tourism Texts: A Multimodal Analysis*, 50, pp. 1–184. doi: 10.1016/j.annals.2014.11.007.

Halliday, M. A. K. (1967) Notes on transitivity and theme in English: Part 2. Journal of Linguistics, 3(2): 199-244.

Halliday, M. A. K. (1964). English system networks, (mimeographed.)

Halliday, M. A. K. (1970). Language Structure and Language Function. In Lyons, J. (Ed.). New Horizons in Linguistics. Harmondsworth: Penguin, 140-165.

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: Edward Arnold.

Hitchcock, M., and Dann, G. M. S. (1998) 'The Language of Tourism: A Sociolinguistic Perspective.', *The Journal of the Royal Anthropological Institute*. doi: 10.2307/3034171.

Hong, L. (2012) 'Application of the Multimodal Discourse Analysis Theory to the Teaching of College English Listening and Speaking', *IERI Procedia*, 2, pp. 319–324. doi: 10.1016/j.ieri.2012.06.095.

Houliston, L., Ivanov, S., and Webster, C. (2021) 'Nationalism in offi cial tourism websites of balkan countries: A multimodal analysis', *Tourism*, 69(1), pp. 83–111. doi: 10.37741/T.69.1.7.

Irimiea, S. B. (2019) 'The Relationship Between the Language of Tourism, Tourism and Sociology', *European Journal of Social Science Education and Research*, 5(2), pp. 128–141. doi: 10.2478/ejser-2018-0041.

Isti'anah, A. (2020) '(Re)evaluating language attitudes on indonesian tourism website: A study on ecolinguistics', *Studies in English Language and Education*, 7(2), pp. 622–641. doi: 10.24815/siele.v7i2.16683.

John, W. C. (2014) *Research Design Qualitative, Quantitative, and Mix Method Approaches*.

Kress, G., and Van Leeuwen, T. (1996). Reading Images : The Grammar of Visual Design. London: Routledge.

Kress, G., and Van Leeuwen, T. (2001). Multimodal Discourse: The Mode and Media of Contemporary Communication. London: Arnold.

Kress, G., and Van Leeuwen, T. (2002). Color as a semiotic mode: Notes for a grammar of colour. Visual Communication, 1(3): 343-368.

Kress, G., and Van Leeuwen, T. (2006). Reading Images: The Grammar of Visual Design (2nd edition). London: Routledge.

Lim, V. F. (2019) 'Analysing the teachers' use of gestures in the classroom: A Systemic Functional Multimodal Discourse Analysis approach', *Social Semiotics*, 29(1), pp. 83–111. doi: 10.1080/10350330.2017.1412168.

Lowry, L. L. (2017) 'Language of Tourism', in *The SAGE International Encyclopedia of Travel and Tourism*. doi: 10.4135/9781483368924.n267.

Papakitsou, V. (2020) 'Qualitative Research: Narrative approach in sciences', *Dialogues in Clinical Neuroscience & Mental Health*, 3(1), pp. 63–70. doi: 10.26386/obrela.v3i1.177.

Salim, M. A. Bin, Ibrahim, N. A. B., and Hassan, H. (2012) 'Language for Tourism: A Review of Literature', *Procedia - Social and Behavioral Sciences*. doi: 10.1016/j.sbspro.2012.11.255.

Shin, D., Shin, Cimasko, T., and Yi, Y. (2020) 'Development of metalanguage for multimodal composing: A case study of an L2 writer's design of multimedia texts', *Journal of Second Language Writing*, 47(February), p. 100714. doi: 10.1016/j.jslw.2020.100714.

Thurlow, C., and Jaworski, A. (2011) 'Tourism discourse: Languages and banal globalization', *Applied Linguistics Review*, 2(2011), pp. 285–312. doi: 10.1515/9783110239331.285.